

الجمهورية الجزائرية الديمقراطية الشعبية  
République Algérienne Démocratique et Populaire

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR  
ET DE LA RECHERCHE SCIENTIFIQUE



وزارة التعليم العالي  
والبحث العلمي

UNIVERSITÉ ABOU BEKR BELKAÏD

جامعة أبو بكر بلقايد  
كلية الطب

DEPARTEMENT DE PHARMACIE

MEMOIRE DE FIN D'ETUDES POUR  
L'OBTENTION DU DIPLOME DE DOCTEUR EN PHARMACIE

THÈME :  
**Relation Structure Activité : Etude Qualitative et Quantitative et  
Développement de Recherche sur les Coumarines**

Présenté par :

**Benazzouz Hicham  
Khebiza Ayoub**

*Soutenu le 03/ 07 /2018*

**Le Jury**

**Président :**

**Dr Merzoug Soumia**

**Maitre Assistante à Faculté de Médecin Tlemcen**

**Membres :**

**Dr Cherif Nassima**

**Maitre Assistante à Faculté Médecin Tlemcen**

**Dr Kada kloucha Asma**

**Assistante à Faculté à de Médecine Tlemcen**

**Encadreur :**

**Dr Beghdadi Sara El Mansouria Maitre Assistante à la Faculté de Médecin Tlemcen**

**ANNÉE UNIVERSITAIRE 2017 - 2018**

*Tout d'abord, nous tenons à remercier Dieu, de nous avoir donné la santé, la volonté et la patience pour mener à terme notre formation et pouvoir réaliser ce travail de recherche.*

***A notre encadreur :***

***Dr Beghdadi Sara El Mansouria maitre assistante à la Faculté de Médecine Tlemcen***

*Nous vous remercions d'avoir été à l'initiative de ce travail. Vous nous avez guidés dans l'élaboration de ce mémoire par votre grande disponibilité, vos conseils et votre précieuse aide. Merci de nous avoir fait partager vos connaissances, nous avons été heureux de travailler avec vous et nous tenons à vous en remercier sincèrement.*

***A notre Président du jury :***

***Dr Merzoug Soumia maitre assistante à Faculté de Médecine Tlemcen***

*Nous vous remercions de nous faire l'honneur de présider le jury de ce travail, Soyez assuré de notre profonde reconnaissance.*

***A nos membres de jury :***

***Dr Cherif Nassima maitre assistante à la Faculté de Médecine Tlemcen***

***Dr Kada kloucha Asma assistante à la Faculté de Médecine Tlemcen***

*Nous vous remercions d'avoir accepté d'examiner ce travail et pour l'honneur que vous nous avez fait en participant au jury.*

*Enfin, j'adresse mes remerciements à tous ceux qui ont contribué de près ou de loin à la réalisation de ce travail.*

**À ma mère**

*Tous les mots ne sauraient exprimer la gratitude, l'amour, Le respect et la reconnaissance que J'ai pour vous. Votre amour et votre patience m'ont accompagné chacun de mes pas. J'ai conscience de tous les sacrifices que vous avez dû faire pour me permettre de mener mes études dans les meilleures conditions possibles. Mon admiration pour la femme si courageuse que vous êtes est sans limite.*

***Puisse dieu, le tout puissant, te préserve, te combler de ses grâce infinies, et t'accorder sante, et longue vie et bonheur***

**À mon père**

*Ce travail est le fruit de tes sacrifices que tu as consenti pour mon éducation et ma formation, Je vous remercie pour votre présence et votre soutien tout au long de ma vie dans les meilleurs moments comme dans les mauvais pour devenir ce que je suis, vous étiez ma source d'inspiration.*

***Puisse dieu, le tout puissant, te préserve, te combler de ses grâce infinies, et t'accorder sante, et longue vie et bonheur***

**À mes chers frères : Hicham et Mohamed jihad**

*Pour les bons moments traversés ensemble, merci d'être toujours là pour moi. Je vous dédie ce travail en vous souhaitant un avenir radieux, plein de bonheur et de succès.*

**À ma chère belle-sœur : Mona**

*Merci de m'avoir accompagné pendant toute ma vie, d'être toujours là pour moi et de m'avoir supporté. Je vous aime .je souhaite une vie pleine de bonheur, de prospérité et de réussite.*

**À ma future femme : Linda**

*Pour votre amour, votre soutient et pour tout l'ambiance que dont vous m'avait entouré je vous dédie ce travail en vous souhaitant une vie pleine de joie et de bonheur.*

***À mon binôme, Hicham :*** *Pour toutes ces années passées ensembles, tous les moments vécus ensemble, pour ton amitié et pour ton soutien. Tu auras toujours une place spéciale dans ma vie.*

***À mes très chers ami(e)s :*** *Habib B, Bilel Z, Abdellatif S, Oussama B, Mourad S, Mourad R, Amine S, Mohammed M, Zoubir K. surtout Yacine B et Younes D je vous aime.*

*À mes collègues de la promotion de pharmacie, À tous les gens qui me connaissent, qui m'aiment et ceux qui me souhaitent la réussite et le bonheur dans ma vie. À tous les gens que j'aime sans exception.*

**KHEBIZA AYOUB**

***A mon très cher père***

*Toutes les lettres ne sauraient trouver les mots qu'il faut, tous les mots ne sauraient exprimer ma gratitude, mon amour, ma reconnaissance. Rien au monde ne vaut les efforts fournis jour et nuit pour mon éducation et mon bien être. Ce travail est le fruit de tes sacrifices que tu as consentis pour mon éducation et ma formation. Tes conseils ont toujours guidé mes pas vers la réussite. Qu'Allah, le tout miséricordieux, te préserve, t'accorde santé, bonheur et te protège de tout mal.*

***A ma très chère mère***

*Affable, honorable, aimable : Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement qui n'a pas cessé de m'encourager et de prier pour moi. Ta prière et ta bénédiction m'ont été d'un grand secours pour mener à bien mes études. Aucune dédicace ne saurait être assez éloquente pour exprimer ce que tu mérites pour tous les sacrifices que tu n'as cessé de me donner depuis ma naissance, durant mon enfance et même à l'âge adulte. Tu as fait plus qu'une mère puisse faire pour que ses enfants suivent le bon chemin dans leur vie et leurs études. Je te dédie ce travail en témoignage de mon profond amour. Puisse Dieu, le tout puissant, te préserver et t'accorder santé, longue vie et bonheur.*

***A mes très chers frères Badreddine ; Aymen***

*Merci de m'avoir accompagné pendant toute ma vie, d'être toujours là pour moi et de m'avoir supporté. Je vous aime tous. Qu'Allah le tout puissant, vous protège et vous exhausse tous vos vœux.*

***A ma très chère grand-mère maternelle, mes tantes, mon oncle Mustapha***

*Que ce modeste travail, soit l'expression des vœux que vous n'avez cessé de formuler.*

*Que Dieu vous préserve santé et longue vie.*

***A tous les membres de la famille Benazouz***

***A monsieur Bellout Djilali, et son épouse Dr Zenagui Nawel***

*Merci de votre confiance, Recevez mes plus belles reconnaissance, témoignage de mon profonde gratitude*

***A mon binôme de cette année***

***À mes collègues de la promotion de pharmacie 2012***

***À mes très chers amis***

*Tachema Abir, Zouag Bilal, Dr Ramdaoui Mourad, Dr Baouch, Belhadj Yacine, Dida Younes, Benazouz Said , Kaddour Imad ,Nair Fouad, Bouchiti Fouad et à tous ceux que je n'ai pas mentionnés.*

*Merci d'être à mes côtés dans mes pires et bons moments*

**Hichem**

|  |      |
|--|------|
| REMERCIEMENTS.....   | I    |
| DEDICACES.....   | II   |
| DEDICACES.....   | II   |
| SOMMAIRE.....  | IV   |
| LISTE DES TABLEAUX.....  | VII  |
| EAUX.....  | VII  |
| LISTE DES FIGURES.....   | VIII |
| LISTE DES ABREVIATIONS.....  | IX   |
| INTRODUCTION GENERALE.....   | 1    |
| Chapitre I : Relation structure activité : généralités.....                                      | 3    |
| 1. Généralité.....   | 3    |
| 2. Historique.....   | 3    |
| 3. Définition de SAR.....  | 5    |
| 4. Etude la relation structure activité qualitative ou pharmacophore : .....                     | 6    |
| 5. Etude de la relation structure activité quantitative (QSAR) / structure propriété (QSPR)..... | 6    |
| 5.1. Définition de la relation structure activité quantitative (QSAR).....                       | 6    |
| 5.2. Principe QSAR.....  | 7    |
| 5.2.1. Qualité des données.....  | 7    |
| 5.2.2. Sélection du descripteur.....   | 8    |
| 5.2.3. Méthodes statistiques.....  | 8    |
| 5.3. Technique et méthode.....   | 8    |
| 5.3.1. Paramètre biologique.....   | 8    |
| 5.3.2. Descripteur moléculaire.....  | 9    |
| 5.3.3. Sélection des descripteurs : .....  | 15   |
| 5.4. Stratégie de réduction des descripteurs.....  | 16   |
| 5.4.1. Les techniques d'encapsulation. ....  | 16   |
| 5.4.2. Les techniques de filtrage.....   | 16   |
| 5.5. Méthode statistique.....  | 16   |
| 5.5.1. Régression linéaire simple (SLR).....   | 16   |
| 5.5.2. Régression linéaire multiple (MLR).....   | 17   |
| 5.5.3. Régression linéaire multiple par étapes.....  | 17   |
| 5.5.4. Partiel des moindres carrés (PLS).....  | 17   |
| 5.5.5. Analyse en composantes principales (ACP).....   | 17   |
| 5.5.6. Approximation de la fonction génétique (GFA).....   | 18   |
| 5.5.7. L'analyse par grappes.....  | 18   |

|  |   |    |
|--|---|----|
| 5.5.8.   | Réseaux de neurones artificiels (RNA).....  | 18 |
| 5.5.9.   | K-Plus proche voisin .....  | 19 |
| 5.5.10.  | Classification des méthodes 3D-QSAR.....  | 19 |
| 5.6.   | Validation du modèle .....  | 21 |
| 5.6.1.   | Validation interne .....  | 22 |
| 5.6.2.   | Validation externe.....   | 22 |
| 5.6.3.   | Coefficients et tests statistiques standards :.....                               | 23 |
| 5.7.   | Application de QSAR.....  | 23 |
| Chapitre II : Coumarine : Généralité et étude de l'aspect qualitative de RSA .....                                   |   | 26 |
| 1.   | Généralités : .....   | 26 |
| 2.   | Historique.....   | 26 |
| 3.   | Définition .....  | 27 |
| 3.1.   | Structure générale et propriétés des coumarines.....                              | 28 |
| 3.2.   | Classification structurale .....  | 28 |
| 3.2.1.   | Coumarines simples .....  | 28 |
| 3.2.2.   | Dicoumarines (coumarines dimériques) .....  | 30 |
| 3.2.3.   | Tricoumarines (coumarines trimériques).....                                       | 31 |
| 4.   | Relation structure activité biologique des coumarines et des analogues simples... | 31 |
| 4.1.   | A. Antimicrobien et molluscicide.....   | 31 |
| 4.2.   | Antivirale .....  | 32 |
| 4.3.   | Anticancer.....   | 32 |
| 4.4.   | Inhibition d'enzymes .....  | 33 |
| 4.5.   | Antioxydant .....   | 33 |
| 4.6.   | Anti-inflammatoire.....   | 33 |
| 4.7.   | Anticoagulant et cardiovasculaire .....   | 34 |
| 4.8.   | Divers.....   | 34 |
| Chapitre III : étude de la relation structure activité quantitatives QSAR sur différents dérivés de coumarines ..... |   | 37 |
| 1.   | Coumarines simples : .....  | 37 |
| 1.1.   | Coumarines simple monosubstitués : .....  | 37 |
| 1.1.1.   | Substitution on C <sub>3</sub> : .....  | 37 |
| 1.1.2.   | Substitution on C <sub>4</sub> : .....  | 41 |
| 1.2.   | Coumarine simple bi substituée .....  | 45 |
| 1.2.1.   | Substitution en C3et C4 .....   | 45 |
| 1.2.2.   | Substitution en C4 et C7 .....  | 53 |
| 1.2.3.   | Substitution en C7 et C8 .....  | 59 |

|  |     |
|--|-----|
| 1.3. Coumarine simple tri substituée.....  | 62  |
| 1.3.1. Substitution en C3 et C4 et C6.....   | 62  |
| 1.3.2. Substitution en C3 et C6 et C 7 .....   | 67  |
| 1.4. Coumarine simple poly substituée .....  | 70  |
| 1.4.1. Substitution en C2, C3, C4, C5, C6, C7 et C8 .....  | 71  |
| 1.4.2. Substitution en C2, C3, C4 et C6 .....  | 74  |
| 1.4.3. Substitution en C3, C6, C7 et C8 .....  | 77  |
| 1.4.4. Substitution en C4, C5, C6, C7 et C8.....   | 81  |
| 1.4.5. Substitution en C5, C6, C7, C8 et C9.....   | 85  |
| 2. Bicomarine .....  | 88  |
| 2.1. Etudes 3D-QSAR CoMFA sur la bis-coumarine analogues comme inhibiteurs de l'uréase. ....   | 88  |
| 2.2. Études QSAR de la 3,3 - (benzylidène substitué) -bis 4-hydroxycoumarine comme un inhibiteur potentiel de l'intégrase du VIH-1. .... | 92  |
| 2.3. C. Etude QSAR 3D des dérivés de coumarinyl triazoles mono et bis-aryloxy en tant qu'agents antituberculeux. ....                    | 96  |
| CONCLUSION GENERALE :.....   | 102 |
| REFERENCES BIBLIOGRAPHIQUES .....  | 104 |
| ANNEXES: .....   | 111 |

|  |    |
|--|----|
| Tableau I : Types de données biologiques utilisées dans l'analyse QSAR. ....   | 9  |
| Tableau II : Les descripteurs géométriques calculés dans étude QSAR.....   | 11 |
| Tableau III : Les descripteurs géométriques dans étude QSAR.....   | 12 |
| Tableau IV : Les descripteurs électroniques calculés dans étude QSAR.....  | 13 |
| Tableau V : Les tests standards pour validation de modèle .....  | 23 |
| Tableau VI : Coumarine simple et ses dérivés .....   | 29 |
| Tableau VII : Les descripteurs 2D et 3D calculés.....  | 39 |
| Tableau VIII: PIC50 prévu de l'ensemble de test pour la validation du modèle QSAR .....  | 39 |
| Tableau IX : valeurs pic50 réelles et prévues du modèle qsar basé sur les champs moléculaires .....  | 43 |
| TABLEAU X : Effet inhibiteur de CDK des composés testés .....  | 47 |
| Tableau XI : Valeurs observées par rapport aux valeurs calculées de l'activité antioxydante des dérivés de la coumarine provenant des études QSAR.....   | 51 |
| Tableau XII : Définition des descripteurs significatifs pour le modèle QSAR .....  | 55 |
| TABLEAU XIII : Les résultats calculés partir de modèle 4.....  | 56 |
| Tableau XIV : Activité des composés étudiés dans l'adp (3 mm) ou l'acide arachidonique (50 mm) induite par l'agrégation plaquettaire (cochon d'Inde) et dans l'essai de rétraction des caillots (rat)..... | 61 |
| Tableau XV : Descripteurs utilisés pour développer les modèles RP .....  | 64 |
| Tableau XVI : Activité antituberculeuse prédite par le modèle RP .....   | 65 |
| TABLEAU XVII : Structure, valeurs des paramètres et chimiluminescence pour les composés.....   | 69 |
| Tableau XVIII : Activité cytotoxique des dérivés de coumarines vis à vis les 2 lignes cellulaire .....   | 76 |
| Tableau XIX : Résultats statistiques du modèle CoMFA .....   | 79 |
| TABLEAU XX : Les statistiques obtenues pour les 2 modèles COMFA .....  | 83 |
| Tableau XXI : Activités observées (Obs.), Calculées (Calc.) Et résiduelles (Res.) .....  | 87 |
| TABLEAU XXII : Activités expérimentales et activités prédictives de CoMFA .....  | 90 |
| Tableau XXIII : Résultats de l'analyse CoMFA.....  | 91 |
| Tableau XXIV : Valeurs pIC50 expérimentales et prédites .....  | 95 |
| TABLEAU XXV : Paramètres statistiques des 4 meilleures équations (le nombre de descripteurs ne dépasse pas 2). .....   | 95 |
| TABLEAU XXVI : Paramètres statistiques des 2 modèles COMFA et Topomer COMFA par analyse PLS.....   | 98 |

|   |    |
|---|----|
| Figure 1 : Modèle de l'étude de relation structure activité. [16] .....   | 6  |
| Figure 2 : Surface (enveloppe) de Van Der Waals.....  | 12 |
| Figure 3 : Architecture des réseaux de neurones.....  | 19 |
| Figure 4 : Le champ stérique et électrostatique dans une grille CoMFA [38].....   | 20 |
| Figure 5 : Coumarine, en tant que noyau central pour l'obtention de coumarines simples [54] .....   | 27 |
| Figure 6 : Structure de coumarine et quelques dérivés [59] .....  | 28 |
| Figure 7 : Structure générale de coumarine simple [61].....   | 29 |
| Figure 8 : Structures de quelques furanocoumarines [64].....  | 30 |
| Figure 9 : Structures de quelques Pyranocoumarines [65].....  | 30 |
| Figure 10 : Structures des quelques Dicoumarines [54] .....   | 30 |
| Figure 11 : Structures de quelques des Tricoumarines [54] .....   | 31 |
| Figure 12: Structures d'antibiotiques coumariniques naturels [64].....  | 31 |
| Figure 13 : Structure de la coumarine naturelle geiparvarin. [54] .....   | 32 |
| Figure 14 : Structures des principaux composés warfarine et dicoumarol. [71] .....  | 34 |
| Figure 15 : Stratégie de conception pour la synthèse des composés cibles [72] .....   | 37 |
| Figure 16 : Stratégie de conception des dérivés de la coumarine-pipérazine [73] .....   | 41 |
| Figure 17 : Stratégie de conception pour la synthèse des composés cibles[74] .....  | 45 |
| Figure 18 : Dérivés de coumarine synthétisés. [75].....   | 48 |
| Figure 19 : Structure chimique générale des coumarines 3,7- et 4,7-disubstitués dont QSAR a été étudiée par Apilak Worachartcheewan et al. [76].....  | 53 |
| Figure 20 : Structure principale de 4- (substitué) phényl-3,4-dihydro-1H-chromino [4,3-d] pyrimidine-2,5-diones et 4- (substitué) phényl-2-thioxo-3,4- les noyaux dihydro-1H-chromino [4,3-d] pyrimidin-5-ones et 2-thiopyrimidone [78] ..... | 63 |
| Figure 21 : Structures chimiques de la quercétine et dérivés de la 3 -phénylcoumarine [79] .....  | 67 |
| Figure 22 : La structure chimique générale des dérivés d'alkyl-, acétyl- et nitro-coumarine [80].....   | 71 |
| Figure 23 : Structure générale des dérivés des coumarines 4a-c et leurs analogues phosphoniques 5a-c [81] .....   | 74 |
| Figure 24 : La structure générale des molécules cibles [82].....  | 78 |
| Figure 25 : A structure générale de Calanolide, B structure générale de benzylidène hydrazides de l'acide coumarin-4-acétique.....  | 81 |
| Figure 26 : Structure générale des coumarines avec numérotation arbitraire. [85].....   | 85 |
| Figure 27 : Exigences des analogues de la coumarine pour l'activité anti-méningo encéphalitique [85].....   | 88 |
| Figure 28 : Structure chimique générale de bis-coumarine étudié [87].....   | 89 |
| Figure 29 : Carte de contour de champ électrostatique et stérique CoMFA [87].....   | 91 |
| Figure 30 : Structure générale des dérivés de coumarines étudiées [88].....   | 93 |
| Figure 31 : Structure générale de coumarinyl triazoles mono et bis-aryloxy [89].....  | 97 |

|                           |   |
|---------------------------|---|
| <b>AChE</b>               | <b>Acetylcholinesterase.</b>  |
| <b>ACP</b>                | <b>Analyse en composantes principales.</b>  |
| <b>ADN</b>                | <b>Acide Désoxyribo-Nucléique.</b>  |
| <b>ADP</b>                | <b>Adénosine diphosphate.</b>   |
| <b>AFMoC<br/>ou CoMFA</b> | <b>Adaptation des champs pour la comparaison moléculaire.</b>   |
| <b>AMPc</b>               | <b>Adénosine monophosphate cyclique.</b>  |
| <b>anti-VHB</b>           | <b>Anti Hépatites virales B.</b>  |
| <b>Anti-VIH</b>           | <b>Anti-virus de l'immunodéficience humaine.</b>  |
| <b>Asc</b>                | <b>Acide ascorbique.</b>  |
| <b>BBB</b>                | <b>Barrière hémato-encéphalique (Blood-brain barrier).</b>  |
| <b>BCF</b>                | <b>Facteur de bioconcentration.</b>   |
| <b>BHT</b>                | <b>Butyl-hydroxy-toluène.</b>   |
| <b>CAA</b>                | <b>Connolly Accessible Area.</b>  |
| <b>CAS</b>                | <b>catalytic active site.</b>   |
| <b>CDK</b>                | <b>Kinase cycline-dépendante.</b>   |
| <b>CMA</b>                | <b>Connolly Molecular Area.</b>   |
| <b>CMI</b>                | <b>Concentration minimale d'inhibition.</b>   |
| <b>COMBINE</b>            | <b>Energie de liaison comparative (Comparative Binding Energy)</b>  |
| <b>CoMFA</b>              | <b>Analyse comparative de champs moléculaires (Comparative Molecular Field Analysis).</b>                           |
| <b>CoMMA</b>              | <b>Analyse comparative des moments moléculaires (Comparative Molecular Moment Analysis)</b>                         |
| <b>CoMSIA</b>             | <b>Analyse comparative des indices de similarité moléculaire (comparative molecular similarity index analysis).</b> |
| <b>CoRIA</b>              | <b>Analyse d'interaction de résidus comparatifs (Comparative residue interaction analysis)</b>                      |
| <b>CoSA</b>               | <b>Analyse de spectres comparatifs (comparative spectraanalysis).</b>   |
| <b>CR</b>                 | <b>Résistance croisée (cross resistance).</b>   |
| <b>CSEV</b>               | <b>Connolly Solvent-Excluded Volume.</b>  |
| <b>CV</b>                 | <b>Validation croisée (cross validation)</b>  |
| <b>COX-1<br/>et COX-2</b> | <b>Cyclo-oxygenases.</b>  |
| <b>D</b>                  | <b>Densité</b>  |
| <b>Da</b>                 | <b>Daltons</b>  |
| <b>DPPH</b>               | <b>1,1-diphényl-2-picrylhydrazyle</b>   |
| <b>EHOMO</b>              | <b>Energie HOMO</b>   |

|                              |   |
|------------------------------|---|
| <b>ELUMO</b>                 | <b>Energie LUMO</b>   |
| <b>F</b>                     | <b>Test de Fisher</b>   |
| <b>FRTL-5</b>                | <b>The Fischer rat thyroid cell line</b>  |
| <b>G / PLS<br/>ou GA-PLS</b> | <b>Moindres carrés partiels génétiques</b>  |
| <b>GERM</b>                  | <b>Genetically evolved receptor models</b>  |
| <b>GFA</b>                   | <b>Approximation de la fonction génétique</b>   |
| <b>GI</b>                    | <b>Gastro-intestinaux</b>   |
| <b>HACHe</b>                 | <b>Human acetyl cholinesterase</b>  |
| <b>HQSAR</b>                 | <b>Hologram QSAR</b>  |
| <b>HSV</b>                   | <b>Hue saturation value ou l'herpès simplex</b>   |
| <b>I</b>                     | <b>Induction</b>  |
| <b>IC50</b>                  | <b>Demi-maximal de la concentration d'inhibition (half-maximal inhibitory concentration)</b>                |
| <b>IUPAC</b>                 | <b>International Union of Pure and Applied Chemistry (Union internationale de chimie pure et appliquée)</b> |
| <b>kNN</b>                   | <b>k-Nearest Neighbor</b>   |
| <b>LFER</b>                  | <b>Relation d'énergie libre linéaire</b>  |
| <b>LGO</b>                   | <b>Leave-group-out</b>  |
| <b>LMO</b>                   | <b>Leave-many-Out</b>   |
| <b>Log</b>                   | <b>Logarithme</b>   |
| <b>LOO</b>                   | <b>Leave-one-out</b>  |
| <b>MABA</b>                  | <b>Bleu Alamaren microplaque</b>  |
| <b>MAO B</b>                 | <b>Monoamine oxidase B</b>  |
| <b>MLR</b>                   | <b>Régression linéaire multiple (multiple linear regression)</b>  |
| <b>MOE</b>                   | <b>Environnement d'exploitation moléculaire (Molecular Operating Environment).</b>                          |
| <b>MR</b>                    | <b>Réfraction molaire.</b>  |
| <b>MSE</b>                   | <b>Erreur quadratique moyenne (Mean Squared Error)</b>  |
| <b>MV</b>                    | <b>Volume moléculaire (molecular volume)</b>  |
| <b>MW</b>                    | <b>Poids moléculaire, poids de formule (molecular weight)</b>   |
| <b>NCBI</b>                  | <b>The National Center for Biotechnology Information</b>  |
| <b>NRB</b>                   | <b>Nombre de liaisons rotatives</b>   |
| <b>OMS</b>                   | <b>Organisation mondiale de la santé</b>  |
| <b>PABA</b>                  | <b>Acide p-aminobenzoïque</b>   |
| <b>PAS</b>                   | <b>Peripheralanionic binding site</b>   |
| <b>PC</b>                    | <b>Composants principaux</b>  |
| <b>PCR</b>                   | <b>Régression des composantes principales (Principal components regression)</b>                             |

|                         |   |
|-------------------------|---|
| <b>PLS</b>              | <b>Partiel des moindres carrés (Partial Least Squares)</b>  |
| <b>QSAR<br/>ou RQSA</b> | <b>Relation quantitative structure-activité (quantitative structure-activity relationship)</b>      |
| <b>QSPR</b>             | <b>Relation quantitative structure propriété (quantitative relationship structure property)</b>     |
| <b>RBF</b>              | <b>Radial Basis Function</b>  |
| <b>RNA</b>              | <b>Réseaux de neurones artificiels</b>  |
| <b>RP</b>               | <b>Rartitionnement récursif</b>   |
| <b>RSA<br/>ou SAR</b>   | <b>Relation structure-activité (structure-activity relationship)</b>                                |
| <b>SLR</b>              | <b>Régression linéaire simple (Simple linear regression)</b>  |
| <b>SNC</b>              | <b>Système nerveux central</b>  |
| <b>SSPE</b>             | <b>Somme des carrés des erreurs de prédiction des carrés (the Sum of Squares Prediction Errors)</b> |
| <b>SVDW</b>             | <b>Surface de Van Der Waals</b>   |
| <b>TB</b>               | <b>Tuberculose</b>  |
| <b>TE</b>               | <b>Energie totale</b>   |
| <b>TI</b>               | <b>Indices topologiques (topological indices)</b>   |
| <b>VIH-PR</b>           | <b>Protéase du VIH-1</b>  |
| <b>Vol</b>              | <b>Volume de van der Waals</b>  |
| <b>WHIM</b>             | <b>Moléculaire invariant holistique pondéré (Weighted Holistic Invariant Molecular)</b>             |
| <b>1D ou 1-D</b>        | <b>Descripteurs monodimensionnelles</b>   |
| <b>2D ou 2-D</b>        | <b>Descripteurs Bidimensionnelles</b>   |
| <b>3D ou 3-D</b>        | <b>Descripteurs tridimensionnelles</b>  |
| <b>4D ou 4-D</b>        | <b>Descripteurs quadridimensionnelles</b>   |

**INTRODUCTION**  
**GENERALE**

Dans les dernières décennies, Le climat financier mondial a radicalement changé Avec le ralentissement économique. L'industrie pharmaceutique est dans un état de changement par l'augmentation des coûts et des investissements dans la découverte et le développement de médicaments et par le manque de capacité à choisir des cibles biologiques pertinentes pour la maladie, car le taux d'échec dans les études de preuve de principe ne s'améliore pas. Les autorités de réglementation et les payeurs fixent des obstacles plus élevés à l'approbation de nouveaux médicaments, donc l'industrie pharmaceutique semble incapable de réagir avec tous ces changements elle a ainsi besoin d'autre méthodes alternatives. Parmi elle, les modèles avec une capacité prédictive réelle.

Les progrès récents dans la chimie, L'avènement et le développement rapide de l'information et des technologies informatiques, la croissance de la puissance de calcul parallèlement au développement des logiciels sophistiqué ont permis d'effectuer des expériences *in silico* pour compléter celles *in vitro* et *in vivo*. Aujourd'hui on ne peut pas parler de la conception de médicaments sans mentionner les méthodes computationnelles. L'analyse de la relation structure activité (SAR) c'est l'un des méthodes les plus utilisées.

La relation structure activité (SAR) est une étude qui cherche à comprendre comment les modifications de la structure moléculaire entraînent des changements sur les propriétés moléculaires, dont certaines sont liées aux changements de la puissance biologique. L'étude SAR suit des règles et des étapes rigoureuses pour donner fin à des modèles valides et fiables permettant de synthétiser un très grand nombre de molécules simultanément et de tester leurs actions sur des cibles thérapeutiques et donner des résultats très attirants.

Les coumarines, des substances naturelles dérivant de la benzo- $\alpha$ -pyrone, sont largement distribuées dans le règne végétal. Une diversité structurelle a noté 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, antioxydantes, et sur le système nerveux central.

Notre travail, traitant ce sujet, est divisé en trois chapitres :

- Le premier chapitre intitulé : "La relation structure activité : généralité" est consacré à une étude bibliographique de la RSA : sa définition, son principe, et ses techniques et méthodes.

- Le deuxième chapitre intitulé : " coumarine : généralité et étude de l'aspect qualitative de RSA" Présente un aperçu général sur la définition, leur chimie leur classification générale et structurelle, puis la relation structure activité biologique des coumarines et des analogues simples.
- Le troisième chapitre est consacré à la partie pratique "Etude de la relation structure activité quantitatives QSAR sur différents dérivés de coumarines" nous présenterons, les démarches QSAR déjà effectuées et publiées sur des classes variées de coumarines.

L'objectif de ce travail :

- Concentre sur les études QSAR et leurs rôles importants dans le découvert des nouveaux médicaments actifs.
- L'importance de la famille des coumarines comme une source inépuisable des molécules actives, et sa diversité structurelle qui peut être le début de plusieurs recherches avancées dans le domaine pharmaceutique.

Enfin, nous finirons par une conclusion générale et les perspectives envisagées pour l'étude RSA.

**Chapitre I :**  
**Relation structure activité :**  
**Généralités**

## 1. Généralité

L'un des défis les plus importants auxquels font face les chimistes médicaux aujourd'hui est la conception de nouveaux médicaments ayant des propriétés améliorées et des effets secondaires diminués. Les chimistes ont commencé le processus en prenant une structure principale, puis en trouvant des analogues présentant les activités biologiques préférées. Ensuite, ils ont utilisé leur expérience et leurs connaissances sur les produits chimiques pour finalement choisir un analogue désigné pour un développement ultérieur. Ce processus est difficile, coûteux et long. [1]

La conception des médicaments est un processus conduit par des percées technologiques impliquant des méthodes expérimentales et informatiques avancées. De nos jours, les techniques de conception des médicaments sont d'une importance primordiale pour la prédiction du profil biologique, l'identification des hits, la génération de pistes et de plus, pour accélérer l'optimisation des prospects en candidats-médicaments. [2]

La relation structure-activité (SAR) est un domaine important pour la conception et la découverte des médicaments ainsi qu'à la recherche des informations relatives à l'influence de la structure chimique sur les activités pharmaceutiques. [3]

## 2. Historique

En 1868, Crum-Brown et Fraser ont publié une équation qui est considérée comme la première formulation générale d'une relation structure-activité. Dans leurs recherches sur différents alcaloïdes, ils ont reconnu que l'alkylation des atomes d'azote basiques produisait des effets biologiques significativement différents des composés d'ammonium quaternaire chargés en permanence, en comparaison avec les amines basiques. Ils ont donc supposé que "l'activité physiologique" doit être fonction de la structure chimique (équation 1). [4]

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

En 1893, Richet a découvert que la toxicité des composés organiques suit leur solubilité dans l'eau. Une telle relation correspond à (équation 2).

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

$\Delta\emptyset$  : les différences dans les valeurs d'activité biologique, causées par des changements correspondants dans les propriétés chimiques et surtout physicochimiques :  $\Delta C$ .

Au tournant du siècle dernier, Meyer et Overton ont indépendamment observé des relations linéaires entre la lipophilie, exprimée en coefficients de partage huile-eau, et les activités narcotiques. [5]

La méthodologie Relation Quantitative Structure Activité (QSAR) s'est rapidement développée à partir du milieu des années 1950 : Bruice, Kharasch et Winzler ont formulé des contributions de groupe aux valeurs d'activité biologique dans une série d'analogues d'hormones thyroïdiennes, qui peuvent être considérées comme une 1<sup>ère</sup> analyse de type Wilson libre. [5]

En 1962, Hansen a dérivé la première (pendant longtemps la seule) relation réelle de type Hammett entre les toxicités des acides benzoïques substitués et les constantes électroniques de leurs substituants. Dans la même année, une première publication QSAR de Corwin Hansch sur « The corrélation of the biological activity of phenoxy acetic acids with Hammett substituent constants and partition coefficients » [6]

1964 peut être considérée comme l'année de naissance de la méthodologie RQSA moderne. Indépendamment, deux articles ont été publiés, un par Hansch et Fujita sur "ρ-σ-π" Analysis. C'est une méthode pour la corrélation de l'activité biologique et la structure chimique, l'autre par Free et Wilson sur " Une contribution mathématique aux études d'activité de structure ". Les contributions ont amorcé le développement de deux nouvelles méthodes de relations quantitatives structure-activité, plus tard appelées analyse de Hansch (approche linéaire liée à l'énergie libre, approche extra thermodynamique) et l'analyse Free Wilson, respectivement [7].

Au début des années 1980, Klopman a développé une approche permettant de fractionner une molécule en fragments 2D constitutifs, d'auto-générer ces fragments pour un grand nombre (de centaines à des milliers) de molécules dans un ensemble d'entraînement et de tenter de corréler la fréquence de chacun de ces fragments. [8]

À mesure que les ensembles de données devenaient plus grands et plus structurellement diversifiés, les descripteurs conçus pour être appliqués dans un cadre de mécanisme de réaction commun ne suffisaient plus [8].

Le travail de Rosenkranz et Klopman pour étendre l'application de CASE (maintenant CASETOX), et plus tard de Multi-CASE, a une grande diversité de défis de modélisation de la toxicité et illustre la puissance de l'approche basée sur la sous-structure pour modéliser des activités non examinées précédemment et élucider la base structurelle de l'activité [9].

Le domaine a également progressé dans le développement des méthodes objectives pour évaluer la fiabilité des modèles et la confiance dans les prévisions. Des procédures statistiques ont été adoptées tôt pour éviter les corrélations de hasard, par exemple, en utilisant des simulations de variables aléatoires, comme expliqué dans l'article fondateur par Topliss et Edwards. [10]

La chimie quantique reste un outil puissant pour explorer les déterminants fondamentaux de la réactivité dans QSAR. Elle est actuellement avec ses implémentations de mécanique semi-moléculaire et moléculaire utilisée dans des approches QSAR 3D, telles que le criblage et le profilage de ligands virtuels in silico dans la découverte de médicaments. En outre, ces méthodes peuvent être utilisées pour examiner la stabilité relative des conformères, ce qui peut à son tour influencer les propriétés dépendant du 3D utilisé dans QSAR. [11]

En plus de 50 ans de développement actif, le domaine de la modélisation QSAR s'est énormément développé par rapport à la diversité des méthodologies et des applications.[11]

### 3. Définition de SAR

Les relations structure-activité (SAR) représentent le lien entre la structure moléculaire et l'activité biologique ou les propriétés physicochimiques des composés qui sont souvent classés ensemble, car ils ont en commun des caractéristiques structurales telles que la forme, la taille, la disposition stéréochimique et la répartition des groupes fonctionnels. La structure moléculaire et l'activité biologique sont corrélées en observant les résultats de la modification structurelle systématique sur des paramètres biologiques définis. [12]

Les relations structure-activité (SAR) sont les pratiques traditionnelles de la chimie médicinale qui tentent de modifier l'effet (c'est-à-dire l'activité) ou la puissance de composés chimiques bioactifs en modifiant leur structure chimique. Les chimistes utilisent des techniques de synthèse chimique pour insérer des nouveaux groupes fonctionnels dans les composés biomédicaux. Cette méthode a été ensuite affinée pour construire des relations mathématiques entre une structure chimique et son activité biologique. [3]

L'hypothèse de base de la SAR est que les molécules similaires ont des activités similaires (l'approche "read-across" voir annexe 01). Le paradoxe fait référence au fait que ce n'est pas le cas que toutes les molécules similaires ont des activités similaires, puisque la similarité est un terme subjectif, et qu'il existe une grande diversité dans le type et l'étendue des similarités.[13]

Le problème sous-jacent est donc de savoir comment définir une petite différence au niveau moléculaire, puisque chaque type d'activité, par ex. La capacité de réaction, la capacité de biotransformation, la solubilité, l'activité cible, etc., peut dépendre de cette différence. [13]

Pour uniformiser la signification du terme « similarité », des relations mathématiques entre une structure chimique et son activité biologique ont été établies, elles sont connues sous le nom de relation quantitative structure activité (QSAR) [12].

#### 4. Etude la relation structure activité qualitative ou pharmacophore :

Les approches pharmacophores sont devenues l'un des principaux outils de la découverte de médicaments. Selon la définition de l'UPAC, un modèle pharmacophore est « un ensemble de caractéristiques stériques et électroniques nécessaires pour assurer les interactions supramoléculaires optimales entre les médicaments, et une cible biologique spécifique pour déclencher (ou bloquer) sa réponse biologique » [10].

Le criblage virtuel est une méthode de calcul précieuse aux côtés du criblage traditionnel à haut débit. C'est une technique pour la découverte de nouveaux composés actifs dans l'industrie pharmaceutique, le succès d'un dépistage virtuel est défini comme la découverte de nouveaux échafaudages intéressants à partir de nombreuses têtes de série [7]. L'amarrage moléculaire est un outil utile pour étudier l'interaction d'un récepteur protéique avec son ligand et démontrer le mécanisme de liaison comme démontré par une série d'études et une revue complète [11].

#### 5. Etude de la relation structure activité quantitative (QSAR) / structure propriété (QSPR)

##### 5.1. Définition de la relation structure activité quantitative (QSAR)

La relation quantitative structure activité (QSAR) est le processus par lequel la structure chimique est corrélée quantitativement avec un processus bien défini, tel que l'activité biologique ou la réactivité chimique.[14]

L'activité biologique peut être exprimée quantitativement par la concentration d'une substance requise pour donner une certaine réponse biologique. L'expression mathématique peut être alors utilisée pour prédire la réponse biologique d'autres structures chimiques [14].

La réactivité chimique peut être exprimée par des propriétés telles que la lipophilicite, la solubilité et la perméabilité.[15]

La forme mathématique (figure 1) la plus générale de QSAR est : [16]

$$\text{Activité} = f(x)$$

X : propriétés physico-chimiques et / ou propriétés structurales

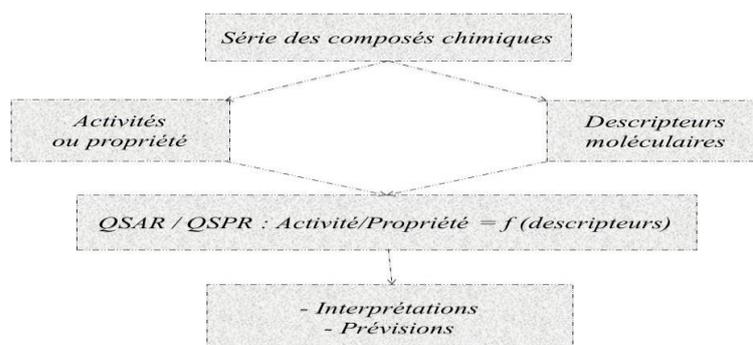


FIGURE 1 : Modèle de l'étude de relation structure activité. [16]

L'hypothèse fondamentale de QSAR est que les variations de l'activité biologique d'une série de substances chimiques qui ciblent un mécanisme d'action commun sont corrélées avec les variations de leurs propriétés structurales, physiques et chimiques. Ces propriétés peuvent être déterminées par des moyens expérimentaux ou informatiques beaucoup plus efficacement que son activité biologique en utilisant des approches in vitro ou in vivo. [17]

Un modèle QSAR statistiquement validé est capable de prédire l'activité biologique d'un nouveau produit chimique dans la même série à la place des processus de synthèse chimique et d'évaluation biologique qui demandent beaucoup plus de temps et de main-d'œuvre. Appliquée judicieusement, QSAR peut économiser du temps, d'argent et de ressources humaines. [8]

Les QSAR sont actuellement appliquées dans de nombreuses disciplines, dont beaucoup concernent la conception de médicaments et l'évaluation des risques environnementaux.

### 5.2. Principe QSAR

La modélisation de la relation structure-activité quantitative comporte généralement trois étapes : [18]

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

L'obtention d'un modèle QSAR de bonne qualité avec la capacité de prédire l'activité d'un produit chimique dépend de nombreux facteurs dans l'approche et l'exécution de chacune des trois étapes.

#### 5.2.1. Qualité des données

Les données devraient provenir du même protocole d'analyse, et il faudrait veiller à éviter la variabilité inter-laboratoire. Tout mauvais point de données aura tendance à corrompre la corrélation correcte de la structure et de l'activité. [19]

Les règles empiriques pour un bon ensemble de données QSAR sont que la relation dose-réponse et l'activité (ou l'affinité) devrait être reproductible, la gamme d'activités devrait s'étendre de deux ordres de grandeur ou plus du produit chimique le moins actif au plus actif de la série, le nombre de produits chimiques utilisés pour construire le modèle QSAR devrait être suffisamment important pour la stabilité, les activités des produits chimiques devraient être réparties uniformément dans toute la gamme d'activités, et les produits chimiques choisis pour

l'ensemble d'entraînement devraient posséder suffisamment de diversité structurelle pour couvrir l'éventail de l'espace de chimie associé à l'activité biologique étudiée . [20]

### 5.2.2. Sélection du descripteur

De nombreux types de descripteurs de structure chimique sont disponibles auprès de logiciels commerciaux. L'obtention d'un modèle statistiquement robuste dépend beaucoup de la capacité des descripteurs sélectionnés à encoder la variation d'activité avec la structure. Plus on en sait au niveau moléculaire sur le mécanisme d'action biologique des produits chimiques, mieux le chimiste peut choisir parmi la grande variété et les types de descripteurs moléculaires spécifiques. [21]

Les programmes de modélisation moléculaire disponibles dans le commerce comprennent souvent des outils statistiques pour aider à évaluer quels descripteurs codent le mieux pour la variation structure-activité. [22]

### 5.2.3. Méthodes statistiques

Faire de la statistique suppose que l'on étudie un ensemble d'objets équivalents sur lesquels on observe des caractéristiques appelées « variables ». Dans notre cas, les objets (ou individus) sont les molécules et les variables sont les descripteurs moléculaires. Après le recueil des descripteurs, la démarche statistique consiste à traiter et interpréter les informations recueillies sur ces molécules (voir plus loin : les méthodes statistiques). [18]

## 5.3. Technique et méthode

### 5.3.1. Paramètre biologique

Les modèles QSAR sont dépendants des données expérimentales utilisées pour leur construction. Ces données devraient, idéalement, être de grande qualité, ce qui signifie qu'elles devraient être fiables et cohérentes. Les données biologiques sont habituellement exprimées sur une échelle logarithmique.

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, utilisées dans l'analyse de QSAR, sont décrits dans le tableau I qui représente les types des données biologiques utilisées dans l'analyse QSAR : [23]

TABLEAU I : Types de données biologiques utilisées dans l'analyse QSAR. [23]

| Source d'activité                             | Paramètres biologiques     |
|---|----------------------------|
| <b>1. Récepteurs isolés</b>                   |                            |
| <b>Constante de vitesse</b>                   | Log $k$                    |
| <b>Constante de Michaelis-Menten</b>          | Log $1/Km$                 |
| <b>Constante d'inhibition</b>                 | Log $1/Ki$                 |
| <b>2. Systèmes cellulaires</b>                |                            |
| <b>Constante d'inhibition</b>                 | Log $1/IC50$               |
| <b>Résistance croisée</b>                     | Log $CR$                   |
| <b>Données biologiques <i>in vitro</i></b>    | Log $1/C$                  |
| <b>Mutation de gène</b>                       | Log $TA98$                 |
| <b>3. Systèmes <i>in vivo</i></b>             |                            |
| <b>Facteur de bioconcentration</b>            | Log $BCF$                  |
| <b>Vitesses de la réaction <i>in vivo</i></b> | Log $I$ (induction)        |
| <b>Vitesses pharmacodynamiques</b>            | Log $T$ (clairance totale) |

### 5.3.2. Descripteur moléculaire

#### 5.3.2.1. Définition

Les descripteurs moléculaires sont des représentations mathématiques formelles d'une molécule, obtenues par un algorithme bien défini et appliquées à une représentation moléculaire définie ou à une procédure expérimentale bien définie : le descripteur moléculaire est le résultat final d'une procédure logique et mathématique qui transforme l'information chimique codée dans une représentation symbolique d'une molécule en un nombre utile ou le résultat d'une expérience standardisée. [24].

L'information codée par les descripteurs dépend généralement du type de représentation moléculaire et de l'algorithme défini pour son calcul.

La classification des descripteurs moléculaires est établie en fonction :

- Leur origine (constitutionnel, géométrique, topologique, thermodynamique, quantique).
- Leur dimensionnalité (1D, 2D, 3D ou 4D).

#### 5.3.2.2. Les descripteurs 1-D

L'activité est corrélée avec les propriétés moléculaires globales (constitutionnelles). Ce sont des descripteurs simples et couramment utilisés qui reflètent la composition moléculaire d'un composé sans aucune information sur sa topologie.

Les descripteurs de constitution les plus courants sont le nombre d'atomes, le nombre de liaisons, le type d'atome, le nombre de cycles et le poids moléculaire (MW) [25].

- **Le poids moléculaire** : noté MW (appelé aussi le poids de formule), mesuré en daltons (Da). C'est la somme des poids atomiques des différents atomes constituant la molécule. Il est utilisé dans l'étude de transport dont la diffusion et le mode de fonctionnement des médicaments.
- **Le pourcentage massique** : indique le pourcentage de chaque élément contenu dans un composé chimique défini par la formule suivante :

$$\text{Pourcentage massique} = (\text{masse de l'élément}/\text{masse totale du composé}) \times 100$$

Ce type de descripteur est employé pour obtenir les modèles QSAR les plus simples, mais dans certains cas ils sont défaillants car ils ne permettent pas la proposition de mécanismes d'interactions mises en jeu pour l'activité étudiée. Par exemple, ils ne permettent pas de distinguer les isomères. Or, pour la majorité des cas, la position d'un substituant modifie la valeur expérimentale de l'activité. [26]

### 5.3.2.3. Les descripteurs 2-D

Les descripteurs 2D sont des descripteurs ou des indices topologiques (TI) qui prennent en compte l'arrangement atomique interne des composés. Ils sont dérivés de la représentation topologique des molécules et peuvent être considérés comme des descripteurs explicitement structurés. Ces indices codent pour des informations sur la taille moléculaire, la forme, la ramification, la présence d'hétéroatomes et de multiples liaisons sous forme numérique, Une représentation topologique de la molécule est présentée sous la forme d'un graphe moléculaire. [27]

Les indices topologiques classiques peuvent être regroupés en deux ensembles selon la matrice dont ils proviennent. Le premier groupe est formé par des indices liés à la matrice d'adjacence du graphe moléculaire. Ces indices sont les indices M1 et M2 du groupe de Zagreb définis sur la base des degrés de vertex, qui sont les sommes des lignes ou des colonnes de la matrice A (2,3) : [28]

$$M_1 = \sum_i (\delta_i)^2 \quad (1)$$

$$M_2 = \sum_k (\delta_i \cdot \delta_j)_k \quad (2)$$

Dans M2, la sommation est effectuée sur tous les sommets (arêtes) adjacents du graphe. Intimement lié à ces indices est l'indice de connectivité de Randić, qui est défini de la même manière que M2 mais en utilisant un exposant -0.5 dans l'invariant :

$$\chi = \sum_k (\delta_i \cdot \delta_j)_k^{-0.5} \quad (3)$$

L'autre groupe de descripteurs est liés à la matrice de distance du graphe moléculaire. Voir le tableau II qui représente les descripteurs géométriques calculés dans l'étude QSAR

Les descripteurs 2-D sont employés pour obtenir des modèles QSAR plus simples, mais leur défaillance, comme pour les descripteurs 1-D, est qu'ils ne permettent pas la bonne interprétation des mécanismes d'interaction mis en jeu pour l'activité/propriété étudiée.

**TABLEAU II : Les descripteurs géométriques calculés dans étude QSAR [28]**

| Descripteur                                     | Définition   |
|---|--|
|   | $\omega = \frac{1}{2} \sum_{ij} d_{ij} = \frac{1}{2} \sum_i s_i$   |
| <b>L'indice de Wiener (<math>\omega</math>)</b> | Est simplement la demi-somme de tous les éléments d'une telle matrice où $S_i$ est la somme des distances                                      |
|   | $J = C \cdot \sum_k (s_i \cdot s_j)^{-0.5}$  |
| <b>L'index J de Balaban (J)</b>                 | Où $C = m(1 + u)$ ; $m$ est le nombre d'arêtes et $u$ est le nombre cyclomatique.  |
| <b>Index de connectivité (<math>X_R</math>)</b> | $X_R \equiv X = \sum_{i=1}^{A-1} \sum_{i+1}^A a_{ij} \cdot (\delta_i \cdot \delta_j)^{-1/2}$   |
| <b>La matrice de distance (D)</b>               | $[D]_{ij} = \begin{cases} d_{ij} = \min & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} P_{ij}$                                      |
| <b>La matrice d'adjacence (A)</b>               | $[A]_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E(G) \\ 0 & \text{otherwise} \end{cases}$<br>Où $E(G)$ est l'ensemble des bords du graphe. |

#### 5.3.2.4. Les descripteurs 3-D

Les relations structure activité quantitatives en trois dimensions (3D QSAR) sont des modèles qui établissent une relation entre une activité biologique et les paramètres structuraux (descripteurs moléculaires) calculés dans l'espace tridimensionnel pour un groupe de molécule. [4]

On distingue plusieurs familles de descripteurs 3D :

▪ **Descripteurs géométriques**

Sont issus de la structure tridimensionnelle de la molécule, et sont évalués à partir des positions relatives des atomes d'une molécule dans l'espace, ainsi que des rayons et des masses atomiques. En général, les descripteurs géométriques sont calculés soit à partir de la géométrie moléculaire optimisée par modélisation moléculaire soit à partir des coordonnées cristallographiques (coordonnées cartésiennes x, y, z des atomes de la molécule). [29]

La plupart de ces descripteurs (descripteurs basés sur la grille) ont besoin de règles d'alignement pour réaliser la comparabilité des molécules. Plusieurs classes de descripteurs peuvent être distinguées dans l'ensemble des descripteurs géométriques. [28]

- La figure 02 ci-après montre une représentation de la surface de Vander Waals

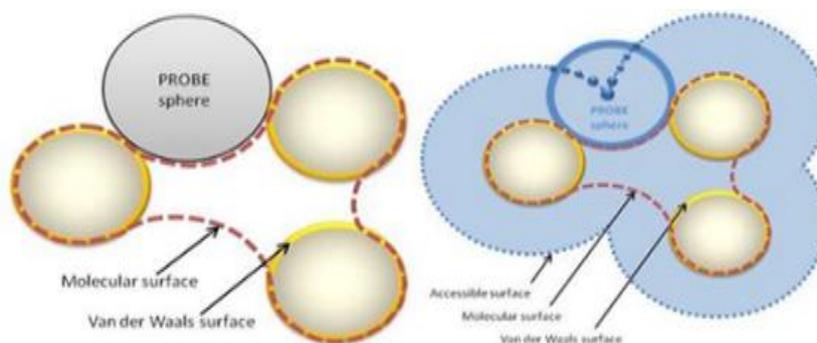


FIGURE 2 : Surface (enveloppe) de VAN DER WAALS

TABEAU III : Les descripteurs géométriques dans étude QSAR [28]

| Descripteur   | Définition  |
|---|---|
| <b>Le volume moléculaire (MV)</b>                                       | $MV = MW/d$   |
| <b>Le nombre de liaisons rotatives (NROT)</b>                           | La liaison d'un composé non cyclique, associée à un atome non lourd (qui n'est pas l'hydrogène) selon la règle de Lipinski, il faut que le nombre de liaisons rotatives soit inférieur ou égal à 5. |
| <b>La surface de Van Der Waals (SVDW) (Figure 2)</b>                    | Il représente une surface à travers laquelle la molécule peut interagir avec d'autres molécules, Plus elle est grande plus la possibilité d'interaction est importante.                             |
| <b>Index de la matrice de géométrie (G)</b>                             | $G \equiv \begin{vmatrix} 0 & r_{12} & \dots & r_{1A} \\ r_{21} & 0 & \dots & r_{2A} \\ \dots & \dots & \dots & \dots \\ r_{A1} & r_{A2} & \dots & 0 \end{vmatrix}$                                 |
| <b>Degré de distance géométrique (G<math>\delta</math><sub>i</sub>)</b> | $G_{\delta_i} = \sum_{j=1}^A r_{ij}$  |

▪ **Descripteurs électroniques**

Sont d'une importance critique dans la détermination des types de forces intermoléculaires qui modulent les interactions des récepteurs médicamenteux, ils décrivent l'influence d'un certain groupe ou d'un substituant sur la distribution de la densité électronique (polarité des molécules). [4]

Les descripteurs les plus importants sont montrés dans le tableau IV qui représente les descripteurs électroniques calculés dans l'étude QSAR :

**TABLEAU IV : Les descripteurs électroniques calculés dans l'étude QSAR [30]**

| Descripteur                              | Définition  |
|--|---|
| <b>L'énergie HOMO (E<sub>HOMO</sub>)</b> | Est le niveau d'énergie le plus élevé dans la molécule qui contient des électrons   |
| <b>L'énergie LUMO (E<sub>LUMO</sub>)</b> | Est le niveau d'énergie le plus bas dans la molécule qui ne contient pas d'électrons.                                       |
| <b>Le moment dipolaire (μ)</b>           | Mesure la polarité nette moléculaire, qui résulte S'il y a une différence d'électronégativité entre les atomes de molécule. |
| <b>L'électronégativité (X)</b>           | $\chi = -\mu = -\left(\frac{\partial E}{\partial N}\right)_{v(r)} = -\frac{(E_{LUMO} + E_{HOMO})}{2}$                       |

▪ **Descripteurs physicochimiques**

Ce sont les propriétés physiques et chimiques d'une molécule qui peuvent être estimées par l'examen de sa structure. Ces propriétés ont un rôle majeur sur l'efficacité des médicaments, le profil de sécurité, et dans la découverte de médicaments en optimisant les composés sélectionnés.

➤ **Solubilité**

Une solubilité aqueuse adéquate est d'une importance primordiale puisque la dissolution du médicament actif dans les fluides gastro-intestinaux (GI) précède son absorption par voie orale du tractus gastro-intestinal. La solubilité aqueuse, à son tour, dépend de plusieurs facteurs tels que la taille et la forme de la molécule, l'hydrophobicité, la liaison hydrogène et l'état cristallin / amorphe. [31]

### ➤ **Perméabilité**

Deux principaux types de perméabilité, à savoir la perméabilité intestinale humaine et la perméabilité de la barrière hémato-encéphalique (BBB) (importante pour la distribution des agents actifs du SNC et la toxicité des médicaments non actif sur SNC), sont importantes pour le profil biologique. En utilisant les descripteurs log P, pKa et PSA.

En ce qui concerne l'étude des médicaments du système nerveux central, un grand ensemble de données de 1593 molécules ainsi qu'un ensemble de 19 descripteurs moléculaires simples ont été utilisés pour construire des modèles de prévision BBB. Les propriétés de liaison hydrogène des molécules ont été trouvées pour moduler la pénétration BBB. [32]

### ➤ **pKa**

La prédiction du pKa à partir de la structure moléculaire est un domaine de recherche intense, où l'état ionique de la molécule de médicament au pH physiologique, représenté par la constante d'ionisation, ou pKa, peut potentiellement affecter son comportement pharmacocinétique. Le coefficient de distribution log D dépendant du pH (à pH 6,5, 7,4) dépend principalement du pKa car les médicaments sont soumis à un environnement variable (pH 1-3 dans l'estomac, pH 5-7 dans le duodénum, pH 8 dans le jéjunum et l'iléon) passage dans le GIT. [27]

### ➤ **Lipophilicite**

La lipophilie, qui fait référence à la capacité d'un composé à se dissoudre dans les graisses, les huiles, les lipides et les solvants non polaires tels que l'hexane ou le toluène, est reconnue depuis longtemps comme un facteur important du passage du médicament au développement clinique.

Un log P calculé (Clog P) est couramment utilisé comme une évaluation de la lipophilie in vivo, ce qui reflète l'événement clé de la désolvatation moléculaire dans le transfert de phases aqueuses aux membranes cellulaires et aux sites de liaison aux protéines. Compte tenu de la valeur et de l'influence de la lipophilie sur le processus de découverte et de conception des médicaments et de son rôle prédictif de succès éventuel des composés, la compréhension de la lipophilie et sa modulation sont essentielles au développement d'un composé thérapeutique efficace [33]. Les descripteurs les plus utilisés sont : [4]

### - **Le coefficient de partage Octanol/Eau :**

Il est défini par la formule suivante :

$$\log P = \log \frac{[\text{Octanol}]}{[\text{H}_2\text{O}]}$$

[Octanol] : concentration du soluté dans l'Octanol

[H<sub>2</sub>O] : concentration du soluté dans l'eau.

Selon la règle de Lipinski (voir annexe 02), les médicaments délivrés par voie orale doivent avoir des valeurs de log P supérieures ou égales à -2 et inférieures ou égales à 5.

Les composés qui ont les valeurs de log P > 0 sont dites lipophiles, et les composés qui ont les valeurs de log P < 0 sont dites hydrophiles.

- **Le coefficient de distribution**

Il est défini par la formule suivante :

$$\text{Log D} = \log P - \log 1 + 10\text{pKa} - \text{pH}$$

- **Le nombre de donneurs de liaisons hydrogène :**

Selon la règle de Lipinski, le nombre de donneurs de liaisons hydrogène doit être inférieur ou égal à 5. (Voir annexe 02)

- **Le nombre d'accepteurs de liaisons hydrogène :**

Selon la règle de Lipinski, le nombre d'accepteurs de liaisons hydrogène doit être inférieur ou égal à 10.

**5.3.3. Sélection des descripteurs :**

Généralement seulement un petit sous-ensemble des descripteurs calculés porte les informations nécessaires pour construire un modèle mathématique du système d'intérêt. Il est possible de construire des modèles incluant tous les descripteurs, mais il peut y avoir de nombreuses raisons de ne sélectionner qu'un sous-ensemble d'entre eux tels que :

- La précision de prédiction du modèle pourrait être améliorée par l'exclusion de la redondance et les descripteurs inutiles.
- Le modèle QSAR à construire est souvent plus simple et potentiellement plus rapide lorsque moins de descripteurs d'entrée sont appliqués, alors l'interprétabilité de la relation entre les descripteurs et l'activité observée pourrait être augmentée.

La sélection de descripteurs vise à se débarrasser des descripteurs calculés qui sont redondants, bruyants ou non pertinents pour les tâches de construction de modèles envisagées, sans perte d'une information importante. [25]

Pour QSAR, il existe deux grandes stratégies qui peuvent être utilisées pour la réduction des descripteurs, des méthodes d'encapsulation et des méthodes de filtrage.

### 5.4. Stratégie de réduction des descripteurs

#### 5.4.1. Les techniques d'encapsulation.

Une approche d'encapsulation consiste essentiellement en deux composantes : la fonction objective, qui peut être un modèle de régression linéaire ou non linéaire, et une méthode d'optimisation (sélection) pour sélectionner les descripteurs de fonction objective. Exemples de la composante d'optimisation inclure des algorithmes génétiques. La performance de l'algorithme de régression est utilisée pour guider la procédure d'optimisation dans la sélection des descripteurs. En conséquence, la procédure de sélection est étroitement liée à l'algorithme de régression utilisé.

#### 5.4.2. Les techniques de filtrage

La technique n'utilise aucun algorithme de régression spécifique pour sélectionner des sous-ensembles de descripteurs. Au lieu de cela, il considère uniquement les caractéristiques des descripteurs calculés pour effectuer la sélection. La procédure standard de réduction des descripteurs, dans laquelle les descripteurs à faible variance corrélés sont retirés d'un grand ensemble initial de descripteurs calculés pour donner un plus petit pool, est un exemple d'une technique de sélection de descripteur filter type.

Les méthodes statistiques sont aussi utilisées pour l'élimination des descripteurs qui n'interviennent pas dans les modèles proposés. [22]

### 5.5. Méthode statistique

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. Les variantes inclusives peuvent être SLR, MLR et MLR par étapes. Une brève explication de ces variantes est la suivante [28] :

#### 5.5.1. Régression linéaire simple (SLR)

Cette méthode effectue comme un calcul de régression linéaire standard dans la génération du modèle QSAR sous la forme d'équations qui incluent un seul descripteur indépendant  $x$  et  $y$  en tant 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 :

$$y = a + bx \quad (1)$$

Où y est la variable dépendante, x est la variable indépendante, a est la constante b, est le coefficient de régression (2).

### 5.5.2. Régression linéaire multiple (MLR)

Cette méthode est l'extension de SLR à plus d'une dimension. Dans cette méthode, des calculs de régression multivariées 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<sup>2</sup> ou q<sup>2</sup> 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 [34] :

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m + e. \quad (2)$$

### 5.5.3. Régression linéaire multiple par étapes

Cette méthode utilise couramment la variante MLR qui crée une équation linéaire à plusieurs termes, mais toutes les variables indépendantes ne sont pas utilisées. Cette méthode a une bonne utilisation lorsque le nombre de descripteurs est important et que les descripteurs principaux sont inconnus. De plus, les variables latentes orthogonales peuvent être utilisées dans MLR. [2]

### 5.5.4. Partiel des moindres carrés (PLS)

Le PLS donne une solution statistiquement robuste même lorsque les variables indépendantes sont fortement interreliées entre elles, ou lorsque les variables indépendantes dépassent le nombre d'observations. PLS est une méthode de régression itérative qui produit ses solutions basées sur la transformation linéaire d'un grand nombre de descripteurs originaux en un petit nombre de nouveaux termes orthogonaux appelés variables latentes. Ainsi, cette méthode est comptée comme statistique standard. [2]

### 5.5.5. Analyse en composantes principales (ACP)

Cette méthode est connue pour créer un nouvel ensemble de descripteurs orthogonaux appelés composants principaux (PC) qui décrivent la plupart des informations contenues dans les variables indépendantes dans l'ordre de la variance décroissante. Dans cette méthode, le

modèle QSAR n'est pas généré mais il témoigne encore de la relation entre des variables différentes. PCA réduit la dimensionnalité de l'ensemble de données des descripteurs à la quantité réelle de données. Pour générer une équation linéaire à plusieurs termes, la PCR applique les scores de la décomposition de l'ACP en tant que régresseurs dans le modèle QSAR. [35]

### **5.5.6. Approximation de la fonction génétique (GFA)**

Cette méthode sert d'alternative à l'analyse de régression standard pour construire des équations QSAR. Elle peut construire des équations linéaires ainsi que des équations non linéaires d'ordre supérieur. Les moindres carrés partiels génétiques (G / PLS ou GA-PLS) sont un outil important qui a évolué en combinant la meilleure caractéristique de GFA et PLS. Cette méthode est largement utilisée par les chercheurs / scientifiques du monde entier [35].

### **5.5.7. L'analyse par grappes**

Cette méthode d'analyse est une méthode de reconnaissance de formes utilisée pour étudier la relation entre les observations associées à de nombreuses autres propriétés et pour partitionner l'ensemble de données en catégories en éléments similaires. Cela implique également d'identifier lesquels des sous-ensembles partagent des propriétés physiques similaires [35].

### **5.5.8. Réseaux de neurones artificiels (RNA)**

Les RNA sont des méthodes utiles dans les études QSAR / QSPR, et en particulier dans les cas où il est difficile de spécifier un modèle mathématique exact pour décrire une relation de propriété de structure donnée. [36]

La méthode des réseaux neuronaux artificiels provient des vrais neurones présents dans un cerveau animal. Les RNA sont des systèmes de calcul parallèles constitués de groupes d'éléments de traitement hautement interconnectés appelés neurones, qui sont disposés en une série de couches. Chaque couche peut faire ses calculs indépendamment et peut transmettre les résultats à une autre couche. En cela, le résultat de la fonction de transfert est communiqué aux neurones dans la couche de sortie. C'est le point où les résultats sont interprétés et présentés finalement. [36]

La figure 3 représente un exemple d'architecture des réseaux de neurones.

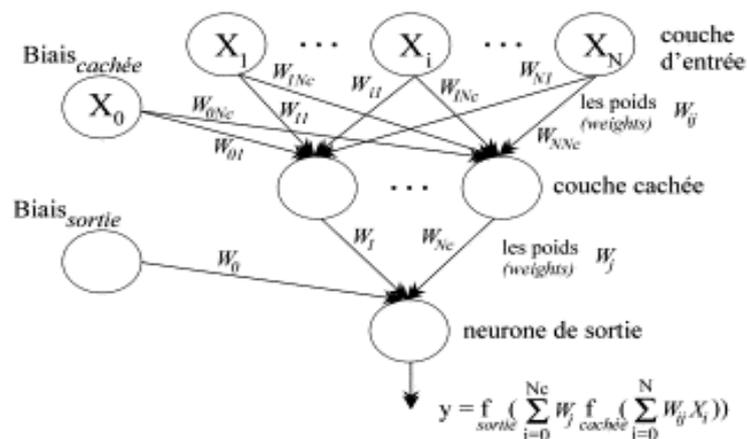


FIGURE 3 : Architecture des réseaux de neurones

### 5.5.9. K-Plus proche voisin

Les approches kNN sont exécutées par des distances entre un objet inconnu et tous les objets de l'ensemble d'apprentissage. Sur la base du calcul des distances, les objets de l'ensemble d'apprentissage les plus similaires à l'objet inconnu sont sélectionnés. La valeur k optimale est sélectionnée par optimisation en catégorisant l'ensemble d'échantillons. [37]

### 5.5.10. Classification des méthodes 3D-QSAR

Cela peut être classé dans diverses catégories. Certains d'entre elles sont basés sur la modélisation intermoléculaire telle que le ligand et le récepteur. Celles-ci peuvent également être classées sur la base de méthodes chimiométriques qui sont principalement utilisées pour la corrélation des propriétés structurales et des activités. De plus, le critère d'alignement est également une base de classification. Peu d'entre eux sont brièvement expliqués [2] :

- Basé sur un modèle de modélisation intermoléculaire : CoMFA, CoMMA, CoMSIA GERM.
- Basé sur le récepteur : CoRIA, COMBINE et AFMoC.
- Basé sur des méthodes chimiométriques Lineaire : CoMFA, CoMSIA et GERM.
- 3D-QSAR non linéaire : COMPASS et QPLS.
- Basé sur le critère d'alignement. Alignement dépendant : CoMFA, CoMMA, CoMSIA, GERM et CoRIA.
- Alignement indépendant : WHIM, CoSA, HQSAR, COMPASS etc.

#### ➤ Analyse comparative de champs moléculaires (CoMFA)

3D-QSAR a été utilisé comme synonyme de CoMFA, qui fut la première méthode à implémenter le concept dans une méthode QSAR, c'est-à-dire que l'activité biologique d'un ligand peut être prédite à partir de ses trois dimensions. [38]

Jusqu'à présent, CoMFA est probablement la méthode 3D-QSAR la plus utilisée. Une étude CoMFA commence normalement avec la modélisation traditionnelle du pharmacophore afin de suggérer une conformation bioactive de chaque molécule et des façons de superposer les molécules étudiées. [38]

L'idée sous-jacente de CoMFA est que les différences dans une propriété cible, par exemple, l'activité biologique, sont souvent étroitement liées à des changements équivalents dans les formes et les forces des champs d'interaction non covalents entourant les molécules. Ou indiqué d'une manière différente, les champs stériques et électrostatiques fournissent toutes les informations nécessaires pour comprendre les propriétés biologiques d'un ensemble de composés [39]. Par conséquent, les molécules sont placées dans une grille cubique et les énergies d'interaction entre la molécule et une sonde définie sont calculées pour chaque point de la grille. [38]

La figure 4 représente le champ stérique et électronique dans une grille CoMFA.

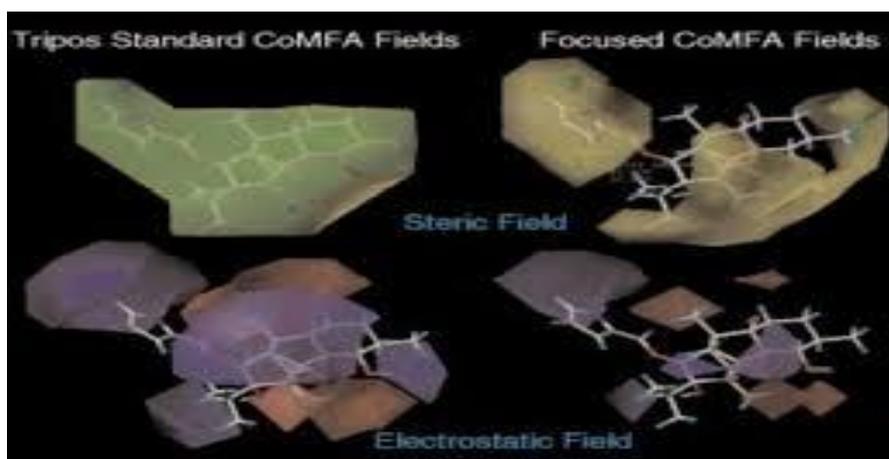


FIGURE 4 : Le champ stérique et électrostatique dans une grille CoMFA [38]

Dans une étude CoMFA, un alignement correct des molécules est essentiel mais souvent problématique. Un alignement optimal peut être défini comme l'obtention de la superposition maximale des champs stériques et électrostatiques d'un ensemble de molécules. L'alignement varie d'une molécule à l'autre en fonction de la similarité structurale ou de la diversité. Par conséquent, l'alignement influence considérablement les résultats du modèle, et des résultats significatifs et pertinents ne devraient être attendus que pour des alignements valides. [40]

#### ➤ CoMSIA

En raison des problèmes associés à la forme fonctionnelle du potentiel de Lennard-Jones utilisé dans la plupart des méthodes de CoMFA, Klebe et al. ont développé une méthode CoMFA basée sur des indices de similarité appelée CoMSIA (analyse comparative des indices

de similarité moléculaire). Au lieu de champs basés sur une grille, CoMSIA est basé sur des indices de similarité obtenus à l'aide d'une forme fonctionnelle adaptée à partir de l'algorithme SEAL.[41]

Trois indices différents liés aux potentiels stérique, électrostatique et hydrophobe ont été utilisés dans leur étude de l'ensemble de données de référence stéroïdien classique. Des modèles de qualité statistique comparable en ce qui concerne la validation croisée de l'ensemble d'apprentissage, ainsi que les prédictivités d'un ensemble de tests, ont été dérivés en utilisant CoMSIA [42].

### ➤ **AFMoC**

Une nouvelle méthode qui pourrait surmonter le problème de la négligence de l'information protéique dans une analyse 3D-QSAR a été récemment développée par Klebe et al. Dans cette approche appelée AFMoC (adaptation des champs pour la comparaison moléculaire) ou « CoMFA inversé », des champs potentiels dérivés d'une fonction de score (Drug Score) sont générés dans la poche de liaison d'une protéine cible. Méthodologiquement, le programme est lié à CoMFA et CoMSIA mais avec l'avantage d'inclure l'environnement de la protéine à l'analyse 3D-QSAR. Au lieu de n'utiliser que le potentiel de Coulomb ou de Lennard-Jones, l'AFMoC commence par une grille de valeurs pré-assignées [43].

Il a été démontré que les modèles QSAR dérivés de l'AFMoC permettent d'obtenir de bien meilleures corrélations entre les activités expérimentales et calculées comparées à la fonction de score originale Drug Score [44].

### ➤ **4D-QSAR and 5D-QSAR**

Les méthodes QSAR récemment développées incluent l'approche dite 4D-QSAR [45], où un ensemble de conformations pour chaque ligand représente la quatrième dimension, et 5D-QSAR, qui considère en plus des hypothèses pour les changements qui pourraient survenir dans une conformation d'un récepteur dû à la liaison du ligand (ajustement induit) comme cinquième dimension; si ces nouvelles approches QSAR montrent une qualité accrue en ce qui concerne la capacité prédictive et l'interprétation des résultats doivent être démontrées par de futures études de cas.[46]

## **5.6. 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 [47].

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

### 5.6.1. Validation interne

La validation interne utilise l'ensemble de données à partir duquel le modèle est construit et vérifie la stabilité interne. La technique de validation croisée (CV) est largement utilisée comme méthode de validation interne des modèles statistiques. Habituellement, un composé de l'ensemble est extrait à chaque fois, puis le modèle est recalculé en utilisant comme ensemble d'entraînement le  $n-1$  (où  $n$  est le nombre de composés) des composés restants, de sorte que la valeur d'activité biologique pour le composé extrait est prédite une fois pour tous les composés. Ce processus est répété  $n$  fois pour tous les composés de l'ensemble initial, obtenant ainsi une prédiction pour chaque objet. [48]

Ce processus fait référence à la méthode LOO (leave-one-out). Une méthode alternative peut également être définie en omettant plus d'un composé de l'ensemble de données à chaque fois. Cette méthode s'appelle la méthode CV leave  $n$  out ou Leave-many-Out (LMO) ou parfois elle est appelée leave-group-out (LGO) [49].

Le calcul du coefficient de corrélation de la procédure de validation croisée  $q^2$ , doit être fait et par définition plus petit ou égal à l'ensemble  $r^2$  (coefficient de corrélation) pour une équation QSAR.

Une autre procédure pour tester la validité du modèle est le test de randomisation. Même avec un grand nombre de composés et un petit nombre de descripteurs, une équation peut toujours avoir un très faible pouvoir prédictif. Une façon de tester cela est par randomisation des composés [50].

### 5.6.2. Validation externe

Un modèle QSAR avec une excellente qualité d'ajustement et des prédictions acceptables peut être déficient dans la relation réelle entre les descripteurs structurels et l'activité biologique.

La validité parfaite du modèle est examinée par validation externe, qui évalue la généralisation du modèle. L'ensemble de données original est divisé en deux sous-groupes, l'ensemble d'apprentissage et l'ensemble d'essai. L'ensemble d'apprentissage ou d'étalonnage est utilisé pour dériver un modèle d'étalonnage qui sera utilisé plus tard pour prédire les activités des composés d'essai ou de validation. D'autre part, un ensemble de test externe qui

n'a été inclus dans aucune étape de la construction du modèle peut être utilisé comme ensemble de test [47]. Les paramètres quantifiant la supériorité de la prédiction de l'ensemble de test externe peuvent être les mêmes que ceux utilisés pour la validation interne. La somme des erreurs de prédiction des carrés (SSPE) est largement utilisée pour rendre compte de l'incohérence.

### 5.6.3. Coefficients et tests statistiques standards :

Plusieurs tests et coefficient peuvent être utilisés pour juger et effectuer la validation de modèle QSAR [51], les plus importants sont cités dans le tableau V qui représente les tests standards pour validation de modèle.

TABLEAU V : Les tests standards pour validation de modèle

| Test                               | Définition  |
|------------------------------------|---|
| Test de Fisher (F)                 | $F(\text{observé}) = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \hat{y}_i)^2} \frac{n - p - 1}{p}$ |
| L'erreur quadratique moyenne (MSE) | $MSE = \frac{\sum(\hat{y}_i - y_i)^2}{n}$   |
| Coefficient de corrélation (r)     | $r = \sqrt{1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}}$                                |
| L'erreur type résiduel (s)         | $s = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n - p - 1}}$  |

$\hat{y}$  et  $y_i$ : les valeurs observées et calculées de la variable dépendante

n : nombre des observations (les molécules).

p : nombre de variables indépendantes (les descripteurs).

### 5.7. 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 [28] :

- 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.

La caractéristique clé du rôle des technologies in silico dans tous ces domaines est que les prédictions peuvent être faites à partir de la seule structure moléculaire.

– Figure n5 représente les différentes étapes d'étude QSAR

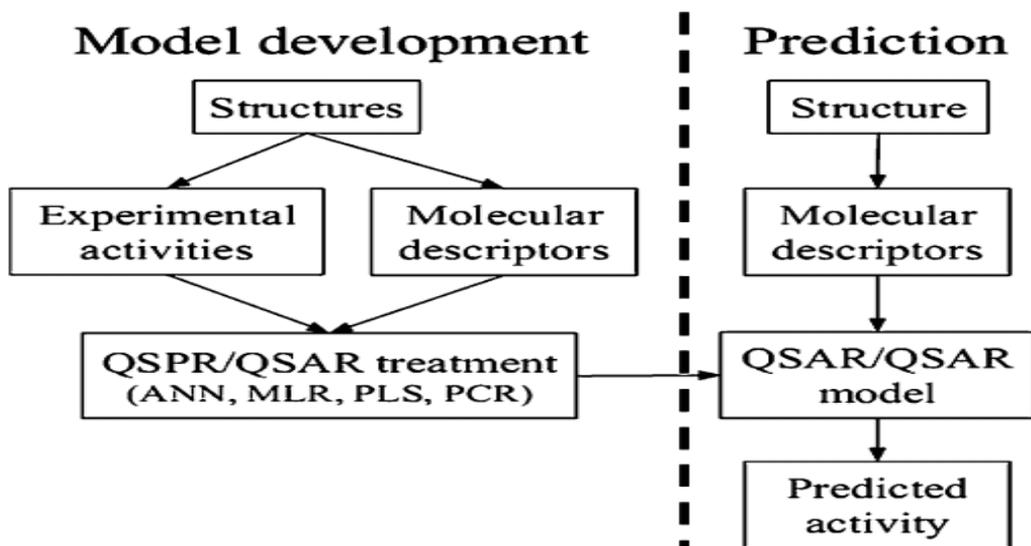


Figure n 5: différentes étapes d'étude QSAR

**Chapitre II : Coumarine :**  
**Chimie et aspect qualitative**  
**de RSA**

## 1. Généralités :

Les coumarines, également connues sous le nom de benzopyrones, sont présentes en quantités remarquables dans les plantes surtout chez certaines familles de Dicotylédones: Rutacées (*bergamote*), Fabacées (*fève Tonka*), Hippocastanacées. bien que leur présence ait également été détectée dans des microorganismes et des sources animales. La diversité structurelle trouvée dans cette famille de composés a conduit à la division en différentes catégories, des coumarines simples à beaucoup d'autres types de coumarines complexes, telles que les furocoumarines et les pyranocoumarines. [52]

Les coumarines simples et leurs analogues sont une grande classe de composés qui ont attiré leur intérêt depuis longtemps en raison de leurs activités biologiques : ils se sont avérés utiles comme agents antitumoraux, anti-VIH et comme composés actifs sur SNC. De plus, ils ont été signalés comme ayant de multiples activités biologiques anticoagulantes, anti-inflammatoires, bien que toutes ces propriétés n'aient pas été systématiquement évaluées. En outre, leurs propriétés d'inhibition enzymatique, leurs activités antimicrobiennes et antioxydantes sont d'autres sujets importants dans ce domaine de recherche. [53]

La figure 5 représente plusieurs substitutions qui peuvent affecter le noyau coumarine

## 2. 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 [54]. 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.

### 3. Définition

L'étude des coumarines a commencé il y a plus de 200 ans [55]. Le nom de Cette famille chimique est dérivé de *Coumarouna 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-a-pyrone (2H-1-benzopirane-2-one) dont la nomenclature systématique a été établie par IUPAC. [56]

Les Coumarines ont été grossièrement classées comme suit :

- a. Coumarines simples
- b. Furanocoumarines
- c. Pyranocoumarines
- d. Biscoumarines et Triscoumarines et

bien que l'on puisse faire des distinctions selon leur structure chimique parmi les divers groupes de coumarines. [57]

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. [58]

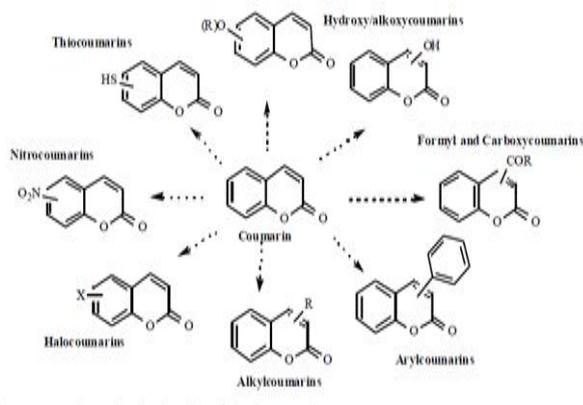


FIGURE 5 : Coumarine, en tant que noyau central pour l'obtention de coumarines simples [54]

### 3.1. Structure générale et propriétés des coumarines

La coumarine et ses dérivés sont les principaux médicaments Anticoagulants oraux. La coumarine est insoluble dans l'eau ; cependant la substitution de 4-hydroxy confère des propriétés faiblement acides à la molécule le rendent soluble dans l'eau et dans des conditions légèrement alcalines. [59]

Le figure 6 représente Les structures de la coumarine et de ses quelques dérivés sont indiquées ci-dessous.

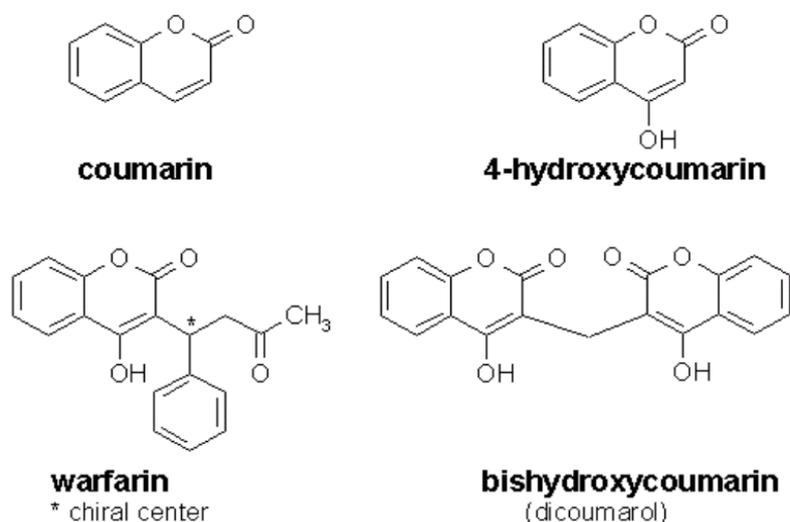


FIGURE 6 : Structure de coumarine et quelques dérivés [59]

### 3.2. 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 [60] :

#### 3.2.1. Coumarines simples

Les coumarines les plus répandues dans le règne végétal possèdent de Substitutions (OH ou OCH<sub>3</sub>) en 6 et 7 [61].

- La figure 7 représente la structure générale de coumarine simple
- Le tableau VI représente la coumarine simple et ses dérivés

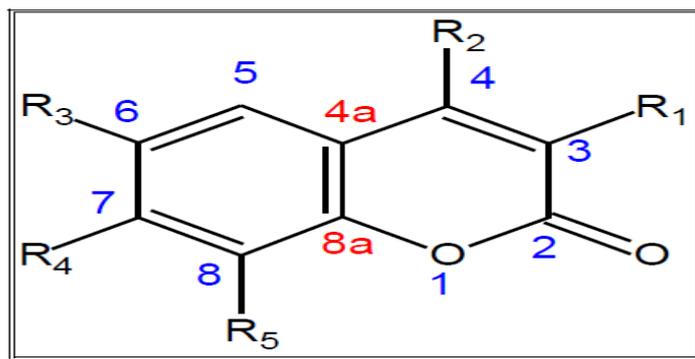


FIGURE 7 : Structure générale de coumarine simple [61]

– Le tableau VI représente les coumarines simples et ses dérivés

TABLEAU VI : Coumarine simple et ses dérivés [52]

| R1 | R2   | R3   | R4     | R5 | Coumarine                     |
|----|------|------|--------|----|-------------------------------|
| H  | H    | H    | H      | H  | Coumarine                     |
| OH | H    | H    | H      | H  | 3-hydroxy coumarine           |
| H  | OH   | H    | H      | H  | 4-hydroxycoumarine            |
| H  | H    | H    | OH     | H  | Ombelliférone                 |
| H  | OCH3 | H    | H      | H  | 4-méthoxy coumarine           |
| H  | H    | H    | OCH3   | H  | Herniarine                    |
| H  | CH3  | H    | OH     | H  | 7-hydroxy-4-méthoxy coumarine |
| H  | H    | CH3  | OH     | H  | Scopolétine                   |
| H  | H    | OCH3 | OCH3   | H  | Scoparone                     |
| H  | CH3  | H    | OCOCH3 | H  | 7-acétoxy-4-méthylcoumarine   |
| H  | H    | OCH3 | OCH3   | OH | Fraxidine                     |

### 3.2.1.1. Furanocoumarines :

Les furocoumarines (appelées encore furanocoumarines) 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 [62]. 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. [63]

Figure 8 représente les structures de quelques furanocoumarines.



Figure 8 : Structures de quelques furanocoumarines [64]

### 3.2.1.2. Pyranocoumarines

Composés formés par la fusion d'un hétérocycle pyrane avec la coumarine [65]

1- soit dans le prolongement (forme linéaire) : xanthylétine

2- soit latéralement (forme angulaire) : séseline, visnadine

La figure 9 représente la structure de quelques pyranocoumarines



FIGURE 9 : Structures de quelques Pyranocoumarines [65]

### 3.2.2. Dicoumarines (coumarines dimériques)

Ce sont des composés formés par la liaison deux unités coumariniques simples. [54]

La figure 10 représente les structures des quelques Dicoumarines



FIGURE 10 : Structures des quelques Dicoumarines [54]

### 3.2.3. Tricoumarines (coumarines trimériques)

Ce sont des composés issus de l'union de trois unités coumariques. [54]

La figure 11 représente les structures des quelques Tricoumarines

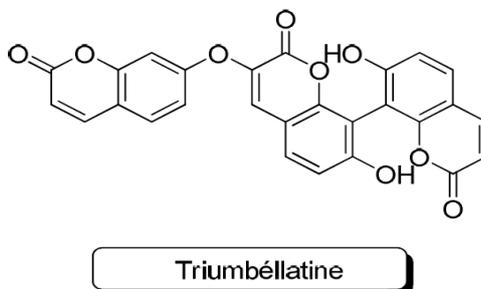


FIGURE 11 : Structures de quelques des Tricoumarines [54]

## 4. Relation structure activité biologique des coumarines et des analogues simples

De nombreuses activités biologiques ont été attribuées à de simples coumarines et ses analogues. Parmi eux, des activités antimicrobiennes, antivirales, anticancéreuses, d'inhibition d'enzymes, anti-inflammatoires, antioxydantes, anticoagulantes et du système nerveux central. [64].

### 4.1. A. Antimicrobien et molluscicide

La novobiocine et la clorobiocine sont des antibiotiques coumariniques d'origine naturelle, qui sont des inhibiteurs de l'ADN girase, et ont un large spectre de bactéries Gram (+), y compris les souches de staphylocoques résistantes à la méthiciline. [66]

La figure 12 représente les structures d'antibiotiques naturels

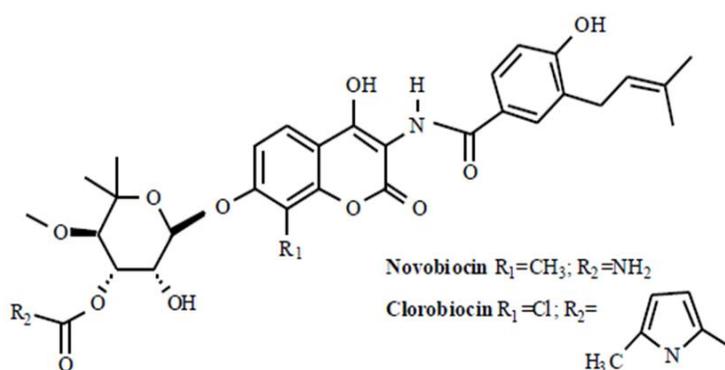


FIGURE 12: Structures d'antibiotiques coumariniques naturels [64]

En raison de certaines limitations de ces composés, en particulier en ce qui concerne la solubilité, la toxicité et le développement d'une résistance, une nouvelle série d'analogues de la coumarine a été synthétisée. [67]

Des références pour l'activité molluscicide de certains dérivés de 4-phénylcoumarine ont déjà été décrites. Cette activité a également été attribuée aux coumarines *d'Ethulia conyzoides* et au dicoumarol.

L'activité acaricide des extraits végétaux et des coumarines a également été étudiée, à partir d'études SAR de composés bioactifs réalisés dans ce but. [68]

#### 4.2. Antivirale

L'activité antivirale des coumarines simples se concentre essentiellement sur l'inhibition de la protéase du VIH-1 (VIH-PR) et de l'intégrase du VIH-1. Les avancées récentes dans le développement de dérivés de la coumarine en tant qu'agents anti-VIH puissants, concernant la découverte, la modification structurale et les études de relations structure-activité, ont fait l'objet de différentes revues ou articles mis à jour. [69]

De cela, phenprocoumon, La warfarine et les dérivés de 4-hydroxy-2-pyrone substitués sont, en fait, appelés la première génération d'inhibiteurs du VIH-PR.

En outre, certains dérivés de la coumarine ont été testés pour leur activité contre le virus de l'herpès simplex (HSV). De ces études, la 5,7,4'-trihydroxy-4 styrylcoumarine s'est avérée présenter une activité antivirale significative. Il est intéressant de noter que la collinine naturelle, greveal, possède une importante activité de réplication de l'ADN anti-VHB. [58]

#### 4.3. Anticancer

Parmi les coumarines dépistées pour l'activité anticancéreuse, la géiparvarine (figure 13) a été jugée la plus représentative. La géiparvarine est une structure naturelle à base de coumarine, isolée des feuilles de *Geijera parviflora Lindl*, qui est connu pour son importante activité cytostatique in vitro. [54]

La figure 13 représente Structure de la coumarine naturelle geiparvarin

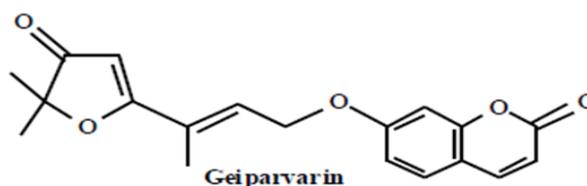


FIGURE 13 : Structure de la coumarine naturelle geiparvarin. [54]

Les effets in vitro de 4 hydroxycoumarine ont été étudiés, en utilisant un mélanome et les lignées cellulaires fibroblastiques non malignes. Il a été proposé que le composé puisse être un adjuvant utile pour le traitement du mélanome. [64]

La mutagénicité des coumarines simples a également été étudiée. Il a été constaté que l'antimutagénicité de cette classe de composés était liée à la présence de fonctions polaires aux carbones 3-, 4- et 7. Il est important de noter que les effets chimiopréventifs de plusieurs coumarines naturelles ont été évalués. [9]

#### **4.4. Inhibition d'enzymes**

Certaines coumarines naturelles et synthétiques se sont révélées être des inhibiteurs du cholinestérase, considérés comme une approche prometteuse pour le traitement de la maladie d'Alzheimer et pour d'éventuelles applications thérapeutiques dans la région.

Une étude QSAR des dérivés de la coumarine agissant sur le MAO B a été réalisée, révélant l'importance des interactions lipophiles dans la modulation de l'inhibition et excluant toute dépendance aux propriétés électroniques. [70]

#### **4.5. Antioxydant**

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. [58]

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. [63]

#### **4.6. Anti-inflammatoire**

Il a été trouvé que plusieurs coumarines isolées de plantes ou d'origine synthétique possèdent des activités anti-inflammatoires et / ou analgésiques significatifs. [59]

Dans une étude QSAR pour l'optimisation du lead dans la conception des coumarines en tant qu'agents anti-inflammatoires non stéroïdiens puissants en inhibent le COX-2, il a été constaté que les substituants dans les positions coumarine C4 et C7 contribuaient à la forte activité. La lipophilicité et l'ionisation se sont révélées être des

paramètres physico-chimiques importants qui étaient corrélés avec l'activité biologique. [57]

#### 4.7. Anticoagulant et cardiovasculaire

La warfarine est considérée, même aujourd'hui, comme l'anticoagulant coumarinique dominant, en raison de son excellente puissance et bon profil pharmacocinétique. Alors que sa forme de marché est le sel de sodium racémique, l'anticoagulant sait que l'activité de l'énantiomère (S) - (-) est six fois plus élevée que celle de l'énantiomère (+).

Néanmoins, il a été démontré que les dérivés de 4-hydroxycoumarine possèdent un effet anticoagulant supérieur à celui de la warfarine. [71]

La figure 14 représente la structure des principaux composés : warfarine, dicoumarol.

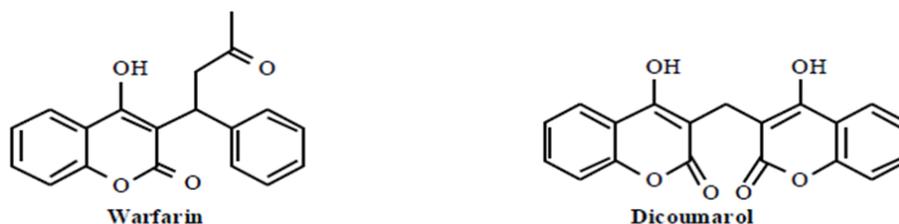


FIGURE 14 : Structures des principaux composés warfarine et dicoumarol. [71]

Il est bien connu que plusieurs dérivés de la coumarine possèdent des propriétés cardiovasculaires. Carbochromène (3-diéthylaminoéthoxycarbonylméthoxy-4-méthyl coumarine), est un vasodilatateur coronarien puissant spécifique, qui a été utilisé pendant de nombreuses années dans le traitement de l'angine de poitrine. Bien que le mécanisme d'action reste inconnu, il a été rapporté que les effets coronariens carbochromen pourraient être médiés par une libération accrue de prostaglandines. [71]

#### 4.8. Divers

L'activité hépato protectrice des coumarines de *Morus alba L.* (*Moraceae*) a été étudiée. En outre, le mécanisme de l'hépatotoxicité induite par la coumarine a été étudié, en utilisant différents substrats de la coumarine. [53]

Osthol, une coumarine naturelle isolée des fruits séchés de *Cnidium monnieri* (*Umbelliferae*), révèle posséder une activité antiallergique.

Certaines coumarines simples montrent la capacité de moduler les réponses au calcium évoquées par l'agoniste dans les cellules thyroïdiennes FRTL-5, une lignée cellulaire non excitable, et de stimuler l'acide aminé endogène libéré par les synaptosomes. [63]



# **Chapitre III :**

**Étude de la relation structure  
activité quantitatives QSAR sur  
différents dérivés de coumarines**

Dans ce chapitre nous avons traités plusieurs études QSAR sur différents dérivés des coumarines, on se basant sur des livres et des recherches publiées par exemple dans : Research Gate – science direct – ELSEVIER – journal of chemical information and modeling – ACS publication– google scholar – PubMed ..... Après nous avons trié et classé ces articles selon :

- Leurs structures principales : coumarine simple – Biscoumarines
- Les différentes substitutions sur une structure coumarinique principale : mono substitution – bi substitution – poly substitutions ...
- Les activités biologiques, nous essayons de couvrir plusieurs et différentes activités thérapeutiques.

## 1. Coumarines simples :

On étudie les coumarines simples mono substituée, bi et tri substituée.

### 1.1. Coumarines simple monosubstitués :

Deux exemples ont été traités : la substitution en C3 puis en C4

#### 1.1.1. Substitution on c3 :

##### 1.1.1.1. Etude QSAR sur nouveaux dérivés de la coumarine en tant qu'inhibiteurs sélectifs puissants de la COX-2 (activité anti-inflammatoire)

###### a. Données expérimentales

Récemment, les dérivés du benzopyrane ont été étudiés comme inhibiteurs puissants de la COX-2, et plus spécifiquement, les dérivés de la coumarine, en tant que classe de benzopyranes, possédaient de puissants effets anti-inflammatoires. Dans le but d'identifier des nouveaux inhibiteurs de la COX-2, **Dina H. Dawood et al** [72] Profitent ces propriétés pour la conception et l'étude QSAR de ce type d'inhibiteurs.

Leur travail était basé sur l'hybridation des deux groupements pharmacophoriques dans une seule molécule, pour cela, deux nouvelles séries de dérivés de coumarine incorporant thiazoline et thiazolidinone ont été conçues, synthétisées, et étudiées in vivo pour leurs activités anti-inflammatoires (figure 15) :

La figure 15 représente la stratégie de conception pour la synthèse des composés cibles

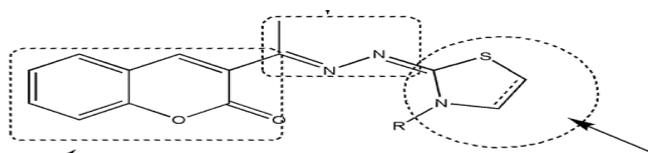


FIGURE 15 : Stratégie de conception pour la synthèse des composés cibles [72]

L'activité anti-inflammatoire des composés cibles a été évaluée in vivo en utilisant des modèles d'œdème de patte de rat induits par la carraghénine.

L'inhibition de la COX et l'effet ulcérogène des composés bioactifs synthétisés ont également été étudiés.

La relation structure-activité quantitative (QSAR) de l'ensemble actif de composés et l'étude d'amarrage moléculaire ont été réalisées pour interpréter et expliquer le mode de liaison possible des composés actifs.

Un ensemble de 21 composés de dérivés de la coumarine ayant montré une puissante activité inhibitrice de la COX-2 a été utilisée pour cette étude.

L'ensemble de données a été divisé en deux ensembles : ensemble d'entraînement contenant 15 composés (2a, 2b, 3a, 3b, 4a, 4b, 5a, 5b, 6a, 6b, 7a, 7c, 7d, 7e et 7f) avec un thiazolinyle substitué échafaudages et ensemble d'essai qui contient six composés (8a, 8b, 9a, 9b, 10a et 10b) avec l'échafaudage thiazolidin-4-one plus les deux médicaments de référence, l'indométhacine et le célécoxib.

#### **b. Descripteurs moléculaires et méthodes statistiques**

##### **▪ Descripteurs moléculaires**

Tous les composés ont été construits en utilisant l'adjuvant moléculaire MOE 2014.09. Une étude QSAR a été menée pour trouver les meilleurs descripteurs auxquels l'activité anti-inflammatoire pourrait être corrélée.

Tous les descripteurs 2D et 3D fournis dans le MEO 2014.09 ont été calculés pour chaque composé de l'ensemble d'apprentissage.

##### **▪ méthodes statistiques**

L'analyse des moindres carrés partiels (PLS) a été utilisée comme méthode statistique pour dériver la signification des modèles.

Les valeurs de pIC<sub>50</sub> ont été calculées pour toutes les valeurs IC<sub>50</sub> des ensembles d'entraînement et d'essai dans une échelle micromolaire en utilisant l'équation suivante :

$$\text{pIC}_{50} = -\log(\text{IC}_{50})$$

#### **c. Validation du modèle**

Afin de valider la puissance du modèle on va calculer pour chaque ensemble de descripteur :

- Le coefficient de corrélation (R2)
- L'erreur quadratique moyenne (MSE) en utilisant l'équation suivante pour les modèles linéaires :

$$MSE(a_0, a) = \frac{1}{W} \sum_{i=1}^m w_i [y_i - (a_0 + a^T x_i)]^2$$

- Test t (utilisé pour vérifier s'il existe une différence significative entre les valeurs pIC50 observées et pIC50 prédites dans l'ensemble d'apprentissage).

#### **d. Résultats et discussion**

##### **▪ Résultats**

Les descripteurs obtenus après optimisation et qui montrent une corrélation significative sont montrés dans le tableau VII.

**TABLEAU VII : Les descripteurs 2D et 3D calculés (72)**

| Les descripteurs 2D  | Les descripteurs 3D                                 |
|--|---|
| <b>Propriétés physiques</b>  | Descripteurs d'énergie potentiels                   |
| <b>Surface subdivisée</b>  | MOPAC   |
| <b>Nombre d'atomes et nombre de bons</b>                             | Descripteurs de surface, de volume et de forme      |
| <b>Connectivité de Kier &amp; Hall et indices de forme de Kapper</b> | Descripteurs de charge dépendant de la conformation |
| <b>Descripteurs d'entités pharmacophores</b>                         |   |
| <b>Charges partielles</b>  |   |

D'après les résultats de calcul et après la validation du pouvoir prédictif de chaque ensemble des descripteurs, les résultats obtenus sont :

Le tableau VIII représente PIC50 prévu de l'ensemble de test pour la validation du modèle QSAR

**TABLEAU VIII : PIC50 prévu de l'ensemble de test pour la validation du modèle QSAR(72)**

| Composé      | IC50 observe | PIC calcule | 50 Pic50 predict | Résiduel |
|--------------|--------------|-------------|------------------|----------|
| 8a.          | 0.78         | 6.107       | 6.178            | - 0.071  |
| 8b           | 0.44         | 6.356       | 6.348            | 0.008    |
| 9a           | 0.39         | 6.408       | 6.399            | 0.009    |
| 9b           | 0.48         | 6.318       | 6.298            | 0.02     |
| 10a          | 0.55         | 6.259       | 6.265            | - 0.006  |
| 0010b        | 0.61         | 6.214       | 6.266            | - 0.052  |
| Indomethacin | 2.63         | 5.580       | 5.583            | - 0.003  |
| Celecoxib    | 0.3          | 6.522       | 6.513            | 0.009    |

- **Les descripteurs 3D** : la surface, le volume et la forme présentent la valeur R<sup>2</sup> la plus élevée (0,908) et la valeur RMSE la plus faible (0,035) par rapport aux autres.
- **Les descripteurs 2D** : charges partielles et comptes d'atomes et comptes de liaisons montraient des valeurs de R<sup>2</sup> presque élevées 0,835 et 0,824 respectivement.

Les résultats de test t montrant une différence insignifiante entre la pIC<sub>50</sub> réelle et la pIC<sub>50</sub> prédite (légèrement <0,05).

#### ▪ Discussion

- Les mesures in vivo ont indiqué que la plupart des composés synthétisés possédaient une activité anti-inflammatoire supérieure à celle de l'indométacine dans le bio essai de l'œdème des pattes induit par la carragénine chez les rats (Tableau VIII).
- Les molécules synthétisées sont testées pour l'activité d'inhibition de COX-1 et COX-2. Les résultats ont indiqué que tous les composés montrent des meilleurs profils inhibiteurs contre la COX-2 que l'indométacine ce qui indique qu'ils sont des inhibiteurs sélectifs de la COX-2 (2b, 3a, 3b, 5b, 6a, 7f et 9a présentent un pouvoir inhibiteur et un taux de sélectivité supérieure au célécoxib de référence).
- L'effet ulcérogène des composés actifs a été évalué. Tous les composés testés se sont révélés avoir des profils d'innocuité GI supérieurs (ulcération de 0 à 7%) par rapport au médicament de référence (indométacine) (ulcération à 100%).
- La plupart descripteurs 3D de la surface, du volume et de la forme sont inversement proportionnels à l'activité telle que celle des descripteurs Vol, qui décrit le volume de van der Waals et qui devrait être de faible valeur pour avoir un composé hautement actif (Vsurf\_S- Vsurf\_W).
- Le rapport hydrophile / lipophile optimal a été décrit par les descripteurs Vsurf\_HL et s'est révélé être 0,12 pour l'activité la plus élevée du composé 2b et du célécoxib.
- Descripteurs 2D : Le nombre d'atomes d'hydrogène (a\_nH) a montré une grande contribution dans l'activité. La plage idéale du nombre d'atomes d'hydrogène s'est révélée être 14-18 (comme 2b et le célécoxib), le composé 7d a plus de 18 atomes d'hydrogène et est moins actif que 2b. En outre, le composé 8a a moins de 14 atomes d'hydrogène et est également moins actif.

- Le nombre de liaisons simples (b-single) est favorisé pour être de 25 à 30 pour une activité plus élevée.
- Le nombre d'atomes d'azote doit être égal dans tous les composés.
- La surface polaire totale positive de van der Waals (Q\_VSA PPOS) doit être de faible valeur pour augmenter l'activité.

#### e. Conclusion

Les résultats obtenus dans ce travail expriment une activité anti-inflammatoire élevée et un pouvoir sélectif COX-2 significatif aux molécules de références (indométhacine-celecoxib) avec un effet ulcérogène minime (ulcération de 0 à 7%).

Tous ces résultats sont tirés à partir de modèle QSAR étudié qui a montré une corrélation importante avec l'activité et un pouvoir prédictif élevé ( $R^2=0.908$ ) capable de prédire l'activité biologique. La sélectivité COX-2 peut être améliorée si on augmente le caractère lipophile. [72]

#### 1.1.2. Substitution on c4 :

##### 1.1.2.1. Etude de la relation structure activité quantitative d'une série des dérivés de la coumarine-pipérazine en tant qu'inhibiteurs de l'acétylcholinestérase

###### a. Données expérimentales :

Dans le cadre une recherche réalisée par **Rahul P. Modh1 et al.** [73] Sur une série d'analogues de coumarine-pipérazine qui a été créée sur la base des caractéristiques structurales de l'ensaculine. Ces molécules ont été synthétisées et évaluées pour leur puissance à inhiber l'enzyme hAChE.

La figur16 représente la stratégie de conception des dérivés de la coumarine-pipérazine.

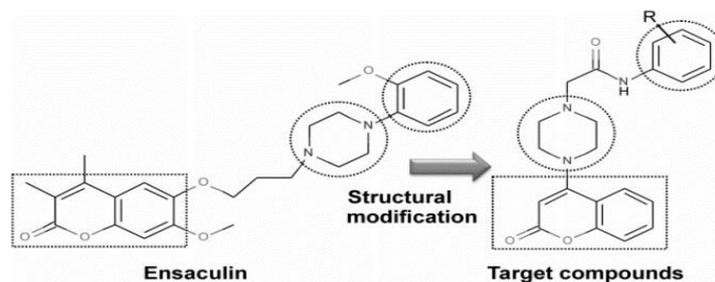


FIGURE 16 : Stratégie de conception des dérivés de la coumarine-pipérazine [73]

Des études de modélisation moléculaire ont été réalisées à l'aide de méthodes d'amarrage moléculaire et de modélisation QSAR basées sur les champs moléculaires (Molecular field-based) pour comprendre le mécanisme d'inhibition.

La puissance inhibitrice de hAChE des analogues de coumarine-pipérazine synthétisés a été mesurée en utilisant la méthode d'Ellman pour déterminer son intérêt potentiel envers les inhibiteurs de hAChE. Pour permettre la comparaison des résultats, l'ensaculine a été utilisée comme composé de référence.

#### **b. Descripteur moléculaire et les méthodes statistiques**

##### **▪ Descripteurs moléculaires**

- Les structures 3D de l'ensemble des composées ont été importées dans l'espace de travail Forge, La conformation 3D de l'ensaculine a été récupérée à partir de la base de données NCBI Pubchem (CID 208923).
- L'ensaculine a été spécifiée comme molécule pivot tandis que les molécules synthétisées (3a-3n) ont été marquées comme ensemble d'entraînement.
- Une recherche conformationnelle systématique a été réalisée avec l'ensemble de données (3a-3n, ensaculine) et les meilleurs conformateurs (triés par la baisse Energie conformationnelle) ont été retenus.
- Cinquante conformères de chaque molécule dans l'ensemble de données ont été générés et autorisés à participer à l'alignement basé sur le champ de la force moléculaire pour l'alignement final  $> 0,5$ .
- 25% des molécules synthétisées ont été choisies au hasard par Forge V10 considéré comme ensemble de test (4 molécules), tandis que les autres composés ont participé à la formation du modèle QSAR.

##### **▪ La méthode statistique : Molecular field-based**

- Le modèle QSAR utilisé dans ce travail est basé sur le champ moléculaire (Molecular field-based QSAR model) généré par Forge V10 (Cresset Biomolecular Discovery Ltd).
- Les activités expérimentales de l'ensemble de données (IC50) ont été converties à son échelle logarithmique en utilisant la formule (1) et définie comme variable dépendante.  $pIC50 = -\log(IC50)$  (1)
- L'équation du modèle QSAR développé est :

$$pIC50(\text{prévu}) = 0.95 (pIC50 \text{ réel}) + 0.194$$

**c. Validation de model :**

Afin d'évaluer la signification des modèles et la capacité de prédiction précise pour les nouveaux composés, **Rahul P. Modh1 et al. [73]** utilisent une procédure de validation interne qui est la technique « leave one out » (LOO) pour optimiser le modèle d'activité. Le meilleur modèle a été validé par le coefficient de corrélation ( $r^2$ ), le coefficient d'inter- corrélation ( $q^2$ ) et le score de similarité (Sim) des descripteurs choisis de chaque ligand par rapport à la molécule pivot.

Sim = le champ global + score de similarité de forme

**d. Résultat et discussion**

▪ **Résultats**

- Les Résultats du calcul du model QSAR sont montrés dans le tableau IX
- Le modèle QSAR construit montre un pouvoir prédictif significatif et une corrélation très importante des descripteurs avec l'activité pour chaque composé testé.

$$r^2 = 0.944$$

$$q^2 = 0.874$$

Les valeurs de score de similarité (Sim) pour chaque compose sont montres dans le tableau IX représente valeurs pic50 réelles et prévues du modèle QSAR basé sur les champs moléculaires

**TABLEAU IX : Valeurs pic50 réelles et prévues du modèle QSAR basé sur les champs moléculaires [73]**

| Composé               | Activité basée sur le domaine moléculaire |             |       |
|-----------------------|---|-------------|-------|
|                       | Pic50 réelle                              | Pic50 prévu | Sim   |
| <b>3a</b>             | 4.26                                      | 4.40        | 0.624 |
| <b>3b<sup>b</sup></b> | 3.51                                      | 4.10        | 0.626 |
| <b>3c</b>             | 4.68                                      | 4.30        | 0.640 |
| <b>3d<sup>b</sup></b> | 3.54                                      | 4.30        | 0.646 |
| <b>3<sup>e</sup></b>  | 3.32                                      | 3.30        | 0.621 |
| <b>3f</b>             | 5.08                                      | 5.10        | 0.636 |
| <b>3g</b>             | 4.39                                      | 4.40        | 0.635 |
| <b>3h</b>             | 4.16                                      | 4.20        | 0.651 |
| <b>3i<sup>b</sup></b> | 4.23                                      | 5.00        | 0.592 |
| <b>3j</b>             | 5.26                                      | 5.30        | 0.625 |
| <b>3k<sup>b</sup></b> | 4.23                                      | 4.70        | 0.643 |
| <b>3l</b>             | 4.71                                      | 4.80        | 0.614 |
| <b>3m</b>             | 5.31                                      | 5.20        | 0.622 |
| <b>3n</b>             | 4.04                                      | 4.10        | 0.650 |
| <b>Ensaculin</b>      | 6.44                                      | 6.34        | 0.630 |

▪ **Discussion**

- Les différentes substitutions sur le cycle phényle dans les molécules de coumarine-pipérazine synthétisées ont montré une activité AChE significative au positions méta qui a diminué remarquablement s'il était déplacé en positions ortho et para.
- L'alignement de l'ensemble des données montre une distribution uniforme de petit champ hydrophobe sur tout la surface (cette hydrophobicité conférée par les groupes aromatiques est très importante dans activité AChE)
- Les fragments aromatiques et le centre positif interagissent spécifiquement avec trois sites biochimiques de hAChE à savoir, CAS, ABS et PAS, respectivement.
- Les composés synthétisés peuvent être améliorés en introduisant divers chémotypes dans le cycle phényle ayant une capacité de liaison Hydrogène de manière à interagir avec Arg296.
- L'arrangement topologique des fragments coumarine et pipérazine s'est avéré être réalisable tandis que la longueur du lieur aliphatique peut être modifiée pour développer des inhibiteurs de l'AChE.

#### **e. Conclusion**

Les auteurs ont conçu et suggéré, sur la base des études QSAR une nouvelle série des dérivés de coumarine-pipérazine qui est synthétisée et suivie d'une substitution par divers acétamides aromatiques.

L'activité inhibitrice de l'enzyme hAChE des composés synthétisés est évaluée en utilisant la méthode Ellman. Où les composés les plus puissants sont : 3f- 3j- 3mN avec des valeurs IC<sub>50</sub> de 8,3- 5,5 et 4,9 mmol / L, respectivement.

A partir du modèle d'activité moléculaire basé sur le champ, ils ont conclu que les caractéristiques aromatiques ainsi que divers groupes substitués des molécules sont des caractéristiques essentielles pour développer une interaction avec le site hAChE PAS et sont fortement associées aux activités inhibitrices de hAChE.

## 1.2. Coumarine simple bi substituée

On a étudié deux exemples de substitutions en C3 et C4 et en C4 et C7 et enfin en C 7 et C8

### 1.2.1. Substitution en C3et C4

#### 1.2.1.1. Etude QSAR sur l'activité inhibitrice de De CDK de 3-Substitued 4-Hydroxycoumarin comme une potentielle thérapeutique anticancéreuse.

##### a. Données expérimentales

Les inhibiteurs spécifiques de la CDK sont des cibles importantes dans la découverte de médicament grâce à son activité anti-tumorale, ils induisent l'apoptose en perturbant le cycle cellulaire et / ou en réprimant la transcription.

La plupart des inhibiteurs de la CDK actuellement disponibles affectent deux ou plusieurs kinases, il était difficile de découvrir de nouveaux inhibiteurs mono-spécifiques.

Une nouvelle série de dérivés de 4-hydroxycoumarine substituée en 3 (voir figure 17)a été conçue, synthétisée et évaluée par **Nehad A. Abdel Latif et al**[74] pour leurs activités inhibitrices et anticancéreuses de CDK in vitro contre dix-huit types de lignées cellulaires tumorales humaines.

Tous les composés cibles synthétisés ont montré une affinité et une sélectivité remarquablement élevées vis-à-vis de CDK1B, par rapport au flavopiridol (molécule de référence), outre les molécules 3a, 6b et 8c ont été évaluées dans le modèle de xéno greffe de souris du cancer du sein MCF-7 et ont montré une efficacité in vivo à une dose de 10 mg / kg. Ces résultats sont confirmés par une étude QSAR pour trouver les descripteurs qui peuvent être corrélés à l'activité.

La figure 17 représente stratégie de conception pour la synthèse des composés cibles

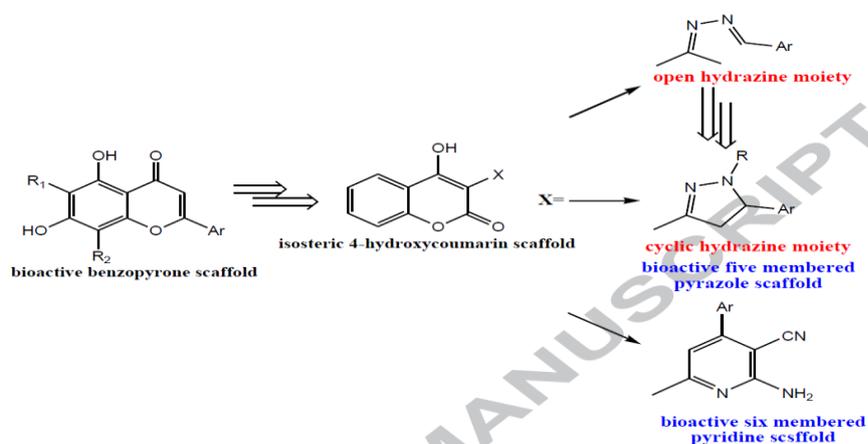


FIGURE 17 : Stratégie de conception pour la synthèse des composés cibles[74]

**b. Descripteurs moléculaires et méthode statistique**

▪ **Descripteurs moléculaires**

Un ensemble de 151 descripteurs 2D et 85 descripteurs 3D fournis par le MEO 2015.10 ont été calculés pour chaque composé de l'ensemble d'apprentissage.

- Les descripteurs 2D calculés : les propriétés physiques, la surface subdivisée, les nombres d'atomes, la connectivité de Kier & Hall et les indices de forme de Kapper, les descripteurs d'entités pharmacophores et les charges partielles.
- Les descripteurs 3D incluaient : les descripteurs d'énergie potentiels, MOPAC, les descripteurs de surface, le volume et la forme et les descripteurs de charge dépendant de la conformation.

▪ **Méthode statistique : Analyse partielle des moindres carrés (PLS)**

L'analyse partielle des moindres carrés (PLS) a été utilisée au cours de ce processus pour assurer l'exactitude du modèle.

Les valeurs de pKi prédites ont été calculées pour toutes les valeurs de Ki en utilisant l'équation suivante :

$$pKi = -\log(Ki)$$

Ki: Kinase inhibition

**c. Validation du modèle QSAR**

Dans le cadre de confirmer le pouvoir prédictif de modèle QSAR **Nehad A. Abdel Latif et al** ont calculés pour chaque ensemble de descripteurs les paramètres suivants :

- R2 : Les valeurs de coefficient de corrélation d'erreur quadratique moyenne
- RMSE d'erreur quadratique moyenne
- MSE a été calculé en utilisant l'équation suivante pour le modèle :

$$MSE(a_0, a) = \frac{1}{W} \sum_{i=1}^m w_i [y_i - (a_0 + a^T x_i)]^2$$

**d. Résultats et discussions**

▪ **Résultats**

- Les différents résultats biologiques de l'inhibition de kinase de différentes molécules synthétisées sont montrés dans le tableau X :
- Le tableau X présente effet inhibiteur de CDK des composés testés.

- Dans cette étude QSAR, les descripteurs de surface subdivisés (2D) jugés comme les descripteurs les plus significatifs et les plus corrélés avec activité grâce à son pouvoir prédictif élevé.
- Un modèle QSAR avec (R2) supérieur à 0,6 est considéré comme très précis, dans le cas de notre modèle R2=0,81 et RMSE la plus faible 0,01
- Donc le modèle construit est hautement précis avec un pouvoir prédictif élevé.

**TABLEAU IX : Effet inhibiteur de CDK des composés testés**

| molécule     | Inhibition de la kinase. Ki (nM) |        |       |       |       |
|--------------|----------------------------------|--------|-------|-------|-------|
|              | CDK6D3                           | CDK9T1 | CDK7H | CDK1B | CDK2E |
| Flavopiridol | 3                                | 113    | 13    | 79    | 265   |
| 3a           | 23.45                            | 44.56  | 0.56  | 24.53 | 39.74 |
| 3b           | 9.87                             | 52.65  | 0.84  | 43.63 | 25.36 |
| 3c           | 4.56                             | 75.09  | 0.66  | 15.45 | 23.55 |
| 3d           | 3.26                             | 56.46  | 0.46  | 23.48 | 25.96 |
| 5a           | 6.78                             | 96.89  | 0.67  | 34.88 | 14.33 |
| 5b           | 5.98                             | 85.77  | 0.59  | 25.77 | 23.24 |
| 5c           | 7.89                             | 72.66  | 0.88  | 16.56 | 14.35 |
| 5d           | 7.09                             | 63.36  | 0.76  | 13.65 | 25.46 |
| 6a           | 7.65                             | 74.64  | 0.77  | 14.14 | 36.78 |
| 6b           | 6.78                             | 86.57  | 0.63  | 16.25 | 25.87 |
| 6c           | 3.46                             | 99.46  | 0.54  | 15.36 | 14.36 |
| 7a           | 6.49                             | 88.32  | 0.35  | 19.57 | 13.40 |
| 7b           | 4.59                             | 77.65  | 0.40  | 10.68 | 22.55 |
| 7c           | 6.49                             | 96.87  | 0.59  | 18.36 | 13.63 |
| 8a           | 4.38                             | 85.67  | 0.68  | 16.57 | 24.31 |
| 8b           | 6.58                             | 92.45  | 0.77  | 14.46 | 15.42 |
| 8c           | 4.68                             | 71.36  | 0.84  | 15.67 | 24.56 |

#### ▪ Discussion

- Dix-sept composés cibles synthétisés inhibaient sélectivement CDK1B, et présentaient une activité inhibitrice supérieure à celle du flavopiridol (Ki = 0,35-0,88 nM vs Kiflav = 13 nM), le composé 7a était l'inhibiteur de CDK1B le plus puissant (Ki = 0,3 nM).
- La diminution des propriétés inhibitrices de CDK9T1 et de CDK1B des composés 5c et 5d (de pyrazoline cyclisés contenant les substituants donneurs d'électrons) par rapport aux congénères d'hydrazone ouverts 3d et 3c.
- Les dérivés de pyrazoline 5b, 6a, 7a avec l'atome de Cl retirant les électrons ont montré une augmentation remarquable de l'effet inhibiteur de CDK9, 2, 1 par rapport à la forme hydrazone ouverte 3b.
- Tous les composés cibles pour leurs activités anticancéreuses, sauf dix composés ont montré une activité remarquable lorsqu'ils ont été testés contre la lignée cellulaire de cancer du sein MCF-7 (3a, 6b étaient les agents les plus puissants IC50 = 0,21 nM).

- L'effet de la fraction hydrazone sur l'activité anticancéreuse a été clairement observé contre la lignée cellulaire HeLa,3a-d présentait les valeurs IC50 les plus basses (IC50 = 2,3-11,0 nM). [74]

### e. Conclusion

Le model QSAR obtenu montre la corrélation importante des descripteurs 2D avec l'activité inhibitrice de CDK, ainsi l'importance de fraction hydrazone sur l'activité et la sélectivité avec des résultats supérieurs à celle de flavopiridol.

Le model choisi est hautement prédictif et donne des résultats et des informations structurelles très intéressantes qui peut être guidé d'autre recherche, sur les médicaments anticancéreux.

### 1.2.1.2. Activité antioxydante in vitro de certains dérivés de 4-hydroxy chromène-2-one : étude SAR et QSAR

#### a. Données expérimentales

Dans l'étude réalisée par Milan Mladenović et al [75], sur l'activité antioxydante des dérivés de coumarines. Un ensemble de 15 molécules dérivés de 4-hydroxycoumarine synthétisés (figure 18) ont été testés in vitro, on examine la capacité antioxydante totale, le 1,1-diphényl-2-picrylhydrazyle (DPPH), le radical hydroxyle, le piégeage des peroxydes et l'activité chélatante, les résultats sont comparés aux valeurs standards de l'acide ascorbique (Asc) et du butylhydroxytoluène (BHT).

Le potentiel de structure de nos dérivés a été confirmé par les études SAR, QSAR qui ont été utilisées pour la conception réussie de 20 nouvelles structures antioxydantes coumariniques améliorées.

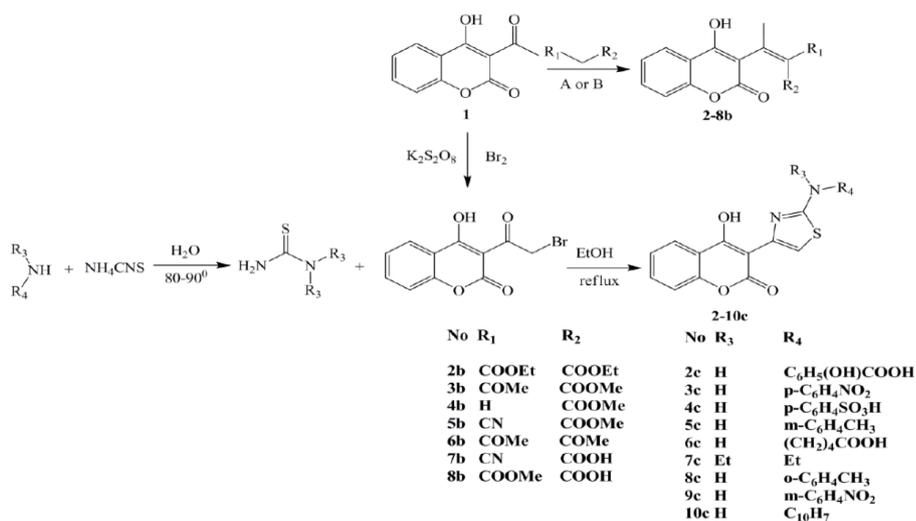


FIGURE 18 : Dérivés de coumarine synthétisés. [75]

**b. Descripteurs moléculaires et méthode statistique**

▪ **Descripteurs moléculaires**

Les descripteurs moléculaires générés pour cette étude étaient électroniques et stériques. Le facteur de solubilité a été pris en compte pour l'environnement de peroxydation des lipides.

- Les descripteurs moléculaires électroniques obtenus après optimisation étaient : énergies HOMO et LUMO,  $\Delta E$  HOMO-LUMO calculés comme la différence des énergies, dipôle, énergie totale, énergie électronique, longueur de liaison O-H, avec l'analyse de population Mulliken appliquée pour le calcul de la charge atomique (Q).
- Les descripteurs moléculaires stériques étaient : Connolly Accessible Area (CAA), Connolly Molecular Area (CMA), Connolly Solvent-Excluded Volume (CSEV) et le moment d'inertie principal par VEGA ZZ, comme descripteur moléculaire utile.

▪ **Méthode statistique : Régression linéaire multiple (MLR)**

Les valeurs logarithmiques de l'activité biologique ont été utilisées pour la génération des équations QSAR avec des descripteurs moléculaires pertinents.

La génération des modèles QSAR a été effectuée avec Origin Pro 8, par l'analyse MLR (Multiple Linear Regression), les résultats de l'analyse de régression sont présentés dans les équations (1), (2), (3), (4) et (5).

- **Étude QSAR sur la capacité antioxydante totale :**

$$pTAC50 = -8.46(\pm 0.32) + 7.32(\pm 0.99) \text{ HOMO} - 2.69(\pm 0.47) \text{ LUMO} - 9.46(\pm 1.48) \text{ Q0H} \\ + 0.6(\pm 0.1) \text{ QN} - 2.9(\pm 0.4) \text{ QS} \quad (1)$$

**TAC** : activité antioxydant totale

- **Étude QSAR sur l'activité de piégeage DPPH in vitro 60 min :**

$$pIC50 = +6.72(\pm 0.48) - 1.23(\pm 0.57) \text{ Q0H} + 4.42(\pm 0.35) \text{ 4-OH bond length} + 3.4(\pm 0.94) \\ \text{QN} - 7.6(\pm 0.26) \text{ QS} \quad (2)$$

- **Étude QSAR sur l'inhibition de la peroxydation lipidique dans l'émulsion d'acide linoléique (24 h) :**

$$pI50 = +8.46(\pm 0.47) + 2.46(\pm 0.28) \log P - 4.92(\pm 0.36) \text{ HOMO} + 1.96(\pm 0.46) \text{ 4-OH bond} \\ \text{length} - 6.49(\pm 0.76) \text{ QS} \quad (3)$$

- **Étude QSAR sur l'activité de piégeage des radicaux hydroxyles in vitro :**

$$pOH50 = +11.46(0.94) - 4.94(0.32) \text{ HOMO} + 7.46(0.24) \text{ LUMO} + 5.96(0.32) \text{ Q0H} \quad (4)$$

- **Étude QSAR sur la capacité de chélation du fer**

$$pCE50 = +1.29(0.52) + 9.01(0.003) \text{ CSEV} + 0.95(0.33) \text{ QN} \quad (5)$$

Tous les résultats ont été exprimés en moyenne  $\pm$  écart type (SD). La signification de la différence a été calculée par un test ANOVA unidirectionnel et les valeurs  $<0,05$  ont été considérées comme significatives.

**c. Validation du modèle QSAR :**

Pour confirmer le pouvoir prédictif de notre modèle QSAR, une méthode de validation croisée (leave-one-out) a été utilisée pour calculer les paramètres suivants :

- n : le nombre de molécules.
- r : la corrélation.
- F : le facteur de signification de Fisher.
- s : l'écart-type.
- Q<sup>2</sup> : carré du pouvoir prédictif du coefficient.
- s-PRESS: somme résiduelle prédictive des carrés.

**d. Résultats et discussions**

▪ **Résultats**

- Les descripteurs obtenus après optimisation ont montrés une corrélation importante avec l'activité, bien que cet ensemble soit petit mais il fournit des équations QSAR qui sont statistiquement significatives. Les rapports des valeurs d'activité antioxydante observées et calculées et la corrélation entre les descripteurs moléculaires et l'activité sont présentés dans le tableau XII :
- Le tableau XII Valeurs observées par rapport aux valeurs calculées de l'activité antioxydante des dérivés de la coumarine provenant des études QSAR

**TABLEAU X : Valeurs observées par rapport aux valeurs calculées de l'activité antioxydante des dérivés de la coumarine provenant des études QSAR. [75]**

| Mole | pTAC50 |       | pIC50  |       | pI50   |       | pOH50  |       | pCE50  |       |
|------|--------|-------|--------|-------|--------|-------|--------|-------|--------|-------|
|      | Obser. | Calc. |
| 1    | 4.01   | 3.89  | 4.06   | 4.26  | 4.57   | 4.78  | 4.71   | 4.80  | 3.32   | 3.42  |
| 2b   | 4.32   | 4.71  | 5.34   | 5.43  | 5.12   | 5.18  | 4.76   | 4.72  | 4.34   | 4.37  |
| 3b   | 370    | 3.69  | 5.06   | 5.11  | 4.92   | 5.14  | 4.49   | 4.62  | 4.24   | 4.36  |
| 4b   | 4.07   | 4.88  | 5.00   | 5.06  | 4.72   | 4.82  | 4.06   | 4.13  | 4.20   | 4.44  |
| 6b   | 4.29   | 4.28  | 5.61   | 5.65  | 5.41   | 5.41  | 4.84   | 5.02  | 4.54   | 4.44  |
| 7b   | 4.45   | 4.76  | 3.87   | 3.75  | 4.12   | 4.18  | 5.00   | 5.13  | 4.22   | 4.52  |
| 8b   | 4.00   | 3.89  | 5.95   | 4.96  | 4.94   | 5.08  | 4.42   | 4.47  | 4.26   | 4.36  |
| 2c   | 4.47   | 4.91  | 5.16   | 5.25  | 5.17   | 5.48  | 5.00   | 5.23  | 4.46   | 4.52  |
| 3c   | 3.88   | 3.92  | 5.03   | 5.13  | 4.60   | 4.92  | 4.70   | 4.86  | 5.26   | 5.31  |
| 4c   | 4.76   | 5.01  | 5.4    | 5.69  | 5.41   | 5.61  | 5.23   | 5.55  | 4.36   | 4.26  |
| 5c   | 4.27   | 4.14  | 4.12   | 4.06  | 4.66   | 4.63  | 4.15   | 4.26  | 3.82   | 3.34  |
| 6c   | 4.14   | 4.08  | 4.63   | 4.80  | 4.47   | 4.73  | 4.52   | 4.72  | 4.21   | 4.21  |
| 7c   | 4.33   | 4.36  | 4.22   | 4.42  | 4.28   | 4.38  | 4.43   | 4.44  | 4.26   | 4.26  |
| 8c   | 4.32   | 4.445 | 5.32   | 5.68  | 4.48   | 4.56  | 4.16   | 4.22  | 4.28   | 4.45  |
| 9c   | 3.86   | 3.15  | 4.12   | 4.12  | 5.23   | 5.53  | 4.61   | 4.71  | 5.11   | 5.34  |
| 10c  | 3.66   | 3.49  | 4.06   | 4.26  | 4.55   | 4.72  | 4.12   | 4.32  | 4.29   | 4.26  |

- Les résultats de validation de chaque équation obtenue dans le modèle QSAR sont les suivants :

➤ **Étude QSAR sur la capacité antioxydante totale**

$n = 15$ ;  $r = 0,992$ ;  $s = 0,032$ ;  $F = 37,426$ ;  $Q2 = 0,993$ ;  $s$ -PRESS = 0,004

➤ **Étude QSAR sur l'activité de piégeage DPPH in vitro 60 min**

Le signe positif avant le 6.72 dans l'équation (2) suggère que le modèle est hautement spécifique vis-à-vis du piégeage de DPPH.

➤ **Étude QSAR sur l'inhibition de la peroxydation lipidique dans l'émulsion d'acide linoléique (24 h) :**

$n = 15$ ;  $r = 0,993$ ;  $s = 0,046$ ;  $F = 103,981$ ;  $Q2 = 0,991$ ;  $s$ -PRESS = 0,007

➤ **Étude QSAR sur l'activité de piégeage des radicaux hydroxyles in vitro**

$n = 15$ ;  $r = 0,997$ ;  $s = 0,028$ ;  $F = 174,273$ ;  $Q2 = 0,994$ ;  $s$ -PRESS = 0,013

➤ **Étude QSAR sur la capacité de chélation du fer**

$n = 15$ ;  $r = 0,993$ ;  $s = 0,054$ ;  $F = 43,383$ ;  $Q2 = 0,894$ ;  $s$ -PRESS = 0,034

▪ **Discussion**

- Selon le modèle qui décrit la capacité antioxydante totale, équation (1), le potentiel des composés dépend fortement des énergies électroniques HOMO et de la charge atomique partielle de l'azote du cycle thiazole (L'équation décrit le centre thiazole comme riche en électrons et participant à la réduction du Mo (VI)).
- La basicité élevée de l'atome d'azote thiazole atténue la force de la liaison C-O du groupe 4-OH, rendant l'atome d'hydrogène plus acide et plus facilement relâché (figure 18). Par conséquent, la basicité de l'azote facilite le piégeage des radicaux DPPH, quantifié par l'équation (2).
- Les équations (2) et (3) caractérisent le facteur de la longueur de la liaison qui est importante pour soulager facilement l'atome d'hydrogène vers le radical libre.
- La capacité des composés à inhiber la réaction en chaîne de peroxydation dépend fortement de la solubilité de la coumarine.
- Une conjugaison continue à travers le système coumarine 1-10c, et la double liaison d'échafaudage C-3 permet la possibilité d'une addition de radicaux hydroxyles sur le niveau orbitalaire de la coumarine LUMO.
- La forte participation de l'énergie LUMO et de la charge partielle d'oxygène du groupe 4-hydroxyle dans l'équation hautement spécifique (4) est en corrélation avec l'activité élevée de nos dérivés.
- Dans la structure des dérivés de coumarine testés, il existe des échafaudages volumineux qui définissent la stéréochimie du composé pendant la chélation de l'atome de fer. Le modèle (équation 5) spécifie la grande importance de l'arrangement spatial des résidus C-3 sur l'activité.

**e. Conclusion**

En fonction des conditions SAR, QSAR et des conditions expérimentales basées sur des études théoriques et sur les résultats expérimentaux présentés par la capacité antioxydante totale et l'affinité pour le DPPH, les peroxydes lipidiques et les radicaux hydroxyles, il a été conclu que les composés 2b, 6b, 2c et 4c sont des antioxydants notables qui pourraient être appliqués en pratique comme agents antioxydants et comme composés de départ pour de futures modifications sélectives de la molécule de coumarine. [75].

### 1.2.2. Substitution en C4 et C7

#### 1.2.2.1. Etude de la QSAR de l'activité inhibitrice de l'aromatase des dérivés de coumarines di substitué

##### a. Données expérimentales

Dans le cadre des recherches sur les cibles thérapeutiques potentielles pour le traitement de cancer de sein. **Apilak Worachartcheewan et al [76]**, étudiaient la relation quantitative structure-activité (QSAR) d'une série de dérivés de coumarine 3,7- et 4,7-disubstitués (1-34) avec des substituants R1 et R2 (**cf annexe 3**) portant une activité inhibitrice de l'aromatase. Ces derniers ont été modélisés en fonction de descripteurs moléculaires et quantiques dérivés du conformère de basse énergie.

La figure 19 présente la structure chimique générale des coumarines 3,7- et 4,7-disubstitués dont QSAR a été étudiée par Apilak Worachartcheewan et al.

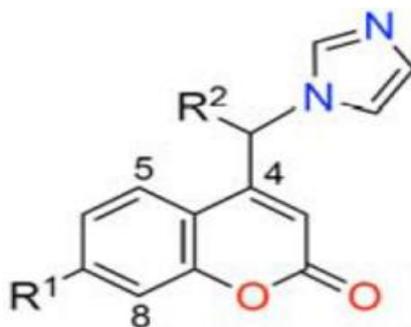


FIGURE 19 : Structure chimique générale des coumarines 3,7- et 4,7-disubstitués dont QSAR a été étudiée par Apilak Worachartcheewan et al. [76]

Ces dérivés sont tirées du travail de **Stefanachi et al. (2011)** où l'activité inhibitrice (IC : concentration inhibitrice) de l'aromatase a été déterminée in vitro en utilisant des microsomes placentaires humains comme source d'aromatase tandis que la [1 $\beta$ -3H] androstènedione a été utilisée comme substrat.

##### b. Descripteurs moléculaires et méthode statistique

###### ▪ Descripteurs moléculaires

Un ensemble initial de 3224 descripteurs moléculaires, obtenus à partir du logiciel Dragon version 5.5 (Talet, 2007), a été soumis à une élimination sélective. Les descripteurs ainsi sélectionnés, sont ensuite combinés avec un ensemble de 6 descripteurs chimiques quantiques. Enfin, ceux qui sont les plus corrélés avec l'activité sont déterminés par la régression multiple linéaire (MLR).

▪ **Méthode statistique : régression linéaire multiple (MLR)**

La régression linéaire multiple (MLR) a été employée pour construire des modèles QSAR comme conceptuellement résumée par l'équation suivante :

$$Y = m_1X_1 + m_2X_2 + \dots + m_nX_n + b$$

Où Y : l'activité inhibitrice prédictive de l'aromatase pour chaque échantillon (pIC50)

m : la valeur du coefficient de régression des descripteurs.

x : le descripteur.

b : la valeur d'ordonnée à l'origine.

Les calculs MLR utilisés sont mis en œuvre par le logiciel WKAKA (Waikato Environment for Knowledge Analysis), version 3.4.5 (Witten et al., 2011).

Une telle analyse multivariée a généré un total de 4 modèles constitués d'équations MLR et d'analyses statistiques.

**c. Validation du modèle QSAR**

Pour confirmer le pouvoir prédictif de notre modèle QSAR, l'ensemble des données a été divisé en deux principaux sous-groupes :

- Un ensemble externe : obtenu en sélectionnant au hasard 15% de l'ensemble des échantillons (au nombre de 34), la validation s'effectue par le calcul du paramètre : Q2 Ext- RMSE Ext.
- L'ensemble interne : constitué de 85% des échantillons restants, la validation interne s'effectue par la méthode LOO-CV en calculent les paramètres suivants : (R2Tr), (RMSETr), (Q2 CV), (RMSECV).

De plus, le ratio de Fisher (F) ainsi que la différence entre R2 et Q2 (R2-Q2) ont été utilisés pour évaluer la qualité des modèles construits.

**d. Résultats et discussions**

▪ **Résultats**

- L'ensemble des descripteurs générés par le logiciel Dragon plus les descripteurs moléculaires ont été soumis à une MLR pour chaque cycle de sélection, il en résulte, à la fin, 7 descripteurs significatifs (cf. tableau XII). Les résultats ont indiqué que ces descripteurs sont indépendants l'un de l'autre et seront utilisés après pour construire le modèle QSAR.

**Tableau XI : Définition des descripteurs significatifs pour le modèle QSAR [76]**

| Symbole             | Définition   | Type                                    |
|---------------------|--|---|
| <b>F10[N-O]</b>     | Fréquence de N-O à la distance topologique<br>10   | Empreintes digitales en<br>fréquence 2D |
| <b>Inflammat-50</b> | Indice anti-inflammatoire de Ghose-<br>Viswanadhan-Wendoloski à 50%  | Propriétés moléculaires                 |
| <b>Psychotic-80</b> | Indice anti-inflammatoire de Ghose-<br>Viswanadhan-Wendoloski à 50%  | Propriétés moléculaires                 |
| <b>H-047</b>        | H attaché à C1 (sp3) / C0 (sp2)  | Fragments centrés sur l'atome           |
| <b>BELe1</b>        | Valeur propre la plus basse n.1 de la matrice<br>de Burden / pondérée par l'électronégativité<br>atomique de Sanderson | Fardeau des valeurs propres             |
| <b>B10[C-O]</b>     | Présence / absence de C-O à distance<br>topologique 10   | Empreintes binaires 2D                  |
| <b>MAXDP</b>        | Variation positive électro topologique<br>maximale   | Descripteurs topologiques               |

- Le tableau XII montre les différents descripteurs moléculaires et les valeurs des activités inhibitrices de l'aromatase expérimentale et prédite (pIC50) des analogues de la coumarine (1-34) obtenues à partir du modèle 4.
- Les résultats suggèrent des bonnes performances prédictives pour le modèle (4), qui a fourni  $Q2CV = 0,9239$ ,  $RMSECV = 0,1104$  et  $Q2Ext = 0,7268$ .
- Le modèle a été sélectionné pour une étude plus poussée en raison de l'absence de composés périphériques et de sa capacité à bien fonctionner sur les ensembles interne et externes.
- R2-Q2 calculé à partir des ensembles d'apprentissage et externes a donné une valeur  $<0,3$ , ce qui suggère la fiabilité du modèle.

## Chapitre III : Etude de la relation structure activité quantitative QSAR sur différents dérivés de coumarines

**TABLEAU XII : Les résultats calculés partir de modèle 4 [76]**

| Cpd | Descripteurs |                  |                   |           |       |                  |       | active                   |                          |                             | Rési-<br>dual |
|-----|--------------|------------------|-------------------|-----------|-------|------------------|-------|--------------------------|--------------------------|-----------------------------|---------------|
|     | F10[N-<br>O] | Inflammat-<br>50 | Psychotique<br>80 | H-<br>047 | BELe1 | B10<br>[C-<br>O] | MAXDP | IC <sub>50</sub><br>(μM) | pIC <sub>50</sub><br>exp | pIC <sub>50</sub><br>prédit |               |
| 1   | 0            | 0                | 1                 | 13        | 1.920 | 1                | 4.891 | 0.150                    | 6.824                    | 6.814                       | -<br>0.010    |
| 2   | 0            | 0                | 1                 | 12        | 1.922 | 1                | 4.928 | 0.114                    | 6.943                    | 6.691                       | -<br>0.252    |
| 3   | 0            | 0                | 1                 | 12        | 1.920 | 1                | 5.234 | 0.113                    | 6.947                    | 6.87                        | -<br>0.075    |
| 4   | 0            | 0                | 1                 | 12        | 1.920 | 1                | 4.914 | 0.130                    | 6.886                    | 6.737                       | -<br>0.149    |
| 5   | 0            | 0                | 1                 | 12        | 1.921 | 1                | 4.916 | 0.235                    | 6.629                    | 6.734                       | -<br>0.105    |
| 6   | 0            | 0                | 0                 | 12        | 1.920 | 1                | 4.937 | 0.207                    | 6.684                    | 6.322                       | -<br>0.362    |
| 7   | 0            | 0                | 1                 | 12        | 1.922 | 1                | 4.929 | 0.141                    | 6.851                    | 6.701                       | -<br>0.150    |
| 8   | 0            | 0                | 1                 | 12        | 1.920 | 1                | 4.947 | 0.267                    | 6.573                    | 6.771                       | -<br>0.198    |
| 9   | 0            | 0                | 1                 | 12        | 1.920 | 1                | 4.910 | 0.178                    | 6.750                    | 6.734                       | -<br>0.016    |
| 10  | 0            | 0                | 1                 | 15        | 1.920 | 1                | 4.940 | 0.127                    | 6.896                    | 7.059                       | -<br>0.163    |
| 11  | 0            | 0                | 0                 | 12        | 1.920 | 1                | 4.930 | 0.481                    | 6.318                    | 6.319                       | -<br>0.001    |
| 12  | 0            | 1                | 0                 | 11        | 1.923 | 1                | 4.870 | 0.169                    | 6.772                    | 7.70                        | -<br>0.398    |
| 13  | 0            | 0                | 1                 | 11        | 1.920 | 1                | 5.271 | 0.165                    | 6.783                    | 6.82                        | -<br>0.037    |
| 14  | 0            | 1                | 0                 | 11        | 1.923 | 1                | 4.780 | 0.051                    | 7.892                    | 7.139                       | -<br>0.153    |
| 15  | 0            | 1                | 0                 | 10        | 1.923 | 1                | 5.262 | 0.072                    | 7.143                    | 7.084                       | -<br>0.059    |
| 16  | 0            | 1                | 0                 | 10        | 1.923 | 1                | 4.897 | 0.072                    | 7.143                    | 7.289                       | -<br>0.146    |
| 17  | 0            | 0                | 1                 | 13        | 1.924 | 1                | 4.937 | 0.292                    | 6.535                    | 6.778                       | -<br>0.243    |
| 18  | 0            | 0                | 1                 | 10        | 1.926 | 1                | 4.907 | 0.690                    | 6.161                    | ND                          | -             |
| 19  | 0            | 1                | 0                 | 10        | 1.923 | 1                | 4.893 | 0.112                    | 6.951                    | 7.120                       | -<br>0.169    |

## Chapitre III : Etude de la relation structure activité quantitative QSAR sur différents dérivés de coumarines

|    |   |   |   |    |       |   |       |       |       |       |       |
|----|---|---|---|----|-------|---|-------|-------|-------|-------|-------|
| 20 | 1 | 1 | 0 | 10 | 1.925 | 1 | 4.910 | 0.164 | 6.785 | ND    | -     |
| 21 | 0 | 1 | 1 | 10 | 1.926 | 1 | 4.939 | 0.296 | 6.529 | 6.469 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.060 |
| 22 | 0 | 0 | 1 | 16 | 1.925 | 1 | 4.967 | 0.081 | 7.092 | 7.012 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.080 |
| 23 | 0 | 0 | 0 | 9  | 1.923 | 1 | 5.283 | 0.070 | 7.155 | 7.189 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.034 |
| 24 | 0 | 0 | 0 | 9  | 1.923 | 1 | 5.299 | 0.047 | 7.328 | 7.146 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.182 |
| 25 | 0 | 0 | 0 | 11 | 1.935 | 1 | 4.890 | 0.105 | 6.979 | 7.013 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.034 |
| 26 | 0 | 0 | 0 | 6  | 1.914 | 0 | 4.471 | 3.750 | 5.426 | 5.666 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.240 |
| 27 | 0 | 0 | 0 | 9  | 1.916 | 0 | 4.600 | 0.280 | 6.553 | ND    | -     |
| 28 | 0 | 0 | 1 | 13 | 1.955 | 1 | 5.177 | 0.455 | 6.342 | 6.481 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.139 |
| 29 | 0 | 0 | 1 | 15 | 1.958 | 1 | 5.447 | 0.067 | 7.174 | ND    | -     |
| 30 | 0 | 0 | 1 | 13 | 1.957 | 1 | 5.503 | 0.317 | 6.499 | 6.583 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.084 |
| 31 | 0 | 0 | 0 | 14 | 1.957 | 1 | 5.480 | 0.532 | 6.274 | 6.095 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.179 |
| 32 | 2 | 0 | 1 | 14 | 1.962 | 1 | 5.504 | 4.010 | 5.397 | 5.207 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.190 |
| 33 | 1 | 0 | 0 | 9  | 1.906 | 1 | 4.918 | 2.820 | 5.550 | 5.744 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.194 |
| 34 | 1 | 0 | 1 | 15 | 1.945 | 1 | 5.972 | 0.313 | 6.504 | 6.530 | -     |
|    |   |   |   |    |       |   |       |       |       |       | 0.026 |

### ▪ Discussion

- Dans la série de 7-benzyloxy-coumarine (1-13, R1 = CH<sub>2</sub>C<sub>6</sub>H<sub>5</sub>, R2 = H), où les groupes phényle (R1) ont été substitués par méta(m) - et para(p)-Me, F, Cl, OMe, OCF<sub>3</sub> et NO<sub>2</sub>, les résultats ont conclu à ce que les substituants méta tels que F (3), Cl (4), OCF<sub>3</sub> (6) présentent une activité inhibitrice de l'aromatase plus élevée que leurs équivalents substituants p correspondants (8, 9, 11). Le composé m-F (3) s'est avéré être l'inhibiteur d'aromatase le plus puissant.
- Il a été noté que pour une série de 7-phénoxy-coumarines (14-24, R1 = C<sub>6</sub>H<sub>5</sub>, R2 = H), le 3',4'-difluorophényle dissymétrique (24) avait une valeur MAXDP supérieure à celle de l'analogue 3', 5'-difluoro symétrique (23).

- Les composés 29 et 30, qui contiennent un effet stérique à partir du groupe phényle sur le carbone méthine C-4, pourraient empêcher sa liaison ou interaction avec le site d'action cible.
- La distance appropriée entre R1 et la structure du noyau de coumarine dans laquelle le phénoxy (R1 = phényle) de l'analogue de coumarine a donné le meilleur ajustement dans l'interaction avec le site cible.
- La propriété lipophile des substituants R1 joue un rôle crucial dans l'obtention des valeurs élevées appropriées de MAXDP et BELe1 qui expliquent l'effet inhibiteur puissant de l'aromatase.
- L'application de la règle des cinq de Lipinski sur cet ensemble de composés a révélé sa conformité dans laquelle il y avait moins de 5 donneurs de liaison Hydrogène moins de 10 accepteurs de liaison Hydrogène, poids moléculaire (MW) inférieur à 500 Da et Ghose-Crippen Coefficient de partage octanol-eau (ALogP) inférieur à 5.
- C'est ainsi que la puissance de cet ensemble de composés est relativement robuste comme déduit les valeurs de CI50.

**e. Conclusion**

A partir de cette étude sur l'activité inhibitrice de l'aromatase des dérivés d'imidazole des coumarines 3,7- et 4,7-disubstituées ayant des substituants R1 et R2 et sur la lumière de model QSAR étudié (bonnes performances prédictives) utilisant la méthode MLR, des relations structure activité ont été discutées et suggèrent que l'activité inhibitrice de l'aromatase des composés dépend principalement des propriétés lipophiles et de la position du substituant (R1) sur la structure du noyau de la coumarine et de l'effet électronique des groupes fonctionnels sur les formes résonantes de charge ionique. Les connaissances structurelles acquises par le modèle QSAR pourraient être utilisées pour guider la conception rationnelle de nouveaux inhibiteurs de l'aromatase [76].

### 1.2.3. Substitution en C7 et C8

#### 1.2.3.1. Études sur le profil antiplaquettaire et antithrombotique du dérivé anti-inflammatoire des dérivés de coumarines 7-hydroxy-8-substitués et dérivés de la coumarine liée à la 7-azométhine.

##### a. Données expérimentales

Dans cette étude faite par **Christos Kontogiorgis et al [77]**, deux séries des dérivés de la coumarine sont déjà synthétisées : bases coumarine de Mannich (7-hydroxy-8 substitués compose 1–11) et dérivés de la coumarine liée à la 7-azométhine (12–19) (voir l'annexe 04).

- Les différentes molécules soumises à des essais biologiques (manipulation du sang
- Agrégation plaquettaire- la viabilité des plaquettes - Rétraction du caillot
- Amélioration du niveau d'AMPc dans des cellules SK-N-MC
- Viabilité cellulaire) où la plupart montrent des activités anti-inflammatoires puissantes sans causer d'ulcérogénicité, activité inhibitrice lipoxigénase. Toutes ces données biologiques soutenues par études QSAR sont exploitées pour étudier l'activité antiplaquettaire et antithrombotique.

##### b. Descripteurs moléculaires et méthode statistique

###### ▪ Descripteurs moléculaires

Les descripteurs 2D étudiés sont obtenus en utilisant le programme C-QSAR, ces derniers sont ajustés et affinés par analyses de multi-régression (MRA) pour déterminer les plus corrélés avec l'activité.

Les descripteurs 3D étudiés :

- Les conformations d'énergie les plus basses de tous les composés synthétisés sont identifiées par l'analyse de conformation de Monte Carlo (MMFF94).
- Les énergies des conformères les plus stables ont été calculées au niveau théorique HF / 6-31G.
- Le potentiel électrostatique et les valeurs énergétiques HOMO et LUMO.

###### ▪ Méthode statistique : analyses de multi-régression (MRA)

Dans cette étude QSAR, ils ont utilisé l'analyse de multi-régression (MRA) pour concevoir le modèle le plus ajusté avec les paramètres statistiques des régressions par l'application de l'équation multivariée suivante :

$$\text{Log1/IC50} = 0.465(\pm 0.212)\text{ClogP} - 22.755(\pm 8.401) E_{\text{LUMO}} + 4.380 (\pm 0.513)$$

ClogP : exprime la lipophilie de la molécule entière.

$E_{LUMO}$  : est un paramètre d'électrophilie.

### c. Validation du modèle QSAR

La validation de modèle a été réalisée par le calcul des paramètres suivants :

**n** : représente le nombre de points de contact.

**r** : coefficient de corrélation entre valeurs observées et calculées.

**s** : écart type.

**q<sup>2</sup>**: carré de coefficient de corrélation. (Calculé à l'aide de procédure leave-one-out)

### d. Résultats et discussions

#### ▪ Résultats

- Les résultats des essais biologiques sont montrés dans le tableau XIV :
- Les études QSAR de l'activité antiplaquettaire de la coumarine induite par l'acide arachidonique ont montré que la lipophilie et  $E_{LUMO}$  étaient les descripteurs les plus significatifs et jouaient un rôle important pour modéliser les données biologiques.
- Les résultats des tests de validation donnaient une bonne performance prédictive pour le modèle :

$$n=9, r=0.948, q^2=0.750, s=0.152.$$

- Le tableau XIII Activité des composés étudiés dans l'ADP (3 mm) ou l'acide arachidonique (50 mm) induite par l'agrégation plaquettaire (cochon d'Inde) et dans l'essai de rétraction des caillots (rat).

**TABLEAU XIII : Activité des composés étudiés dans l'ADP (3  $\mu$ M) ou l'acide arachidonique (50  $\mu$ M) induite par l'agrégation plaquettaire (cochon d'Inde) et dans l'essai de rétraction des caillots (rat) [77].**

|     | <u>Agrégation plaquettaire IC50</u> |                   | Rétraction du caillot IC50 | Plaquette de cochon      |
|-----|-------------------------------------|-------------------|----------------------------|--------------------------|
|     | <u>(mM)</u>                         |                   | (mM)                       | d'Inde viabilité (1 h) % |
|     | ADP(3 $\mu$ M)                      | AA(5mM)           | Thrombine (20U/ml)         | Test PNPP (500mM)        |
| 1   | Inac                                | Inac              | Inac                       | 100                      |
| 2   | Inac                                | Inac              | 398 (292-543)              | 100                      |
| 3   | 390 (240-630)                       | 193 (161-232)     | 129 (93-179)               | 74                       |
| 4   | Inac                                | Inac              | Inac                       | 100                      |
| 5   | 385 (243-610)                       | 335 (210-534)     | 146 (84-255)               | 87                       |
| 6   | Inac                                | Inac              | 405 (136-239)              | 100                      |
| 7   | Inac                                | Inac              | Inac                       | 100                      |
| 8   | Inac                                | Inac              | 180 (136-239)              | 100                      |
| 9   | Inac                                | 299 (243-369)     | Inac                       | 100                      |
| 10  | Inac                                | Inac              | Inac                       | 100                      |
| 11  | Inac                                | Inac              | Inac                       | 100                      |
| 12  | 252 (150-327)                       | 138 (117-170)     | Inac                       | 100                      |
| 13  | 337 (286-496)                       | 103 (48-222)      | 131 (109-157)              | 100                      |
| 14  | 135 (120-154)                       | 100 (69-143)      | Inac                       | 100                      |
| 15  | 232 (194-277)                       | 70 (61-49)        | Inac                       | 100                      |
| 16  | Inac                                | 19.9 ( 12.1-32.7) | Inac                       | 100                      |
| 17  | Inac                                | Inac              | Inac                       | 100                      |
| 18  | >500                                | 275 (190-401)     | 192 (128-290)              | 100                      |
| 19  | 363 (264-498)                       | 253 (205-314)     | 428 (271-675)              | 100                      |
| ASA | >500                                | 55 (39-77)        | Inac                       | 100                      |

▪ **Discussion**

- Dans le groupe des bases de Mannich de coumarine (1-11), les composés 3 et 5 se sont avérés légèrement actifs avec des valeurs d'IC50 égales à 390 et 385 mM, respectivement. On peut en déduire qu'une lipophilie modérée (Clog P près de 2,5) peut jouer un rôle dans cette activité.
- Le second groupe de composés, y compris 12-19, l'agrégation plaquettaire induite par l'ADP s'est avérée réduite de manière significative. Parmi eux, le composé 14,

l'agent anti-inflammatoire le plus puissant, qui suggère que l'activité antiplaquettaire est améliorée par la présence d'un groupe oxime en position 7.

- En utilisant AA comme inducteur, les composés 3, 5 et 11 se sont révélés être les plus actifs. Parmi eux, le composé 3 est le plus actif et ceci pourrait probablement être attribué à sa valeur de réfractivité molaire, la plus élevée.
- Dans l'autre groupe le composé 16, c'est-à-dire celui ayant la teneur lipophile la plus élevée, s'est avéré être le plus actif (IC<sub>50</sub>=19.9mM) et, plus important encore, plus actif que l'acide acétylsalicylique (IC<sub>50</sub>=55mM) utilisé comme référence.

#### **e. Conclusion**

Les résultats précédents et études QSAR nous amènent à conclure que les composés étudiés et leur puissance multi-cible pourrait également être utilisés pour concevoir de nouvelles séries d'agents antiplaquettaires et anti thrombotiques prometteurs. [77]

### **1.3. Coumarine simple tri substituée**

Deux exemples ont été traités : substitution en C3, C4 et C6 et substitution en C3 et C6 et C7

#### **1.3.1. Substitution en C3 et C4 et C6**

**1.3.1.1. Evaluation de l'activité antituberculeuse chez des dérivés des coumarines hybridés avec un noyau de pyrimidine (phényl-2-thioxo-3,4-dihydro-1H-chromino [4,3-d] pyrimidin-5-one et de 4- (substitué) phényl-3,4 substitué (substitué) -dihydro-1H-chromino [4,3-d] pyrimidine-2,5-dione).**

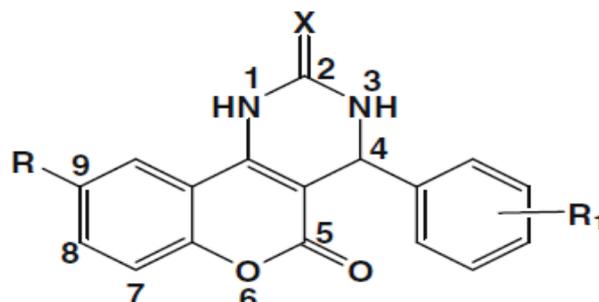
#### **a. Données expérimentales**

Le travail actuel s'inscrit dans la poursuite de recherche des nouveaux agents antituberculeux, où l'hypothèse de base est qu'une fusion des noyaux coumarine[coumarines possèdent une activité antituberculeuse prouvée par des études 3D-QSAR antérieures (Manvar et al., 2008)] et pyrimidine-2,5-dione [benzopyranopyrimidines sont intéressantes en raison de leurs activités antifongiques et antibactériennes (Kidwai et al., 2006)] et dont les deux noyaux se sont montrés actifs contre la tuberculose, donnerait de puissants agents antituberculeux (effet synergie).

Sur la base des informations contenues dans les cartes CoMFA rapportées dans les études précédentes 3D-QSAR sur les analogues de la coumarine (Manvar et al., 2008; Virsdoia et al., 2010), **Premlata K. Ambreet al [78]**, ont synthétisé et testé 16 variantes moléculaires contenant 4- (substitué) phényl-3,4-dihydro-1H-chromino [4,3-

d] pyrimidine-2,5-diones et 4- (substitué) phényl-2-thioxo-3,4- les noyaux dihydro-1H-chromino [4,3-d] pyrimidin-5-ones et 2-thiopyrimidone.

- La figure 20 représente la structure étudiée



Class A: 4a-h  
Class B: 5a-h

FIGURE 20 : Structure principale de 4- (substitué) phényl-3,4-dihydro-1H-chromino [4,3-d] pyrimidine-2,5-diones et 4- (substitué) phényl-2-thioxo-3,4- les noyaux dihydro-1H-chromino [4,3-d] pyrimidin-5-ones et 2-thiopyrimidone [78]

L'activité antituberculeuse des composés synthétisés a été évaluée par des essais et des dosages biologiques (bleu Alamar en microplaque (MABA) et dosage par LORA (Low Oxygen-Recovery Assay) ainsi que leur cytotoxicité mesurée in vitro.

Les tests sont effectués sur une souche de *Mycobacterium tuberculosis* H37Rv et la rifampicine a été considérée comme molécule de référence.

La relation structure activité est expliquée par un modèle de partitionnement récursif (RP). Deux modèles RP différents ont été construits, l'un basé sur l'activité antituberculeuse et l'autre basé sur la toxicité des molécules.

L'arbre de décision pourrait identifier des descripteurs discriminants actifs et inactifs ainsi que des analogues de la coumarine 3,4-annelés toxiques et moins toxiques. Ce modèle RP sera utilisé dans d'autres travaux pour concevoir des molécules plus puissantes.

#### b. Descripteurs moléculaires et méthode statistique

##### ▪ Descripteurs moléculaires

Des différents descripteurs sont utilisés pour construire le modèle RP. Le logiciel de modélisation moléculaire Cerius 2 a été utilisé pour calculer les descripteurs 2D des molécules.

Le tableau XIV représente descripteurs utilisés pour développer les modèles RP

**Tableau XIV : Descripteurs utilisés pour développer les modèles RP [78]**

| <b>Descripteur</b>            | <b>Définition</b>   |
|-------------------------------|---|
| Descripteur de structure      | Le nombre de donneurs et d'accepteurs de liaisons H, le nombre de liaisons rotatives, le poids moléculaire                    |
| Descripteurs topologiques     | Indice de Wiener (W), indice de Zagreb (Zagreb), indice de Hosoya (Z), indice de connectivité moléculaire de Kier et Hall (X) |
| Descripteurs spatiaux         | Rayon de giration, volume moléculaire, surface, densité, moment d'inertie principal   |
| Descripteurs électroniques    | Charge, Apol, HOMO, LUMO, moment dipolaire  |
| Descripteurs thermodynamiques | AlogP98, réfractivité moléculaire   |

▪ **Méthode statistique : partitionnement récursif (RP)**

L'analyse RP a été effectuée en utilisant la méthode CSAR RP et l'arbre de décision a été généré à partir des résultats de RP.

La méthode CARTTM (arbres de classification et de régression) a été appliquée pour générer le modèle de classification de l'arbre de décision.

**c. Validation du modèle QSAR**

Le modèle d'arbre de décision a été validé en interne par validation croisée avec le nombre de groupes de validation croisée défini sur 10.

**d. Résultats et discussions**

▪ **Résultats**

- RP catégorise les molécules en résultant un arbre de décision binaire dans lequel les descripteurs sont utilisés pour diviser l'ensemble de données en sous-ensembles plus petits et homogènes. Un arbre de décision RP à 4 feuilles a été obtenu avec 4 nœuds terminaux et 3 nœuds non terminaux pour les analogues de chrominopyrimidine substitués basés sur des descripteurs 2D.
- L'ensemble de données a été catégorisé en deux ensembles :(classification selon les résultats des tests MABA et LORA) :

- Les inactifs (0) : l'ensemble Inactifs avait 11 molécules avec une inhibition <50%
- Les actifs (1) : l'ensemble des Actifs avait cinq molécules avec une inhibition > 50%
- Les résultats statistiques du modèle RP et les prédictions de l'activité antituberculeuse à partir de celui-ci sont résumés dans le tableau XV qui représente activité anti tuberculeuse par le modèle RP ou 0 représente la molécule inactive tandis que 1 représente la molécule active.

**TABLEAU XV : Activité antituberculeuse prédite par le modèle RP [78]**

| Composé        | Activité | Activité RP prévu | Descripteurs moléculaires |                  |         |
|----------------|----------|-------------------|---------------------------|------------------|---------|
|                |          |                   | Apol                      | Moment dipolaire | Densité |
| 4a             | 0        | 0                 | 13509.62                  | 3.10             | 1.20    |
| 4b             | 0        | 0                 | 14415.84                  | 4.14             | 1.20    |
| 4c             | 0        | 0                 | 45322.06                  | 2.25             | 1.20    |
| 4d             | 0        | 0                 | 14977.26                  | 4.76             | 1.18    |
| 4 <sup>e</sup> | 0        | 0                 | 15883.48                  | 2.98             | 1.18    |
| 4f             | 0        | 0                 | 14808.80                  | 4.01             | 1.222   |
| 4g             | 1        | 1                 | 13487.40                  | 8.70             | 1.41    |
| 4h             | 1        | 1                 | 14048.82                  | 10.05            | 1.38    |
| 5a             | 0        | 0                 | 12191.82                  | 3.10             | 1.18    |
| 5b             | 0        | 0                 | 13098.04                  | 4.12             | 1.18    |
| 5c             | 0        | 0                 | 14004.26                  | 2.29             | 1.18    |
| 5d             | 1        | 1                 | 13659.46                  | 4.77             | 1.16    |
| 5 <sup>e</sup> | 1        | 1                 | 14565.68                  | 2.86             | 1.16    |
| 5f             | 0        | 0                 | 13491.00                  | 4.02             | 1.20    |
| 5g             | 1        | 1                 | 12584.78                  | 5.33             | 1.21    |
| 5h             | 0        | 0                 | 12169.60                  | 8.64             | 1.40    |

- Afin d'éviter un sur-ajustement et d'améliorer la généralisation du modèle là, validation interne à l'aide de la technique de validation croisée décuplée exclut 10% les molécules réalisées. Un pourcentage de classification acceptable observé pour les ensembles suggère la nature stable de ce modèle pour les tests externes. Compte tenu de la taille et des autres limites de l'ensemble de données, ce taux de prédiction est acceptable et les descripteurs clés déterminés dans l'ensemble d'essais sont cohérents avec ceux observés dans l'ensemble.

▪ **Discussion**

- Sur la base des structures, les molécules sont divisées en deux classes : la phényl-2-thioxo-3,4-dihydro-1H-chromino [4,3-d] pyrimidin-5-one (classe A, 4a-h) et le phényle3,4-dihydro-1H-chromino [4,3-d] pyrimidine-2,5-dione (Classe B, 5a-h) avec huit molécules par classe.
- La majorité des molécules de classe A sont non cytotoxiques avec une activité antituberculeuse relativement bonne contre les cellules H37Rv.
- Les molécules appartenant à la classe B (5a-h) inhibent la croissance de M. tuberculosis >50% ; Cependant, ils ont des profils de cytotoxicité élevés.
- La présence de groupes donneurs d'électrons tels que méthyle, méthoxy ou hydroxyle sur le cycle 4-phényle du noyau chromino [4,3-d] pyrimidine-2,5-dione confère une plus grande activité antituberculeuse tandis que le brome entraîne une perte de l'activité.
- Une perte de l'activité et simultanément une augmentation de la cytotoxicité sont notées lorsqu'aucun substituant n'est présent sur le cycle 4-phényle.
- Les molécules qui ont une substitution méthyle à la neuvième position du noyau chromino [4,3-d] pyrimidine-2,5-dione possèdent une plus grande activité antituberculeuse et sont dépourvues de cytotoxicité comparée aux analogues non substitués.

**e. Conclusion**

La tuberculose (TB) est une maladie infectieuse potentiellement grave. De nombreuses souches de tuberculose résistent aux médicaments les plus utilisés. Les personnes atteintes de tuberculose active doivent prendre plusieurs types de médicaments pendant plusieurs mois pour éradiquer l'infection et prévenir le développement d'une résistance aux antibiotiques.

Des cibles supplémentaires doivent être recherchées des nouvelles pistes avec de nouveaux modes d'action sont nécessaires d'urgence.

La fusion du noyau chromino avec le cadre pyrimidique a révélé un nouveau noyau et des molécules prometteuses. Dans la lumière des résultats obtenus à partir de l'étude QSAR les molécules 5d, 5 e, 5g méritent d'être considérées pour un développement ultérieur. [78]

### 1.3.2. Substitution en C3 et C6 et C 7

#### 1.3.2.1. Evaluation de l'activité Inhibitrice du métabolisme oxydatif des neutrophiles à médiation immunitaire : étude QSAR 3D et 2D basées sur GRIND pour les dérivés de 3-phénylcoumarine.

##### a. Données expérimentales

Les coumarines naturelles de faible poids moléculaire et leurs dérivés synthétiques sont une nouvelle classe de médicaments anti-inflammatoires potentiels étudiés avec un grand intérêt.

Dans ce travail **L.M. Kabeya et al [79]**, ont étudié les effets modulateurs d'une série de 20 dérivés de 3-phénylcoumarine hydroxylés et acétoxylés sur le métabolisme oxydatif des neutrophiles à médiation immunitaire (les complexes immuns (CI) sont des processus essentiels dans les mécanismes de défense immunitaire humorale. Le dépôt de CI dans les tissus déclenche une réaction inflammatoire appelée hypersensibilité de type III caractérisée par une importante infiltration des neutrophiles et une libération des grandes quantités des ROS), ainsi que leur effet toxique sur ces cellules.

Des tests de chimioluminescence améliorés dépendants du luminol et de la lucigénine (CL-lum et CL-luc, respectivement) sont utilisés pour mesurer les ROS présents donc ils évalueront l'activité des composés synthétisés (le quercétin est utilisé comme molécule de référence).

Ce travail est complété par des études QSAR 2D et 3D Afin d'analyser une relation possible entre les activités des composés étudiés ici et leurs structures et discuter leurs aspects quantitatifs.

Le figure 21 représente Structures chimiques de la quercétine et dérivés de la 3 – phénylcoumarine.

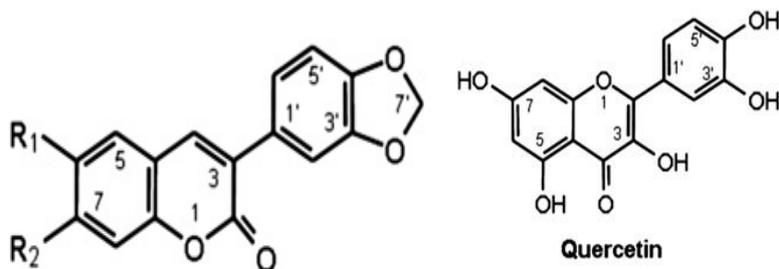


FIGURE 21 : Structures chimiques de la quercétine et dérivés de la 3 –phénylcoumarine [79]

**b. Descripteurs moléculaires et méthode statistique**

▪ **Descripteurs moléculaires**

- QSAR 3D : l'activité biologique des composés a été indiquée par les valeurs de pIC<sub>50</sub>. En utilisant cette approche, le modèle peut être obtenu sans avoir besoin d'un alignement 3D des structures.

Des descripteurs indépendants de la grille (GRIND) ont été générés, analysés et interprétés en utilisant ALMOND 3.3. Pour un ensemble de données contenant 13 composés, 570 descripteurs ont été obtenus.

- QSAR 2D : le modèle 2D étudié a donné un ensemble de 6 équations différentes présentant les différentes substitutions R et les différentes propriétés (hydrophobicité, lipophilie).

$$pIC_{50} \text{ CL-luc} = 4.21 + 2.92\sigma_1R_6 + \sigma_1R_7 + \sigma_1R_2 + \sigma_1R_3 + \sigma_1R_3 + \sigma_1R_4$$

▪ **Méthode statistique : partiel des moindres carrés (PLS)**

L'analyse statistique des relations entre CL-luc et les propriétés des substituants a été réalisée par l'algorithme des partiels des moindres carrés en utilisant le module QSAR du package SYBYL 7.3. Les résultats sont présentés dans le diagramme de scores PLS et le graphe des données Y expérimental vs calculé.

**c. Validation du modèle QSAR**

La capacité prédictive des modèles a été évaluée sur les 17 composés de l'ensemble d'apprentissage en utilisant l'approche de validation croisée et mesurée en termes des valeurs :

$$r^2\text{-SDEC} - q^2_{LOO} - \text{SDEP}.$$

**d. Résultats et discussions**

▪ **Résultats**

Dans le modèle final 3D (GRIND), un total de 213 descripteurs a été obtenu après sélection de la variable via FFD (Fractal factorial design).

Les résultats de calcul sont présents dans le tableau XVII qui représente la structure, valeurs des paramètres et chimioluminescence :

**TABLEAU XVI : Structure, valeurs des paramètres et chimioluminescence pour les composés [79]**

| Comp | $\sigma_1R_6$     | $\sigma_1R_7$     | $\sigma_1R_2$     | $\sigma_1R_3$     | $\pi R_3$          | $\sigma_1R_4$     | $\rho IC_{50}$<br>CL-luc | $\rho IC_{50}$<br>CL-<br>LUM |
|------|-------------------|-------------------|-------------------|-------------------|--------------------|-------------------|--------------------------|------------------------------|
| 20   | 0.33 <sup>d</sup> | 0.33 <sup>d</sup> | 0.00              | 0.33 <sup>d</sup> | -0.64 <sup>e</sup> | 0.33 <sup>d</sup> | 6.44                     | 6.41                         |
| 6    | 0.33 <sup>d</sup> | 0.33 <sup>d</sup> | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | 6.16                     | 6.18                         |
| 5    | 0.29 <sup>d</sup> | 0.29 <sup>d</sup> | 0.00              | 0.27 <sup>f</sup> | -0.64 <sup>e</sup> | 0.27 <sup>f</sup> | 6.14                     | 6.14                         |
| 18   | 0.33 <sup>d</sup> | 0.00              | 0.00              | 0.33 <sup>d</sup> | -0.64 <sup>e</sup> | 0.33 <sup>d</sup> | 5.75                     | 5.69                         |
| 15   | 0.29 <sup>d</sup> | 0.00              | 0.29 <sup>d</sup> | 0.00              | 0.00               | 0.00              | 5.60                     | 5.55                         |
| 16   | 0.33 <sup>d</sup> | 0.00              | 0.33 <sup>d</sup> | 0.00              | 0.00               | 0.00              | 5.59                     | 5.54                         |
| 14   | 0.33 <sup>d</sup> | 0.00              | 0.00              | 0.00              | 0.00               | 0.33 <sup>d</sup> | 5.40                     | 4.56                         |
| 13   | 0.29 <sup>d</sup> | 0.00              | 0.00              | 0.00              | 0.00               | 0.29 <sup>d</sup> | 5.36                     | 4.53                         |
| 17   | 0.29 <sup>d</sup> | 0.00              | 0.00              | 0.29 <sup>d</sup> | -0.67 <sup>e</sup> | 0.29 <sup>d</sup> | 5.35                     | 5.1                          |
| 11   | 0.29 <sup>d</sup> | 0.00              | 0.00              | 0.29 <sup>d</sup> | -0.67 <sup>e</sup> | 0.00              | 5.28                     | -g                           |
| 19   | 0.29 <sup>d</sup> | 0.29 <sup>d</sup> | 0.00              | 0.29 <sup>d</sup> | -0.67 <sup>e</sup> | 0.29 <sup>d</sup> | 5.28                     | 4.87                         |
| 2    | 0.33 <sup>d</sup> | 0.00              | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | 5.27                     | -g                           |
| 1    | 0.29 <sup>d</sup> | 0.00              | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | 5.25                     | -g                           |
| 12   | 0.33 <sup>d</sup> | 0.00              | 0.00              | 0.33 <sup>d</sup> | -0.64 <sup>e</sup> | 0.00              | 5.16                     | 4.66                         |
| 4    | 0.00              | 0.33 <sup>d</sup> | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | 5.14                     | -g                           |
| 10   | 0.33 <sup>d</sup> | 0.00              | 0.00              | 0.00              | 0.00               | 0.00              | 5.13                     | 4.51                         |
| 9    | 0.29 <sup>d</sup> | 0.00              | 0.00              | 0.00              | 0.00               | 0.00              | 5.10                     | -g                           |
| 3    | 0.00              | 0.29 <sup>d</sup> | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | 5.07                     | -g                           |
| 7    | 0.00              | 0.00              | 0.00              | 0.27 <sup>f</sup> | -0.02 <sup>e</sup> | 0.27 <sup>f</sup> | -g                       | -g                           |
| 8    | 0.00              | 0.00              | 0.00              | 0.00              | 0.00               | 0.00              | -g                       | -g                           |

L'étude 3D (GRIND) est significativement validée en utilisant la validation croisée, Il a donné une quantité importante de variance (75.7%) et les valeurs des paramètres suivants :

$$r^2 = 0.95; SDEC = 0.14; q^2_{LOO} = 0.76 \text{ par validation interne.}$$

$$SDEP = 0,32.$$

L'étude 2D : L'approche de validation croisée et mesurée en termes de valeurs de  $q^2$ . Les équations obtenues ont donné une bonne valeur  $q^2$  de 0,958. Un modèle est considéré comme significatif lorsque  $q^2 > 0,3$ .

▪ **Discussion**

- Les descripteurs les plus importants dans notre modèle PLS peuvent être arrangés pour obtenir un pharmacophore approximatif valable pour les composés étudiés ici. Le pharmacophore à base de MIF résultant a suggéré l'importance des groupes accepteurs de liaisons hydrophobes et de liaisons hydrogènes inversement aux positions C-7 et C-3, et a également indiqué la taille idéale du ligand.
- Les résultats 2D-QSAR ont fourni une meilleure corrélation pour le CL-luc que pour le CL-lum.
- L'influence de R6, R7, R2' et R4' était inductive, tandis que celle de R3' était hydrophobe à travers la constante de substituant au carré dans le CL-luc. Bien que les effets des groupes R7, R2', R3' et R4' pour le CL-lum soient similaires à ceux de CL-luc, la contribution hydrophobe pour le premier était beaucoup plus importante, tandis que le substituant R6 n'a pas contribué, soit électroniquement ou hydrophobiquement, au CL-lum.

**e. Conclusion**

Bien que l'ensemble de données ne soit pas très grand, les auteurs ont construit un modèle 3D-QSAR avec une signification statistique élevée, ce qui a permis de proposer un site récepteur virtuel considérant les aspects qualitatifs (régions pharmacophoriques) et quantitatifs (distances mutuelles).

L'étude 2D a montré une forte capacité à expliquer les relations entre les caractéristiques structurelles des dérivés de 3-phénylcoumarine et leur activité biologique.

Cette étude pourrait contribuer au développement de nouvelles molécules pour contrôler l'inflammation et les lésions tissulaires induites par les radicaux libres.

**1.4. Coumarine simple poly substituée**

Cinq exemples ont été traités :

- la substitution en C2, C3, C4, C5, C6, C7 et C8,
- la substitution en C2, C3, C4, C6,
- la substitution en C3, C6, C7, C8,
- la substitution C4, C5, C6, C7 et C8
- et la substitution en C5, C6, C7, C8 et C9.

### 1.4.1. Substitution en C2, C3, C4, C5, C6, C7 et C8

#### 1.4.1.1. Synthèse, relations structure-activité (SAR) et études silico des dérivés de la coumarine à activité antifongique.

##### a. Données expérimentales

L'augmentation de l'incidence des infections fongiques opportunistes, associée à une plus grande résistance aux médicaments antifongiques actuellement utilisés a obligé les recherches vers de nouveaux achèvements thérapeutiques.

Dans cette étude de **Rodrigo S. A. de Araújo et al**, vingt-quatre dérivés de la coumarine ont été examinés in vitro pour l'activité antifongique contre les souches d'*Aspergillus*, Certains des composés présentaient une activité antifongique significative.

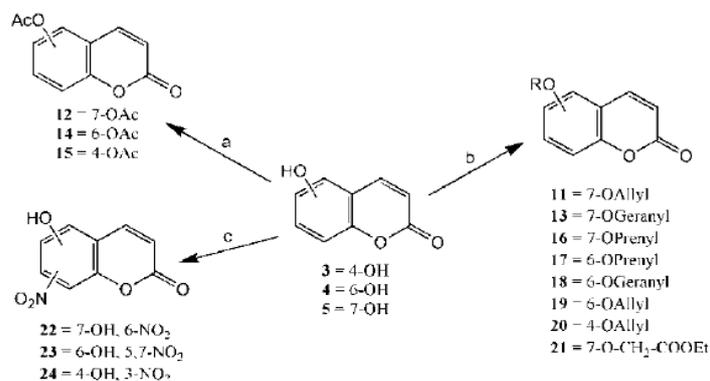


FIGURE 22 : La structure chimique générale des dérivés d'alkyl-, acétyl- et nitro-coumarine [80]

L'activité antifongique in vitro des composés synthétisés a été étudiée en utilisant huit souches d'*Aspergillus*, Les valeurs de CMI ont été déterminées par la méthode du bouillon de microdilution en utilisant des microplaques à 96 puits et l'amphotéricine B a été utilisée comme fongicide de référence.

Après une évaluation antifongique, les dérivés ont été soumis à une analyse de relation structure-activité, à des analyses de surface électroniques, à la modélisation moléculaire et à la chimiométrie (Analyse en composantes principales (ACP) et Régression partielle des moindres carrés (PLS)).

##### b. Descripteurs moléculaires et méthode statistique

###### ▪ Descripteurs moléculaires

**Rodrigo S. A. de Araújo et al** [80], utilisent le logiciel Pentacle Version 1.5 est un outil de calcul pour déterminer les descripteurs moléculaires sans alignement, ils sont également appelés descripteurs GRID-IN ou GRIND, Les descripteurs GRIND sont

très pertinents pour décrire les propriétés biologiques des composés étudiés. Plus de 405 variables ou descripteurs indépendants sont étudiés dans cette recherche.

▪ **Méthodes statistiques : Analyse en composantes principales (ACP) et régression des moindres carrés partiels (PLS)**

L'analyse en composantes principales (ACP) et les moindres carrés partiels (PLS) sont des outils chimiométriques permettant d'extraire et de rationaliser l'information de toute description multivariée d'un système biologique.

La réduction de la complexité et la simplification des données sont les deux caractéristiques les plus importantes de ces outils.

Ces outils chimiométriques ont été développés dans le logiciel Pentacle, Les valeurs de CMI ont été converties en logarithmes négatifs et utilisées comme variables dépendantes dans cette étude.

$$pMIC = -\log MIC$$

**c. Validation du modèle QSAR**

Le pouvoir prédictif et la validation du modèle étudié est déterminé par la méthode leave-one-out en calculent les paramètres suivants :  $r^2$  et  $q^2$ .

**d. Résultats et discussions**

▪ **Résultats**

- Les données présentées dans l'annexe 05 montrent que les coumarines commerciales (1-10) sont généralement inefficaces contre les souches testées, Certains dérivés ont montré des activités antifongiques significatives avec des CMI comprises entre 16 et 32  $\mu\text{g} / \text{ml}$ , incluant sept composés (2, 11, 12, 14, 15, 20 et 22) plus actifs que le médicament de référence Amphotéricine B pour le LM- 26 souche (*A. flavus*).
- Le meilleur modèle de PLS a été obtenu avec 72 descripteurs sélectionnés parmi 405, Les principaux descripteurs mis en évidence (générés par des interactions avec la sonde TIP) dans le modèle PLS étaient :  
350 (5.6-6 A), 531 (10.8-11.2 A), 527 (7.2-7.6 A), et 201 (13.2-13.6 A)
- Les descripteurs moléculaires obtenus peuvent être observés dans des diagrammes graphiques appelés "corrélogrammes" nous avons observé le plus grand nombre d'interactions des composés actifs avec la sonde N1. Cela peut indiquer que les

coumarines dépourvues d'électrons ont une plus grande attraction avec l'atome électronégatif de cette sonde.

- Le modèle obtenu a donné un pouvoir prédictif significatif ou les valeurs des paramètres étudiés sont :

$$r^2 = 0.86 \quad \text{et} \quad q^2 = 0.64$$

- **Discussion**

- L'étude SAR a permis deux conclusions : la substitution O est essentielle pour l'activité antifongique et la présence d'une courte chaîne aliphatique et / ou de groupes électroattracteurs (NO<sub>2</sub> et / ou acétate) favorise l'activité.
- Les résultats de la chimiométrie (ACP et PLS) ont été satisfaisants et en corrélation avec l'analyse SAR. Les composés les plus actifs ont montré une plus grande interaction avec la sonde N1, qui montre une plus grande concentration électronique.
- Des descripteurs générés par la sonde TIP ont également été mis en évidence, indiquant que le contour moléculaire est, en effet, important pour l'activité antifongique, comme observé dans l'analyse SAR.

- e. **Conclusion**

Il existe un besoin urgent de découvrir et de tester de nouveaux composés ayant des propriétés antifongiques. Les produits naturels sont inépuisables en tant que source de structures chimiques et ont été utilisés pendant plus d'un siècle comme plateformes pour la synthèse des nouveaux médicaments. Les résultats d'étude QSAR obtenue suggèrent des dérivés de la coumarine comme composés prometteurs pour le développement de nouveaux agents anti-Aspergillus et ouvrent la porte pour découvrir d'autres molécules dans le futur [80].

### 1.4.2. Substitution en C2, C3, C4 et C6

#### 1.4.2.1. Effets cytotoxiques, propriétés alkylantes et modélisation moléculaire des dérivés de la coumarine et de leur analogue phosphonique.

##### a. Données expérimentales

Elzbieta Budzisz et al [81], ont étudiés l'activité cytotoxique d'une série des dérivés de benzopyrone 4, 5 et 6 et leurs analogues phosphoniques qui sont déjà synthétisé avant sur 2 lignées cellulaire, de leucémie humaine.

La cytotoxicité des dérivés de benzopyrone 4 $\beta$  / 6 a été déterminée sur les deux lignées cellulaires de leucémie humaine, HL-60 promyélocytaire et NALM-6 lymphoblastique. Les valeurs d'IC50 (la concentration des composés test requis pour réduire la fraction de survie cellulaire à 50% du témoin) ont été calculées à partir des courbes dose-réponse et utilisées comme mesure de la sensibilité cellulaire à un traitement donné. (Warfarine a été utilisée comme référence).

Les composés 4a / c, 5a $\beta$  / c et 6a / c ont été criblés pour leur activité d'alkylation vis-à-vis de la 4- (4 $\beta$ -nitrobenzyl) pyridine (NBP) (Le taux d'alkylation a été quantifié par spectrophotométrie à 560 nm)

Une étude QSAR a été envisagée pour déterminer la relation entre les propriétés alkylantes des composés et leurs cytotoxicités vis-à-vis les lignes étudiées, ainsi expliquer la relation structure activité.

La figure 23 représente des dérivés des coumarines 4a–c et leurs analogues phosphoniques 5a–c

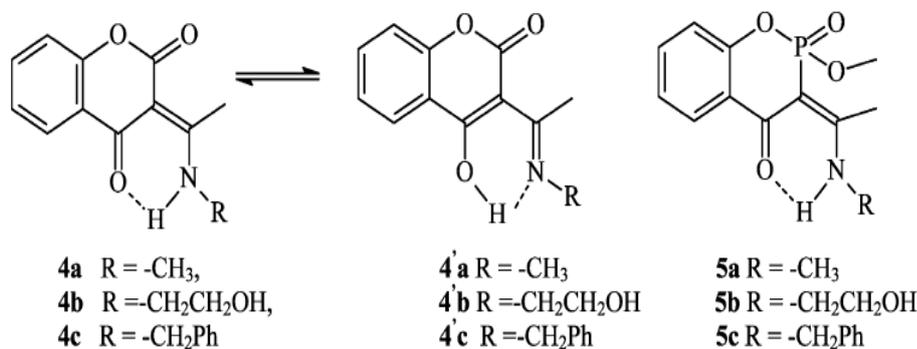


FIGURE 23 : Structure générale des dérivés des coumarines 4a–c et leurs analogues phosphoniques 5a–c [81]

### b. Descripteurs moléculaires et méthode statistique

#### ▪ Descripteurs moléculaires

Les descripteurs physico-chimiques suivants sont utilisés pour les analyses mathématiques dans cette étude QSAR :

Volume de van der Waals (Vol), réfraction molaire (MR), charge électrique focalisée sur un atome électronégatif du substituant en positions 2, 3 et 4 (Q), moment dipolaire (DM), la valeur de lipophilie (log P) et le poids moléculaire-MW (pMW).

Ces paramètres sont calculés pour les molécules entières et pour leurs substituants particuliers aux positions 2, 3 et 4.

#### ▪ Méthode statistique : méthodes de régression linéaire et multiple

L'analyse de régression a été réalisée à l'aide d'un programme STATISTICA 5.1. L'utilisation de plus d'une variable dans une équation multivariée est justifiée par l'étude d'inter-corrélation envisagés.

Dans l'analyse de régression multiple, nous avons utilisé exclusivement des paramètres ne montrant pas d'inter-corrélation mutuelle. Les équations obtenues sont :

➤ Pour l'activité cytotoxique :

$$pIC50 = 1.546(\pm 0.709) + 0.007(\pm 0.002) V_2 - 1.140(\pm 0.338) pMW_3 \quad (1)$$

$$pIC50 = -6.50(\pm 1.10) - 2.00(\pm 0.48) pMW_{2-4} \quad (2)$$

➤ Pour la propriété alkylante:

$$\log A = 1.43(\pm 0.15) - 1.42(\pm 0.23) Q_3 \quad (3)$$

$$\log A = 1.94(\pm 0.11) - 0.02(\pm 0.000) V_2 - 1.72(\pm 1.13) Q_3(4)$$

$$\log A = 1.180(\pm 0.440) - 0.002(\pm 0.000) V_3 - 1.357(\pm 0.263) Q_3 \quad (5)$$

A : Absorbance calculée par spectrophotométrie

### c. Validation du modèle QSAR

La validation des modèles obtenues, le pouvoir prédictif, la corrélation des descripteurs avec l'activité et l'absence d'inter-corrélation mutuelle entre eux sont déterminés par le calcul des paramètres suivants :

$$F, r^2, r, n, SEE, P$$

### d. Résultats et discussions

#### ▪ Résultats

- D'abord, le premier but était de chercher et expliquer l'existence d'une relation entre la propriété alkylante et l'activité cytotoxique des composés synthétisés, les

données issues des études pharmacologiques (voir le tableau XVIII qui montre l'activité cytotoxique des dérivés vis-à-vis les 2 lignées cellulaires) et les coefficients de corrélation calculés ( $r$ ) étaient respectivement de 0,42 et 0,19.

- Les valeurs de cytotoxicité montrent que les composés testés sont plus actifs sur la ligne HL-60 que la ligne NALM-6 et on peut diviser l'ensemble des composés en : activité cytotoxique modérée (4a, 6b) –groupe à faible cytotoxicité –groupe très toxique (5c, 6c les plus toxiques) les différents résultats sont montrés dans le tableau XVIII.
- La validation du modèle s'effectue par le calcul des paramètres. Pour chaque équation obtenue, les résultats sont :
  - Equation 1:  $r=0.84, r^2=0.71 ; n=9, F=7.282, SEE=0.125, p<0.02$
  - Equation 2:  $r=0.90; r^2=0.81; n=6; F=17.087; SEE=0.117; P<0.01$
  - Equation 3:  $r=0.91 \quad r^2=0.83 ; n=10 \quad F=38.198, SEE=0.281 \quad p<0.0002$
  - Equation 4:  $r=0.99, r^2=0.97; n=10, F=114.940, SEE=0.124, p<0.0000$
  - Equation 5:  $r=0.91, r^2=0.84 ; n=10, F=17.798, SEE=0.293 \quad p<0.001$

**Tableau XVII : Activité cytotoxique des dérivés de coumarines vis à vis les deux lignées cellulaires [81]**

| Comp      | HL-60          |            |             |              |                         | NALM-6         |            |             |              |                         |
|-----------|----------------|------------|-------------|--------------|-------------------------|----------------|------------|-------------|--------------|-------------------------|
|           | Viabilité<br>% | IC50       |             |              | PIC50                   | Viabilité<br>% | IC50       |             |              | PIC50                   |
|           |                | 10<br>(mM) | 100<br>(mM) | 1000<br>(mM) | $\times 10^{-6}$<br>(M) |                | 10<br>(mM) | 100<br>(mM) | 1000<br>(mM) | $\times 10^{-6}$<br>(M) |
| 4a        | 87             | 67         | 17          | 408.0        | 3.3893                  | 86             | 62         | 0           | 276.4        | 3.5585                  |
| 4b        | 91             | 88         | 19          | 590.8        | 3.2286                  | 94             | 44         | 0           | 89.4         | 4.0487                  |
| 4c        | 106            | 103        | 35          | 801.0        | 3.0964                  | 93             | 40         | 0           | 83.0         | 4.0809                  |
| 5a        | 101            | 96         | 20          | 649.3        | 3.1876                  | 93             | 43         | 0           | 87.9         | 4.0560                  |
| 5a        | 101            | 89         | 36          | 772.4        | 3.1122                  | 107            | 77         | 1           | 420.0        | 3.3768                  |
| 5c        | 86             | 29         | 0           | 66.8         | 4.1752                  | 91             | 8          | 0           | 54.4         | 4.2644                  |
| 6a        | 102            | 92         | 7           | 542.3        | 3.2658                  | 89             | 40         | 0           | 81.9         | 4.0867                  |
| 6b        | 87             | 78         | 1           | 428.2        | 3.3684                  | 84             | 52         | 0           | 133.         | 3.8735                  |
| 6c        | 83             | 45         | 6           | 88.0         | 4.0555                  | 99             | 37         | 3           | 80.9         | 4.0921                  |
| Warfarine | 84             | 83         | 5           | 486.0        | 3.3134                  | 95             | 41         | 1           | 74.8         | 4.1261                  |

▪ **Discussion**

- L'augmentation de la valeur de Vol2 (volume de van der Waals) favorise une activité biologique plus élevée des dérivés de coumarine testés (équation 1).
- La cytotoxicité s'accompagne d'une augmentation de la valeur des paramètres hydrophobes et elle est corrélée avec le poids moléculaire (MW) des substituants aux positions : 2-, 3- et 4 (équation 2).
- L'effet d'alkylation est augmenté avec la charge de l'atome électronégatif (N ou O) lié au substituant en position 3, et le paramètre volumique de van der Waals, et le paramètre stérique de la région 2 (équation 4 et 5).
- Pas de corrélation entre la propriété alkylante des dérivés et leur activité cytotoxique. Ce résultat peut indiquer que l'effet toxique est causé par un mécanisme indépendant de l'alkylation.

**e. Conclusion**

Le travail d'**Elzbieta Budzisz et al [81]** a donné des informations pharmacologiques et structurales très importantes qui peuvent être la base d'autres recherches plus approfondies sur les dérivés des coumarines et leurs activités cytotoxiques.

**1.4.3. Substitution en C3, C6, C7 et C8**

**1.4.3.1. Relations structure-activité et études d'amarrage de dérivés de la coumarine en tant que nouveaux ligands sélectifs pour le récepteur CB2**

**a. Données expérimentales**

Les récepteurs aux cannabinoïdes, y compris les sous-types CB1 et CB2 sont des membres de la superfamille des récepteurs couplés aux protéines G. Le récepteur CB2 a été considéré comme une cible thérapeutique inspirante pour le traitement de la douleur et des maladies immunitaires.

Des analogues de 2-pyridone (S-444823 et S-777469) et des dérivés de 2-quinolone (JTE-907 [études Iwamura et al]), qui ont été déjà connues comme ligands sélectifs des récepteurs CB2, ont été sélectionnées pour construire un modèle CoMFA, afin d'élucider les exigences structurales responsables de la liaison des ligands au récepteur CB2. Ces ligands sont divisés en ensemble d'entraînement contenant 45 composés et un ensemble d'essai avec 10 composés.

Guidé par les caractéristiques pharmacophoriques dérivées des composés générés par CoMFA et les caractéristiques structurales des ligands précités, série des dérivés de 8-alcoucoumarine-3-carboxamide N Substitués (Fig.24 représente la structure générale des molécules cibles) a été conçue sous la direction du modèle de calcul, les études de relation structure-activité (SAR), et le principe combiné de bioisostère et d'hybridation par **Shuang Hanune et al** [82]

Les dosages de mobilisation de calcium sur cellules entières sont utilisés pour évaluer les activités fonctionnelles in vitro des composés synthétisés et les composés de référence GW842166X et JTE-907.

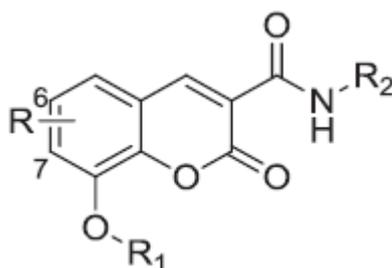


FIGURE 24 : La structure générale des molécules cibles [82]

#### b. Etude QSAR : Analyses CoMFA

Afin de générer le modèle CoMFA, la base des données (les ligands) est divisée en ensemble d'entraînement contenant 45 composés et un ensemble d'essai avec 10 composés.

Des ligands pour les analyses CoMFA sont créés en modifiant la structure de JTE-907 en utilisant le module Sybyl / Sketch, puis alignés dans un espace 3D basé sur la conformation de JTE-907 comme modèle, puis les composés des ensembles d'entraînement et de test sont ensuite superposés sur la conformation de liaison potentielle prévue de JTE-907 pour obtenir leur alignement structural.

Le module Sybyl x1.3 / CoMFA est ensuite utilisé pour générer le modèle 3D-QSAR.

Les valeurs de  $K_i$  des composés dans les ensembles d'entraînement et de test sont converties en valeurs de  $pK_i$  ( $-\log K_i$ ), qui ont été utilisées comme variables dépendantes dans les études de CoMFA.

Les analyses initiales des moindres carrés partiels (PLS) sont effectuées pour générer le modèle COMFA.

**c. Validation du modèle QSAR**

La validation statistique est une étape préliminaire pour déterminer et confirmer la capacité et fiabilité de modèle pour prédire et expliquer les relations structure activité pour les composés sélectionnés en tant que ligands CB2.

Une validation croisée leave-one-out (LOO) et des paramètres statistiques sont calculés :

$$\text{LOO } r_{cv}^2 - \text{ONC} - R^2 - \text{SEE} - F - \text{Steric contribution} - \text{Electrostatic contribution}$$

**d. Résultats et discussions**

▪ **Résultats**

- Le modèle généré remplit le critère de validité statistique CoMFA, les résultats statistiques sont montrés dans le tableau XIX qui représente les résultats statistiques de modèle COMFA :

**TABLEAU XVIII : Résultats statistiques du modèle CoMFA [82]**

| <b>Modèle CoMFA</b>        |         |
|----------------------------|---------|
| LOO $r_{cv}^2$             | 0.610   |
| ONC                        | 5       |
| R <sup>2</sup>             | 0.970   |
| SEE                        | 0.250   |
| F                          | 256.352 |
| Steric contribution        | 0.468   |
| Electrostatic contribution | 0.532   |

- Les valeurs prédites se rapprochent des valeurs pKi expérimentales avec des résidus inférieurs à l'unité logarithmique, indiquant que le modèle a une bonne capacité prédictive. Les résultats sont montrés dans l'annexe 06.
- Les valeurs EC50 ou IC50 ont été déterminées pour les composés actifs identifiés au stade de criblage. En conséquence, de nombreux composés présentent une sélectivité élevée vis-à-vis du récepteur CB2 contre le récepteur CB1.
- Parmi les ligands bioactifs, le meilleur agoniste 30 et l'antagoniste présente des bioactivités 4 et 34 fois plus élevées que les composés de référence correspondants GW842166X ou JTE-907, respectivement.

▪ **Discussion**

- La longueur de chaîne carbonique R1 influe directement sur l'activité agoniste du récepteur CB2 (le composé 15 plus actif que le composé 14,16 et 17).
- Les composés préparés 18 et 19 présentent des activités faiblement agonistes sur le récepteur CB2 en remplaçant la chaîne carbonée par un groupe cyclopropylméthyle ou benzyle.
- La molécule 15 est considérée comme molécule de référence, le R1 est fixé et le R2 va être modifié. Les différentes substitutions ont donné des molécules moins actives que le composé 15 sauf le composé 30 avec un substituant isopropyle qui présente la meilleure activité (4 fois plus que GW842166X).
- Le bloc de construction cyclohexylamine a été remplacé par de la pipéridine. Le nouveau composé 33 entraîne une diminution significative de l'activité du récepteur CB2, ce qui indique que le groupe NH de l'amide pourrait être crucial pour un agoniste puissant.
- Des analogues de 6-bromo et de 7-méthoxycoumarine ont été également synthétisés pour étudier l'influence de l'introduction d'un groupe C6 ou C7- sur la bioactivité du dérivé apparenté.
- Les résultats trouvés montrent de manière intéressante que ces composés présentent les activités antagonistes du récepteur CB2, on pourrait induire qu'un substituant en position C7 pouvait avoir une plus grande influence sur la puissance agoniste inverse d'un ligand CB2 que sur sa position C6.

**e. Conclusion**

Les études SAR ont indiqué que la chaîne droite R1 et le groupe aliphatique R2 et un substituant hydrophobe du groupe fonctionnel amide C3 étaient favorables à la bioactivité du récepteur CB2 d'un ligand et seraient responsables de l'activité in vitro d'un ligand.

De plus, les composés sans substitution en position C6 ou C7 sont les agonistes du récepteur CB2, tandis que d'autres composés portant un groupe 6-bromo ou 7-méthoxy démontrent des activités antagonistes au récepteur CB2.

Dans l'ensemble, les dérivés de la coumarine constitueraient des points de départ efficaces pour le développement de puissants ligands du récepteur CB2, et les résultats bioactifs et de calcul pourraient fournir une ligne directrice pour caractériser les activités fonctionnelles des nouveaux ligands du récepteur CB2. [82]

#### 1.4.4. Substitution en C4, C5, C6, C7 et C8

##### 1.4.4.1. Activité antituberculeuse d'acide coumarin-4-acétique benzylidène hydrazides : Etude 3D-QSAR

###### a. Données expérimentales

Selon l'OMS, de 2002 à 2020, il y aura environ 1 milliard de personnes nouvellement infectées par la tuberculose et environ 36 millions de décès si le ravage mondial de la tuberculose reste incontrôlé. En ce qui concerne le développement de nouveaux médicaments, aucun nouvel agent antituberculeux n'a été lancé depuis l'introduction de la rifampicine en 1965, malgré les progrès considérables qui ont été accomplis dans le découvert des médicaments.

Les dérivés naturels de la coumarine comme le Calanolide A et B présentent une puissante activité antituberculeuse avec une CMI de 3,13 mg / ml.

Dans ce travail, **Atul Manvar et al [84]**, créent un programme de fabrication d'agents antituberculeux en utilisant l'anneau coumarinique central, dans le Calanolide A (structure A du Schéma) comme matrice et ajoutant des substituants sur les différentes positions si nécessaires pour conférer de l'activité.

Une série de 25 composés dérivés de benzylidène hydrazides de l'acide coumarin-4-acétique (structure B) sont synthétisés et évalués pour leurs activités antituberculeuses contre *M. tuberculosis* H37Rv en utilisant la rifampicine comme référence et Les valeurs MIC ont été déterminées.

La recherche est poursuite par une étude QSAR-3D basée sur l'analyse du champ moléculaire comparatif (CoMFA) pour établir la relation entre la structure et l'activité de cette classe de composés.

La figure 25 présente structure de Calanolide A, structure hydrazides de l'acide coumarin-4-acétique structure B.

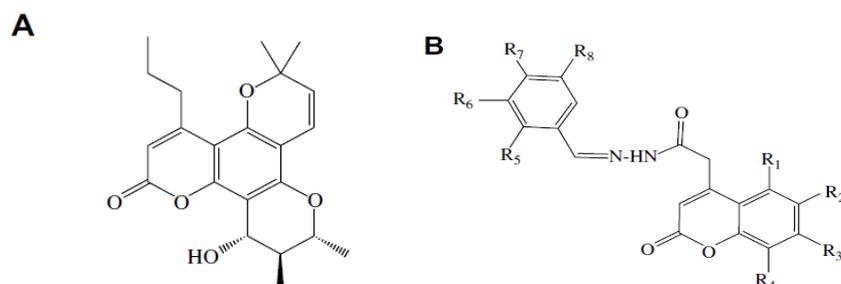


FIGURE 25 : A structure générale de Calanolide, B structure générale de benzylidène hydrazides de l'acide coumarin-4-acétique.

## b. Etude QSAR–3D

### ▪ Etude CoMFA

Les études CoMFA ont été réalisées avec Sybyl 7.1 où les structures des molécules ont été construites avec le module Sketcher.

L'ensemble de données a été divisé en un ensemble d'apprentissage (19 molécules) et un ensemble de test 6 molécules (5D, 5D, 5J, 5Q, 5U, 5W) au moyen de la diversité chimique ainsi que biologique.

Les champs stériques et électrostatiques, le potentiel de van der Waals et l'énergie coulombienne sont les descripteurs les plus importants déterminés dans COMFA.

La phase la plus cruciale pour CoMFA est l'alignement des molécules, Divers modèles CoMFA ont été générés pour deux alignements : la base de données et l'ajustement sur le terrain (la molécule 5y la plus active est choisie comme matrice).(84)

Le meilleur modèle CoMFA est représenté par Les courbes de contour stérique et électrostatique (Les contours colorés sur la carte représentent les zones dans l'espace 3D où les changements dans les valeurs de champ stérique et électrostatique d'un composé sont fortement corrélés avec le changement concomitant de son activité biologique).

### ▪ Méthode statistique : Analyse des moindres carrés partiels (PLS)

Méthode PLS a été utilisée pour établir une corrélation entre les champs moléculaires et l'activité inhibitrice des molécules.

Le nombre optimal de composants a été déterminé avec SAMPLS (Samples-distance Partial Least Square) et le calcul statistique est effectué sur chacun des échantillonnages.

Les valeurs d'activité biologique ont été transformées comme suit :(84)

$$\text{Activité} = -\log c + \text{logit}$$

Où c : est la concentration molaire = concentration (mg / ml)  $\times$  0,001 / poids moléculaire

$$\text{Logit} : \log [\% \text{inhibition} / (100 - \% \text{inhibition})]$$

## c. Validation du modèle QSAR

La capacité prédictive de chaque modèle 3D-QSAR a été déterminée à partir de l'ensemble de test (6 molécules non incluses dans la génération du modèle), une validation croisée a été par la méthode leave-one-out.

- On calcule le coefficient de corrélation prédictif

$$r^2_{\text{pred}} = (\text{SD} - \text{PRESS}) / \text{SD}$$

Où SD : la somme des écarts quadratiques entre l'activité biologique de l'ensemble de test et l'activité moyenne des molécules de l'ensemble d'entraînement

PRESS est la somme des écarts quadratiques entre les valeurs d'activité prévues et réelles pour chaque molécule dans l'ensemble de test

- Les autres paramètres :  $r^2$  et  $q^2$

#### d. Résultats et discussions

##### ▪ Résultats

- Les deux meilleurs modèles de COMFA sont rapportés et considérés pour une analyse plus approfondie, le modèle 1 obtenu par l'alignement de la base de données et le modèle 2 dérivé de l'alignement du champ.
- Les 2 modèles ont des paramètres qui sont statistiquement significatifs. Ces statistiques sont présentées dans le tableau XX qui montre les statistiques obtenues pour les 2 modèles.
- Le modèle 1 présente des meilleures qualités statistiques globales :  
Un coefficient de corrélation conventionnel ( $r^2$ ) de 0,983, un coefficient de corrélation croisée ( $q^2$ ) de 0,636 et un coefficient de corrélation ( $r^2_{\text{pred}}$ ) de 0,338.
- Le modèle 2 présente une valeur plus élevée de ( $r^2_{\text{pred}}$ ) 0,464, les valeurs de  $r^2$  et de  $rcv^2$  sont relativement plus faibles.
- L'analyse des contours stériques du modèle 1 révèle :
  - De larges contours colorés verts entourant presque toutes les positions du cycle phényle du fragment benzylidène hydrazide.
  - Aucun contour de couleur jaune n'est visible, ni aucun contour stérique significatif vu autour de l'anneau coumarine.
  - Des contours de couleur bleue près des positions méta et para du cycle phényle de la fraction benzylidène hydrazide et Aucun contour électrostatique n'est observé autour de l'anneau de coumarine.
- Les contours stériques et électrostatiques du modèle 2 conduisent également à la même interprétation.

TABLEAU XIX : Les statistiques obtenues pour les 2 modèles COMFA [84]

| Paramètre                      | Alignement de la base | Alignement sur le terrain |
|--------------------------------|-----------------------|---------------------------|
|                                | CoMFA model 1         | CoMFA model 2             |
| N                              | 6                     | 5                         |
| r <sup>2</sup>                 | 0.983                 | 0.951                     |
| r <sup>2</sup> <sub>cv</sub>   | 0.636                 | 0.403                     |
| LGO (5)                        | 0.522                 | 0.473                     |
| F-value                        | 116                   | 51                        |
| r <sup>2</sup> <sub>pred</sub> | 0.338                 | 0.464                     |
| r <sup>2</sup> <sub>bs</sub>   | 0.992                 | 0.991                     |
| SD                             | 0.052                 | 0.056                     |
| Contributions %                |                       |                           |
| Stérique                       | 59.5                  | 53.1                      |
| Electrostatique                | 40.5                  | 46.9                      |

▪ **Discussion**

- Les molécules avec des substituants de taille moyenne, par ex. groupe méthoxy, sur le noyau phényle, présente une bonne activité inhibitrice (5h et 5u).
- La molécule hautement active (5y) de la série possède un substituant OH électronégatif à la position C7 du cycle coumarine.
- Les molécules ayant un substituant méthyle en C6 (5n et 5x) présentent une activité inhibitrice modérée à bonne car le groupe méthyle est proche du contour électropositif favorable de couleur bleue.
- Les molécules 5j et 5n ont également un substituant méthyle en C6 mais ceux-ci ont une activité faible car le substituant méthyle en C7 supplémentaire est proche du contour coloré rouge défavorisé.

**e. Conclusion**

Parmi les différents modèles évalués, le modèle CoMFA dérivé de l'alignement de la base de données était le meilleur avec une bonne corrélation et un bon pouvoir prédictif. L'analyse des contours de CoMFA fournit des détails sur la relation entre la structure et l'activité, et fournit des indices pour les modifications structurelles qui peuvent améliorer l'activité.

Les résultats fournis montrent la capacité de dérivé de benzylidène hydrazides de l'acide coumarin-4-acétique à proposer des composés prometteurs qui pourraient être la base des recherches plus avancées. [84]

### 1.4.5. Substitution en C5, C6, C7, C8 et C9

#### 1.4.5.1. Étude QSAR sur les coumarines en tant qu'agents anti-méningo-encéphalitique

##### a. Données expérimentales

Dans le cadre de conception, de découverte et de développement des nouveaux agents antifongiques, antiméningoencéphalitique (La méningoencéphalite est une maladie causée par le champignon virulent *Cryptococcus neoformans*), **Tarun Jha et al [85]** ont réalisé une étude quantitative de la relation structure-activité (QSAR) sur une série des dérivés de la coumarine qui présentent une activité anti-méningoencéphalitique exprimée en valeurs MIC50 ( $\mu\text{g} / \text{mL}$ ).

Les valeurs MIC50 ont été recueillies à partir des travaux publiés de **Sourish et al**, ils ont converti en logarithmes négatifs de MIC50 (pMIC50) pour obtenir la relation linéaire avec les variables indépendantes.

Les études QSAR de 25 coumarines ont été réalisées en utilisant différents paramètres tels que des descripteurs chimiques quantique physicochimiques, constitutionnels et semi-empiriques ainsi que des descripteurs topologiques, moléculaires entiers.

La figure 26 représente structure générale avec numérotation arbitraire.

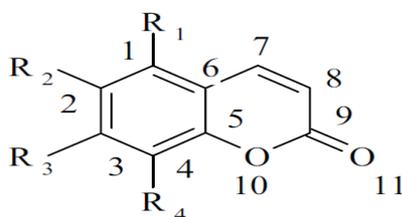


FIGURE 26 : Structure générale des coumarines avec numérotation arbitraire. [85]

##### b. Descripteurs moléculaires et méthode statistique

###### ▪ Descripteurs moléculaires

Les descripteurs constitutionnels et les descripteurs topologiques moléculaires entiers ont été calculés par le logiciel DRAGON version 2.1. Les descripteurs chimiques quantiques sont calculés en utilisant le paquet Hyperchem Pro Release 7.0 Pro.

- Les descripteurs constitutionnels : nO, nR06 et RBF
- Les descripteurs topologiques moléculaires entiers : IC2, T (N... O), TI2, SIC1, SCI2, Lop, PJI2 et D/Dr05.
- Les descripteurs chimiques quantiques : les charges atomiques et la surface moléculaire, Le champ de force de la mécanique moléculaire (MM +), FEH et LEL.

▪ **Méthode statistique : régression linéaire multiple (MLR)**

Le modèle QSAR étudié est généré avec l'analyse par régression linéaire multiple (MLR) en utilisant les descripteurs calculés. Les paramètres intercorrélés ont été éliminés par étapes en fonction de leur corrélation avec l'activité biologique. Plusieurs équations sont obtenues.

**c. Validation du modèle QSAR**

Pour justifier la qualité statistique de toutes les équations obtenues, des paramètres tels que le coefficient de corrélation (R), le R<sup>2</sup> ajusté (R<sup>2</sup> A), le ratio de variance (F), l'erreur type d'estimation (EES) sont calculés.

Des analyses de corrélation sont effectuées et les paramètres intercorrélés ont été éliminés par étapes.

La validation des modèles est réalisée à l'aide de la méthode de validation croisée Leave-out (LOO) et des paramètres tels que PRESS, SSY, R<sup>2</sup> CV, SDEP et SPRESS.

**d. Résultats et discussions**

▪ **Résultats**

Les modèles générés dans l'étude QSAR sont les suivants :(85)

-  $pMIC50 = 4,1783 (\pm 0,694815) - 0,2399 (\pm 0,045) nO + 0,558 (\pm 0,193) IC2 + 0,0498 (\pm 0,010) T (N..O) - 3,642 (\pm 1,314) FEH11 - 22,303 (\pm 6,320) FEL9$

L'équation mentionnée ci-dessus est une équation statistiquement significative : n = 25 ; R = 0,883; R<sup>2</sup> = 0,781; RA<sup>2</sup> = 0,723; F (5, 19) = 13,524; p < 0,00001 SEE = 0,238; SSY = 4,8891 ; PRESS = 2.197; R<sup>2</sup>CV = 0,551; SDEP = 0,296; SPRESS = 0,340

- Un autre modèle statistiquement validé a été développé :

$pMIC50 = 4,097 (\pm 0,405) + 0,709 (\pm 0,121) Lop - 1,974 (\pm 0,539) FEH3 + 75,987 (\pm 35,937) FEH9 - 5.092 (\pm 1.095) FEL1$

n = 25; R = 0,841; R<sup>2</sup> = 0,707; RA<sup>2</sup> = 0,649; F (4, 20) = 12,094; p < .00004; VOIR = 0,267; SSY = 4,8891; PRESS = 2.006; R<sup>2</sup>CV = 0,590; SDEP = 0,283; SPRESS = 0.317

- A partir de ces 2 équations, d'autres modèles ont été validés on base sur la déletion des composés qui ont des résidus les plus grands.

- Les activités calculées de ces composés en utilisant toutes les équations finales sont montrées dans le tableau XXI.

**TABLEAU XX : Activités observées (Obs.), Calculées (Calc.) et résiduelles (Res.) [85]**

| Comp | Obs. | Eq. (9) |       | Eq. (11) |       | Eq. (13) |       | Eq. (15) |       | Eq. (17) |       |
|------|------|---------|-------|----------|-------|----------|-------|----------|-------|----------|-------|
|      |      | Calc.   | Res.  | Calc.    | Res.  | Calc.    | Res.  | Calc.    | Res.  | Calc.    | Res.  |
| 1    | 3.60 | 3.52    | 0.08  | 3.57     | 0.04  | 3.60     | 0.00  | 3.45     | 0.16  | 3.52     | 0.08  |
| 2    | 4.20 | 4.26    | -0.06 | 4.28     | -0.08 | 4.08     | 0.13  | 4.23     | -0.02 | 4.05     | 0.15  |
| 3    | 3.30 | 3.25    | 0.06  | 3.37     | -0.07 | 3.72     | -0.42 | 3.34     | -0.04 | 3.28     | 0.02  |
| 4    | 3.00 | 3.13    | -0.13 | 3.24     | -0.24 | -        | -     | 3.24     | -0.24 | -        | -     |
| 5    | 3.30 | 3.27    | 0.03  | 2.96     | 0.34  | 3.30     | 0.00  | 3.33     | -0.03 | 3.01     | 0.30  |
| 6    | 2.70 | 2.96    | -0.26 | 2.97     | -0.27 | -        | -     | 2.92     | -0.22 | 2.91     | -0.21 |
| 7    | 3.60 | -       | -     | 3.36     | 0.25  | 3.48     | 0.12  | 3.29     | 0.31  | 3.16     | 0.44  |
| 8    | 3.90 | 4.19    | -0.29 | 3.85     | 0.06  | 3.65     | 0.25  | 3.73     | 0.17  | 3.91     | -0.01 |
| 9    | 4.50 | 4.15    | 0.36  | 4.52     | -0.02 | 4.04     | 0.46  | 4.51     | -0.01 | 4.28     | 0.23  |
| 10   | 4.50 | 4.57    | -0.07 | 4.59     | -0.08 | 4.57     | -0.06 | -        | -     | 4.58     | -0.08 |
| 11   | 3.60 | 3.51    | 0.09  | 3.41     | 0.19  | 3.46     | 0.15  | 3.47     | 0.13  | 3.64     | -0.04 |
| 12   | 3.30 | 3.07    | 0.23  | 3.55     | -0.25 | 3.22     | 0.09  | 3.30     | 0.00  | 3.27     | 0.03  |
| 13   | 3.60 | 3.66    | -0.05 | 3.55     | 0.06  | 3.63     | -0.03 | 3.56     | 0.04  | 3.65     | -0.04 |
| 14   | 3.30 | 3.50    | -0.20 | 3.23     | 0.07  | 3.76     | -0.46 | 3.46     | -0.16 | 3.38     | -0.08 |
| 15   | 3.90 | 3.85    | 0.05  | 3.69     | 0.21  | 3.56     | 0.34  | 3.55     | 0.36  | 3.77     | 0.14  |
| 16   | 3.00 | 2.99    | 0.01  | 3.27     | -0.27 | 3.52     | -0.52 | 3.44     | -0.44 | 3.39     | -0.39 |
| 17   | 3.90 | 3.84    | 0.06  | 3.89     | 0.01  | 3.96     | -0.06 | 3.60     | 0.30  | 3.80     | 0.10  |
| 18   | 3.60 | 3.58    | 0.02  | 3.54     | 0.06  | 3.75     | -0.15 | 3.55     | 0.05  | 3.74     | -0.14 |
| 19   | 3.90 | 3.55    | 0.35  | 3.71     | 0.19  | 3.60     | 0.30  | 3.48     | 0.43  | 4.08     | -0.17 |
| 20   | 3.60 | 3.56    | 0.04  | 3.86     | -0.26 | 3.56     | 0.05  | 3.65     | -0.05 | 3.56     | 0.04  |
| 21   | 3.30 | 3.30    | 0.00  | 3.50     | -0.20 | 3.24     | 0.06  | 3.54     | -0.24 | 3.42     | -0.12 |
| 22   | 3.30 | 3.62    | -0.32 | -        | -     | 3.56     | -0.26 | -        | -     | -        | -     |
| 23   | 4.20 | -       | -     | 3.84     | 0.36  | 4.22     | -0.01 | 4.36     | -0.16 | 4.00     | 0.20  |
| 24   | 3.60 | 3.60    | 0.00  | -        | -     | 3.55     | 0.05  | 3.84     | -0.24 | 3.84     | -0.23 |
| 25   | 3.30 | 3.32    | -0.02 | 3.41     | -0.11 | 3.33     | -0.03 | 3.42     | -0.12 | 3.51     | -0.21 |

▪ **Discussion**

- L'indice de contenu informationnel IC2 et le descripteur topologique moléculaire T (N ... O) ont des effets positifs sur l'activité biologique. Les coefficients négatifs de FEH11 et FEL9 impliquent que, avec la diminution de ces valeurs, il peut y avoir des chances plus élevées d'attaque électrophile au nombre d'atomes 11 ainsi qu'une attaque nucléophile au nombre d'atomes 9 et peut conduire à de meilleurs composés antifongiques actifs.
- Lop et FEH9 ont des effets positifs sur l'activité biologique, FEH3 et FEL1 ont des effets négatifs. Ceux-ci signifient que l'attaque nucléophile peut être favorable à l'atome numéro 1 alors que l'attaque électrophile peut conduire à l'atome numéro 3.
- L'équation a également montré que si le nombre total d'atomes d'oxygène diminue, l'activité antifongique peut augmenter. La diminution de la densité électronique frontalière liée à l'orbitale moléculaire inoccupée la plus basse à l'atome numéro 9 suggère que l'attaque nucléophile peut être favorable à l'atome numéro 9. Cela peut être bénéfique en ce qui concerne l'activité antifongique.

- Les nombres d'atome 1 et 9 peuvent être favorables pour les attaques nucléophiles alors que les nombres d'atomes 3 et 11 peuvent être favorables à l'attaque électrophile.
- L'augmentation de la flexibilité moléculaire par incorporation de liaisons rotatives (RBF) peut être favorable à un effet antifongique.
- La figure ci-après illustre exigences des analogues de la coumarine pour l'activité

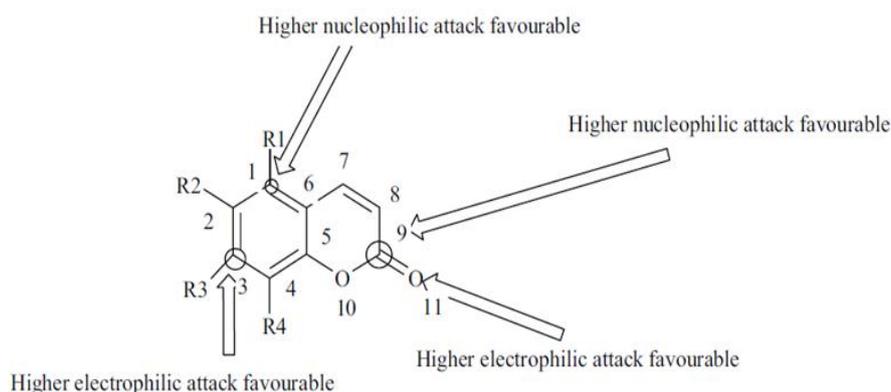


FIGURE 27 : Exigences des analogues de la coumarine pour l'activité anti-méningo encéphalitique [85].

### e. Conclusion

L'étude QSAR présentée par Tarun Jha et al [85] a réussi à nous expliquer les caractéristiques structurales requises pour les composés les plus actifs et a donné une carte géométrique pour la synthèse d'autres agents antifongiques pour traiter telles infections.

## 2. Bicoumarine

### 2.1. Etudes 3D-QSAR CoMFA sur la bis-coumarine analogues comme inhibiteurs de l'uréase.

#### a. Données expérimentales

L'uréase est une enzyme responsable de la transformation de l'urée en dioxyde de carbone et ammoniac. Elle est responsable de la formation de l'infection des calculs et de la lithiase urinaire, de la pyélonéphrite de l'ammoniaque et de l'encéphalopathie hépatique, et de l'incrustation urinaire des cathéters. Elle représente également une cause majeure de pathologies induites par *Helicobacter pylori* (HP).

Les stratégies basées sur l'inhibition de l'uréase sont maintenant considérées comme la première ligne de traitement des infections causées par l'uréase (produite par des bactéries).

**Zaheer-ul-Haq et al [87]**, ont réalisé une étude 3D-QSAR sur 30 dérivés bis-coumarine [ont été prises à partir de la littérature et a été récemment publiée par **Khan, K et al**] pour relier les structures chimiques avec leur activité inhibitrice d'uréase observée. (Figure 28 illustre la structure générale de bis coumarines étudié)

Le mécanisme actif du ligand et de la protéine ciblée reste inconnu pour la plupart des inhibiteurs d'uréase, c'est pour ça une étude de champ moléculaire comparatif analyse (CoMFA) a été utilisée.

L'étude est basée sur deux propriétés différentes : stérique, électrostatique, prétendue couvrir les principales contributions de l'activité. Une carte de contour électrostatique a été construite pour mieux comprendre les mécanismes potentiels des interactions intermoléculaires entre le ligand et le récepteur.

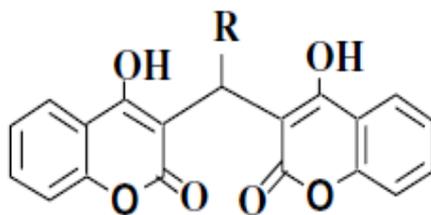


FIGURE 28 : Structure chimique générale de bis-coumarine étudié [87]

#### b. Etude QSAR –3D :

##### ▪ Analyses CoMFA

Les structures moléculaires de 30 dérivés sont modélisées avec le SYBYL 6.8 programme de modélisation moléculaire (Tripos Associates, Saint Louis, MO).

L'alignement structurel est l'un des plus importants paramètres dans les analyses 3D-QSAR. deux différentes méthodes d'alignement superposition et champ sont réalisées.

CoMFA est effectuée avec les champs stériques et électrostatiques habituellement utilisés dans SYBYL (Tripos Associate Ltd., St. Louis, MO (1999)).

Le QSAR produit par CoMFA, avec ses centaines ou milliers de termes, était habituellement représenté comme des contours, Les régions colorées dans les cartes représentent les zones dans l'espace 3D où les changements dans les valeurs sont fortement corrélés avec les changements concomitants dans les activités.

▪ **Méthode statistique : moindres carrés partiels PLS**

La méthode statistique des moindres carrés partiels PLS est utilisée pour simplifier les données de ce modèle QSAR, elle très utile dans les cas où le nombre de descripteurs est supérieur au nombre d'échantillons (composés).

**c. Validation du modèle QSAR**

Une validation croisée (LOO) est utilisée pour vérifier la capacité prédictive des modèles, et pour déterminer le nombre optimal de composants à utiliser dans les modèles QSAR finaux.

**d. Résultats et discussions**

▪ **Résultats**

Les résultats biologiques calculés par CoMFA sont montrés dans le tableau XXII qui montre activités expérimentales et activités prédictives de COMFA :

**TABLEAU XXI : Activités expérimentales et activités prédictives de CoMFA [87]**

| Composé   | Activité expr<br>( $\mu\text{M}$ ) | Activité prédit<br>( $\mu\text{M}$ ) | Composé | Activité expr<br>( $\mu\text{M}$ ) | Activité prédit<br>( $\mu\text{M}$ ) |
|-----------|------------------------------------|--------------------------------------|---------|------------------------------------|--------------------------------------|
| <b>1</b>  | 4.82                               | 4.81                                 | 16      | 4.13                               | 4.09                                 |
| <b>2</b>  | 4.20                               | 4.16                                 | 17      | 4.08                               | 4.05                                 |
| <b>3</b>  | 4.09                               | 4.10                                 | 18      | 4.17                               | 4.17                                 |
| <b>4</b>  | 4.10                               | 4.09                                 | 19      | 4.19                               | 4.19                                 |
| <b>5</b>  | 4.08                               | 4.08                                 | 20      | 4.36                               | 4.38                                 |
| <b>6</b>  | 4.28                               | 4.28                                 | 21      | 4.22                               | 4.23                                 |
| <b>7</b>  | 4.13                               | 4.14                                 | 22      | 4.23                               | 4.22                                 |
| <b>8</b>  | 4.07                               | 4.09                                 | 23      | 4.45                               | 4.47                                 |
| <b>9</b>  | 4.17                               | 4.20                                 | 24      | 4.14                               | 4.12                                 |
| <b>10</b> | 4.07                               | 4.06                                 | 25      | 4.12                               | 4.12                                 |
| <b>11</b> | 4.08                               | 4.07                                 | 26      | 4.09                               | 4.10                                 |
| <b>12</b> | 4.25                               | 4.22                                 | 27      | 4.15                               | 4.13                                 |
| <b>13</b> | 4.22                               | 4.24                                 | 28      | 4.21                               | 4.21                                 |
| <b>14</b> | 4.06                               | 4.06                                 | 29      | 4.04                               | 4.03                                 |
| <b>15</b> | 4.13                               | 4.13                                 | 30      | 4.33                               | 4.30                                 |

Le modèle étudié a donné de bons résultats prédictifs, ils sont représentés dans le tableau XXII qui montre les résultats de l'analyse COMFA.

TABLEAU XXII : Résultats de l'analyse CoMFA [87]

| COMFA                        |         |
|------------------------------|---------|
| q <sup>2</sup>               | 0.558   |
| r <sup>2</sup>               | 0.992   |
| Erreur standard d'estimation | 0.018   |
| F                            | 367.045 |
| Nombre de composés           | 30      |
| Fraction Stérique            | 0.711   |
| Electrostatique              | 0.289   |
| Espacement de la grille (Å°) | 2.0     |

Selon les résultats de CoMFA, les cartes de contours CoMFA ont été créées en utilisant les données de l'analyse PLS. Ces cartes nous aident à expliquer les caractéristiques stériques et électrostatiques des composés inclus dans l'analyse.

La figure 29 illustre les cartes de contours de champs électrostatiques et stériques COMFA.

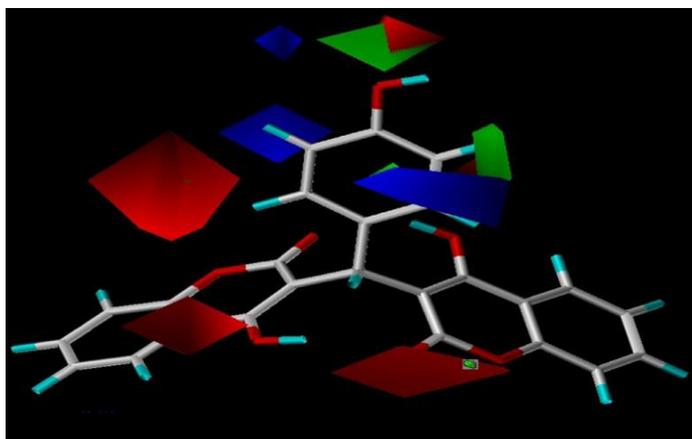


FIGURE 29 : Carte de contour de champ électrostatique et stérique CoMFA [87]

#### ▪ Discussion

Les cartes de contours électrostatiques sont montrées en tant que contours rouges et bleus.

- Les régions rouges indiquant que les groupes électronégatifs dans les ligands (la région benzénique du cycle et les molécules de lactone) interagissent favorablement avec les ions nickel positivement chargés dans le site actif, et ils donnent des activités d'inhibition accrue.

- La carte des contours bleus indique que les groupes chargés positivement dans le ligand ont une activité biologique améliorée et abaissent ainsi les IC<sub>50</sub> comme les composés 3, 4, 5. Ceci indique qu'une charge plus positive est favorisée dans cette région.

Les cartes de CoMFA stériques sont montrées en tant que contours verts et jaunes.

- Le contour jaune représente les régions dans lesquelles la masse stérique est préjudiciable à la liaison et augmente donc les IC<sub>50</sub> en diminuant l'activité inhibitrice globale.
- Le contour vert très petit peut être un site hydrophobe dans le récepteur, qui peut recevoir des groupes hydrophobes, par exemple les composés 2, 9, 12, 13, 17 et 13, avec des valeurs de pIC<sub>50</sub> de 4.21, 4.17, 4.22, 4.06, 4.08, et 4,45, respectivement ; ceci supporte également, tous les autres résultats précédemment publiés sur le SAR de l'uréase que les groupements hydrophobes très volumineux sont responsables de la diminution de l'activité.

#### e. Conclusion

Le modèle COMFA étudié a montré une bonne capacité prédictive ( $q^2$ ). Les courbes de contours 3D COMFA dans cette étude nous aident à comprendre les mécanismes sous-jacents de l'interaction des récepteurs avec les ligands.

Le modèle étudié a fourni des informations rationnelles pour la conception de nouveaux médicaments potentiels. [87]

## 2.2. Études QSAR de la 3,3 - (benzylidène substitué) -bis 4-hydroxycoumarine comme un inhibiteur potentiel de l'intégrase du VIH-1.

### a. Données expérimentales

Les médicaments anti-VIH (également appelés antirétroviraux) sont utilisés pour contrôler la reproduction du virus et ralentir ou stopper la progression de la maladie liée au VIH, l'intégrase du VIH-1 (VIH-1 IN) est apparue comme une cible thérapeutique importante pour la conception d'agents anti-VIH car l'intégration est essentielle pour la réplication du VIH (Le virus de l'immunodéficience humaine).

Récemment, des dérivés de la coumarine sont évalués dans le traitement du virus de l'immunodéficience humaine, en raison de leur capacité à inhiber le VIH-1.

Une série(9 molécules) de bis coumarines contre l'intégrase du VIH-1 sont soumise à une analyse quantitative de la relation structure-activité (QSAR) pour expliquer la cause de la différence de l'activité des composés contre HIV-1 IN en utilisant l'analyse de régression linéaire multivariable et d'obtenir l'information du mécanisme d'action au niveau moléculaire. [88]

Les IC<sub>50</sub> (concentration inhibitrice à 50%) des composés 1-9 contre HIV-1 IN ont été prises à partir de travail rapporté antérieur (Chiang, C et al), ces derniers sont convertis en pIC<sub>50</sub>.

- La figure 30 illustre la structure générale des composés étudiés.

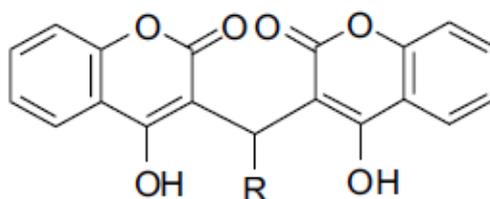


FIGURE 30 : Structure générale des dérivés de coumarines étudiées [88]

#### b. Descripteurs moléculaires et méthode statistique

##### ▪ Descripteurs moléculaires

Les structures moléculaires des composés dans les séries sélectionnées ont été dessinées en utilisant Chem Draw ultra 10.0 et ont été transférées au module Chem3D pour générer une structure tridimensionnelle (3D).

Les valeurs des descripteurs pour toutes les molécules ont été calculées en utilisant le module de programme "propriétés de calcul".

- Les descripteurs thermodynamiques : la température critique (T), l'énergie totale (TE),
- Les descripteurs physico-chimiques : coefficient de partage log P (lipophilie)
- Les descripteurs stériques : nombre de liaisons rotatives (NRB), réfractivité molaire (MR), ....
- Les descripteurs électroniques : énergie orbitale moléculaire occupée la plus élevée (HOMO), énergie orbitale moléculaire inoccupée la plus basse (LUMO), énergie de répulsion (RE), ...

##### ▪ Méthode statistique : moindres carrés partiel (PLS) et ACP

Pour obtenir un modèle fiable utilisant une régression linéaire multiple, Le nombre de composés devrait être quatre fois plus élevé que le nombre de descripteurs.

Dans cette étude publiée par **Bo-Jian Li et al** [88], le pool des descripteurs a été réduit en utilisant l'analyse des composantes principales (ACP) pour construire des modèles QSAR multilinéaires.

L'analyse de régression multiple qui consiste à trouver le meilleur ajustement de la variable dépendante (activités inhibitrices contre HIV-1 IN) à une combinaison linéaire de variables indépendantes (descripteurs) est utilisée par la méthode des moindres carrés.

L'analyse de corrélation et de régression de l'étude QSAR a donné 4 modèles. Chaque modèle est représenté par une équation :

➤ Modèle 1 :  $pIC_{50} = 2,942 (\pm 0,515) + 0,391 (\pm 0,084) \log P(1)$

➤ Modèle 2 :  $pIC_{50} = 2.139(\pm 1.060) + 0.279(\pm 0.155) \log P + 0.008(\pm .009) MR (2)$

➤ Modèle 3 :  $pIC_{50} = 3,673 (\pm 0,527) + 0,495 (\pm 0,081) \log P - 0,135 (\pm 0,061) NRB (3)$

➤ Modèle 4 :  $pIC_{50} = -1,903 (\pm 1,962) + 0,442 (\pm 0,066) \log P - 4,504 (\pm 1,788) \epsilon_{LUMO} (4)$

#### c. Validation du modèle QSAR

La Validation du modèle est une étape préliminaire pour déterminer le pouvoir prédictif et la capacité de corrélation. Pour cela, la validation croisée et les paramètres suivantes sont calculés :

Q2: validation croisée leave-one-out (LOO)

R : coefficient de corrélation multiple

R<sup>2</sup> : coefficient de détermination

F : test de signification

SEE : l'erreur-type de l'estimation

#### d. Résultats et discussions

##### ▪ Résultats

- La sélection effectuée par l'analyse des composantes principales (ACP) a donné les descripteurs les plus corrélés avec l'activité :

$\log P - LUMO - HOMO - NRB - MR$

- Les valeurs IC<sub>50</sub> prédites calculées par chaque modèle sont présentées dans le tableau XXIV qui présente les valeurs pIC<sub>50</sub> expérimentales et prédites.

**Tableau XXIII : Valeurs pIC50 expérimentales et prédites [88]**

| Composé | pIC50 exp | pIC50 prédit |       |       |       | Résiduel |        |        |        |
|---------|-----------|--------------|-------|-------|-------|----------|--------|--------|--------|
|         |           | Eq 1         | Eq 2  | Eq 3  | Eq 4  | Eq 1     | Eq 2   | Eq 3   | Eq 3   |
| N°      |           |              |       |       |       |          |        |        |        |
| 1       | 5.509     | 4.896        | 4.774 | 5.197 | 5.258 | -0.613   | -0.735 | -0.312 | -0.251 |
| 2       | 4.636     | 4.630        | 4.601 | 4.861 | 4.958 | -0.006   | -0.036 | 0.225  | 0.322  |
| 3       | 4.08      | 4.202        | 4.150 | 4.186 | 3.979 | 0.185    | 0.132  | 0.168  | -0.038 |
| 4       | 4.636     | 5.136        | 5.022 | 5.231 | 5.034 | 0.500    | 0.385  | 0.594  | 0.397  |
| 5       | 3.991     | 4.251        | 4.377 | 3.707 | 3.899 | 0.260    | 0.385  | -0.284 | -0.092 |
| 6       | 5.745     | 6.118        | 6.121 | 5.797 | 5.827 | 0.374    | 0.376  | 0.052  | 0.082  |
| 7       | 6.018     | 6.386        | 6.354 | 6.135 | 6.354 | 0.368    | 0.336  | 0.117  | 0.337  |
| 8       | 6.237     | 5.908        | 5.782 | 6.206 | 5.679 | -0.329   | -0.455 | -0.031 | -0.557 |
| 9       | 6.310     | 5.572        | 5.920 | 5.781 | 6.111 | -0.738   | -0.390 | -0.529 | -0.199 |

- Les résultats de validation obtenus (tableau 26) ont montré que les 4 modèles présentent un très bon pouvoir prédictif avec signification globale supérieur et une capacité d'expliquer la variance de l'inhibition entre les différentes molécules. Le modèle 4 est le meilleur.
- Le tableau XXV présente les paramètres statistiques des 4 meilleures équations.

**TABLEAU XXIV : Paramètres statistiques des 4 meilleures équations (le nombre de descripteurs ne dépasse pas 2). [88]**

| Eq  | ND <sup>a</sup> | R <sup>b</sup> | Rb <sup>2c</sup> | Q <sup>2d</sup> | F <sup>e</sup>             | P <sup>f</sup> | SEE <sup>g</sup> |
|-----|-----------------|----------------|------------------|-----------------|----------------------------|----------------|------------------|
| (1) | 1               | 0.871          | 0.758            | 0.634           | F <sub>(1,7)</sub> = 21.97 | 0.0022         | 0.486            |
| (2) | 2               | 0.886          | 0.786            | 0.151           | F <sub>(2,6)</sub> = 10.99 | 0.0098         | 0.495            |
| (3) | 2               | 0.931          | 0.868            | 0.960           | F <sub>(2,6)</sub> = 19.65 | 0.0023         | 0.388            |
| (4) | 2               | 0.939          | 0.993            | 0.768           | F <sub>(2,6)</sub> = 22.56 | 0.0016         | 0.366            |

#### ▪ Discussion

- Selon (eq1), il existe une relation significative entre log P et l'activité biologique (lipophilie)
- Selon (eq 2), log P et MR et ont été considérés comme des mauvais prédicteurs de l'activité en raison de sa mauvaise validation (Q2 = 0,151).

- Selon (eq 3), le coefficient négatif de NRB a révélé que la flexibilité moléculaire des biscoumarines est préjudiciable à l'activité.
- Selon (eq 4) le Lumo est fortement corrélé à l'activité. Le signe négatif du coefficient de Lumo indique que l'activité inhibitrice augmente lorsqu'il est diminué.

#### e. Conclusion

Le résultat de l'étude a montré que l'activité inhibitrice, était fortement corrélée avec les effets électroniques (LUMO) et la lipophiles présentés par la substitution (R).

L'analyse QSAR réalisée peut être utile pour élucider les mécanismes d'interaction chimio-biologique pour HIV-1 IN et dans la conception des puissants analogues inhibiteurs de la biscoumarine.

### 2.3. C. Etude QSAR 3D des dérives de coumarinyl triazoles mono et bis-aryloxy en tant qu'agents antituberculeux.

#### a. Données expérimentales

La conception de nouvelles entités moléculaires antituberculeuse reste un but essentiel en raison de l'émergence des souches de TB résistantes conduisant à des taux de mortalité élevés associés à cette maladie.

La combinaison d'un système hétérocyclique d'oxygène couplé à une fonctionnalité azotée ou vice versa est apparue comme une stratégie efficace dans la conception de nouvelles entités moléculaires pour la tuberculose.

Une série d'hybrides mono et bis-triazole-coumarine 6(a-u) et 9(a-f) est synthétisée dans les conditions de Click Chemistry par **Ashish Anand et al [89]**, l'activité antituberculeuse de ces composés est testée contre *M. tuberculosis* (H37Rv) en utilisant une microplaque Alamar Blue Assay (MABA). Les résultats sont exprimés en CIM.

Une étude QSAR 3D par COMFA et Topomer CoMFA, des cartes de contours 3D pour les champs stériques et électrostatiques de CoMFA sont tracées pour expliquer et montrer la relation entre les structures des composés synthétisés et l'activité antituberculeuse.

La figure 3 illustre la structure générale de dérivés des coumarines étudiés

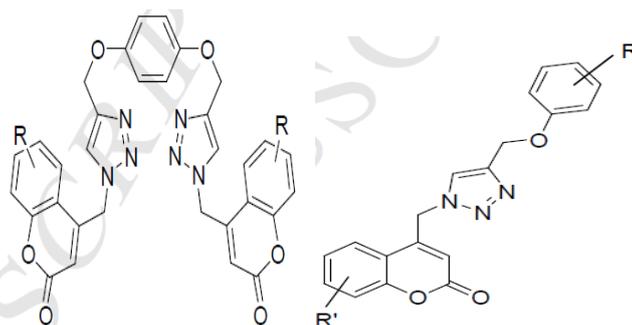


FIGURE 31 : Structure générale de coumarinyl triazoles mono et bis-aryloxy [89]

### b. Etude QSAR –3D :

#### ▪ Analyses CoMFA

Une analyse 3D-QSAR par CoMFA a été réalisée pour prédire les fractions favorables et défavorables des coumarinyl triazoles 6 et 9. L'ensemble de données utilisées dans cette étude contient 38 inhibiteurs de souches de *M. tuberculosis* H37Rv.

Les structures 3D des composés de Coumarinyl triazole sont construites en utilisant les paramètres géométriques standard du logiciel de modélisation moléculaire SYBYL-X 2.0. L'ensemble de données est divisé en ensemble d'apprentissage qui est utilisé pour produire des modèles 3D-QSAR, tandis que l'ensemble d'essai est utilisé pour valider la qualité du modèle, l'alignement des composés est effectué par la méthode base des données.

Topomer CoMFA est appliqué sur l'ensemble d'entraînement des composés en se divisant en fragments R1 et R2, le modèle Topomer CoMFA donne à la fois des résultats statistiques et graphiques (des courbes de niveau).

La carte de contours 3D représente la différence dans les domaines moléculaires qui sont associés aux différences dans l'activité biologique, qui à son tour, indique le lieu où les molécules alignées interagiraient favorablement ou défavorablement avec le récepteur.

#### ▪ Méthode statistique : Analyse par régression des moindres carrés partiels (PLS)

La régression PLS est une méthode statistique qui a un certain rapport avec la régression des composantes principales, cette méthode est utilisée pour analyser l'ensemble d'apprentissage en corrélant les valeurs des variables dépendantes (activité antituberculeuse) avec les variations de leurs variables indépendantes (champs d'interaction CoMFA / Topomer CoMFA). Le modèle avec un nombre optimal de composants (Q2 le plus élevé) et avec le SEP le plus bas a été sélectionné pour une analyse plus approfondie.

**c. Validation du modèle QSAR**

Afin d'obtenir le nombre optimal d'éléments conduisant au coefficient de validation croisée le plus élevé, la méthode q<sup>2</sup> leave one (LOO) a été utilisée et SEP qui indique la fiabilité et la capacité prédictive des modèles.

- Le q<sup>2</sup> et l'erreur standard de prédiction la plus basse (SEP) ont été calculés en utilisant

$$:q^2 = 1 - \frac{\sum_y (Y_{pred} - Y_{actual})^2}{\sum_y (Y_{actual} - Y_{signifier})^2} \quad (1)$$

Y<sub>pred</sub> est l'activité prédite, Y<sub>actual</sub> est l'activité expérimentale et Y<sub>signifiée</sub> est la meilleure estimation de la moyenne.

- L'erreur type d'estimation (SEE) ou SEP a été calculée comme suit :

$$SEE, SEP = \sqrt{\frac{PRESS}{n-c-1}} \quad (2)$$

Où n est le nombre de composés et c le nombre de composants

- La valeur de PRESS a été calculée comme suit :

$$PRESS = \sum_y (Y_{pred} - Y_{actual})^2 \quad (3)$$

**d. Résultats et discussions**

▪ **Résultats**

- Les résultats de validation de modèle QSAR sont illustrés dans le tableau XXV

**TABLEAU XXV : Paramètres statistiques des 2 modèles CoMFA et Topomer CoMFA par analyse PLS. [89]**

| <b>Paramètres statiques</b>    | <b>CoMFA<sup>a</sup></b> | <b>CoMFA<sup>b</sup></b> | <b>Topomer CoMFA</b> |
|--------------------------------|--------------------------|--------------------------|----------------------|
| q <sup>2</sup>                 |                          | 0.704                    | 0.496                |
| N                              | 7                        |                          | 6                    |
| r <sup>2</sup>                 | 0.985                    |                          | 0.997                |
| SEE                            | 0.037                    |                          | 0.027                |
| F                              | 18889.005                |                          | 12091.470            |
| r <sup>2</sup> <sub>pred</sub> | 0.991                    |                          | 0.991                |
| r <sup>2</sup> <sub>LOO</sub>  | 0.966                    |                          | 0.976                |
| r <sup>2</sup> <sub>CV</sub>   | 0.989                    |                          | 0.978                |
| r <sup>2</sup> <sub>bs</sub>   | 0.990                    |                          | 0.995                |
| SD <sub>bs</sub>               | 0.037                    |                          | 0.025                |

- Les meilleures prédictions ont été obtenues pour le modèle CoMFA<sup>a</sup> avec  $q^2 = 0,70$ ,  $r^2 = 0,98$ ,  $SEE = 0,03$ ,  $F = 18889,005$ .
- Tous les composés, 6 (a-u) ont présenté une activité modérée avec une CMI de 50 à 100  $\mu\text{g} / \text{ml}$ . L'activité modérée observée des mono triazoles nous a conduit à étendre le travail aux bis triazoles qui se sont révélés être de meilleurs candidats contre *M. tuberculosis*. (cf l'annexe 07)
- Tous les composés de la série 9 (a-f) ont montré une activité bonne-excellente avec un composé 9b ayant une CMI aussi faible que 0,2  $\mu\text{g} / \text{ml}$  qui s'est avérée meilleure que le médicament standard pyrazinamide ayant une CMI de 3,125  $\mu\text{g}/\text{ml}$ . Le composé 9c était hautement actif en tant que pyrazinamide avec une CMI de 3,125  $\mu\text{g}/\text{mL}$  (Voir annexe 07).

#### ▪ Discussion

- L'analyse des cartes de contour CoMFA stériques et électrostatiques mettant en évidence des régions vertes sur le cycle coumarinyl triazole et les phényles C2 et C3, a suggéré que des groupes stériquement encombrants sont favorables. Cependant, un grand contour rouge sur la 5ème et la 6ème position du cycle phényle et du groupe carbonyle proche du cycle coumarine du composé 9c montre que la présence d'un groupe chargé négativement augmenterait l'activité et un type similaire de carte de contour a été observé pour le composé 9f.
- La présence d'azotes dans le cycle triazole était impliquée dans l'interaction avec GLY96 et LYS165 au site de liaison de l'enzyme, ce qui a été confirmé par l'analyse Topomer CoMFA.
- Les bis triazoles 9 se sont avérés être des meilleurs candidats contre *M. tuberculosis* H37Rv ils ont montrés une meilleure liaison hydrogène en présence de deux cycles triazole, l'effet de la présence d'un groupe / atome plus encombrant attaché à l'anneau aryloxy, ce qui ressort clairement de l'activité, en outre la présence de deux cycles triazole ajoute jusqu'à la région favorisée du donneur de liaison hydrogène améliorant l'activité biologique globale.

#### e. Conclusion

Les modèles COMFA et Topomer COMFA étudiés ont montré une bonne capacité prédictive. Les courbes de contours 3D COMFA dans cette étude ont aidé à comprendre les mécanismes d'action nécessaires pour l'activité antituberculeuse.

Le modèle étudié a fourni des informations rationnelles pour la conception de nouveaux médicaments potentiels et doit être sélectionné pour une analyse plus approfondie. [89]

# **CONCLUSION GENERALE**

Notre présent travail se place dans le contexte d'une recherche fondamentale sur l'étude de relation structure activité QSAR sur différentes dérivés coumariniques.

L'importance de l'étude de la relation structure activité SAR (QSAR) a été prouvée par sa capacité à aider pour développer des modèles fiables pour la prédiction de quelques propriétés et activités biologiques des dérivés coumariniques. L'importance des dérivés des coumarines a été présentée aussi pour qu'ils constituent des sources naturelles des composés à intérêt thérapeutiques et dont l'étude fournit des informations structurelles qui vont guides les recherches dans le futur.

Différents articles publier et accepter sur les dérivés coumarines ont étaient rassemblées, ces études sont classées selon la substitution sur structure générale de coumarine, selon les activités biologiques (différentes activités biologiques sont étudiées).

Tous les modèles cités sont des modèles prédictifs, valides, et montrent une corrélation importante des descripteurs avec l'activité biologique étudiée.

Les auteurs peuvent proposer de nouveaux composés avec des valeurs d'activité supérieurs ou similaires que celles existantes (molécule de référence), en ajoutant des substituants appropriés.

Les modèles proposés donnent des orientations, dans le domaine de la recherche pharmaceutique, pour concevoir et synthétiser de nouvelles molécules susceptibles de devenir des médicaments.

Les connaissances structurelles acquises par le modèles QSAR pourraient être utilisées pour guider la conception rationnelle des nouvelles séries des composés actifs. Par conséquent, les modèles proposés réduiront le temps et le coût de la synthèse.

L'étude QSAR ne devrait pas être considérée comme un outil académique permettant la post-rationalisation des données. Les progrès dans les méthodologies QSAR dans les dernières années ont orienté vers des applications étendues dans la conception de médicaments et dans les nouveaux domaines tels que nano-QSAR (relations quantitatives entre nanostructures et activités), Modélisation QSAR des peptides (Les peptides antimicrobiens), QSAR Modélisation des mélanges chimiques.

Donc il convient d'utiliser d'avantages des modèles QSAR et de développer d'autre techniques et approches *in silico* pour découvrir d'autres molécules dans le futur, et réduire l'utilisation des animaux dans la recherche biomédicale.

**REFERENCES**  
**BIBLIOGRAPHIQUES**

1. Deeb O. Recent applications of quantitative structure-Activity relationships in drug design. *Medicinal Chemistry and Drug Design: In Tech*; 2012.
2. Patel HM, Noolvi MN, Sharma P, Jaiswal V, Bansal S, Lohan S, et al. Quantitative structure-activity relationship (QSAR) studies as strategic approach in drug discovery. *Medicinal Chemistry Research*. 2014;23(12):4991-5007.
3. Hulzebos E, Janssen P, Maslankiewicz L, Meijerink M, Muller J, Pelgrom S, et al. The application of structure-activity relationships in human hazard assessment: a first approach. *RIVM Rapport 601516008*. 2001.
4. Mannhold R, Krogsgaard-Larsen P, Timmerman H. *QSAR: Hansch analysis and related approaches*: John Wiley & Sons; 2008.
5. NEELY WB. *The Hansch Structure—Activity Approach As an Aid in Designing New Biologically Active Chemicals*. ACS Publications; 1972.
6. Cherkasov A, Muratov EN, Fourches D, Varnek A, Baskin II, Cronin M, et al. QSAR modeling: where have you been? Where are you going to? *Journal of medicinal chemistry*. 2014;57(12):4977-5010.
7. Topliss J. *Quantitative structure-activity relationships of drugs*: Elsevier; 2012.
8. Yousefinejad S, Hemmatee nejad B. Chemometrics tools in QSAR/QSPR studies: A historical perspective. *Chemometrics and Intelligent Laboratory Systems*. 2015;149:177-204.
9. Chuman H, Karasawa M, Fujita T. A Novel Three-Dimensional QSAR Procedure: Voronoi Field Analysis. *Quantitative Structure-Activity Relationships*. 1998;17(04):313-26.
10. Selassie C, Verma RP. History of quantitative structure-activity relationships. *Burger's Medicinal Chemistry and Drug Discovery*. 2003.
11. Pastor M, Cruciani G, McLay I, Pickett S, Clementi S. GRIND-IN dependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *Journal of medicinal chemistry*. 2000;43(17):3233-43.
12. Nikolova N, Jaworska J. SAR applicability domain.
13. Nikolova N, Jaworska J. Approaches to measure chemical similarity—a review. *Molecular Informatics*. 2003;22(9-10):1006-26.
14. Hansen OC. *Quantitative structure-activity relationships (QSAR) and pesticides*: Ministry of the Environment, Environmental Protection Agency; 2004.
15. Brown AC, Fraser TR. V.—On the connection between chemical constitution and physiological action. Part. I.—On the physiological action of the salts of the ammonium bases, derived from strychnia, brucia, thebaia, codeia, morphia, and nicotia. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*. 1868;25(1):151-203.
16. Free SM, Wilson JW. A mathematical contribution to structure-activity studies. *Journal of Medicinal Chemistry*. 1964;7(4):395-9.
17. De Benedetti PG, Fanelli F. Multiscale quantum chemical approaches to QSAR modeling and drug design. *Drug discovery today*. 2014;19(12):1921-7.

18. Perkins R, Fang H, Tong W, Welsh WJ. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. *Environmental Toxicology and Chemistry*. 2003;22(8):1666-79.
19. Blaney JM, Martin EJ. Computational approaches for combinatorial library design and molecular diversity analysis. *Current opinion in chemical biology*. 1997;1(1):54-9.
20. Hopfinger AJ, Duca JS. Extraction of pharmacophore information from high-throughput screens. *Current opinion in biotechnology*. 2000;11(1):97-103.
21. Karelson M. *Molecular descriptors in QSAR/QSPR*: Wiley-Interscience; 2000.
22. Todeschini R, Consonni V. *Molecular descriptors for chemoinformatics: volume I: alphabetical listing/volume II: appendices, references*: John Wiley & Sons; 2009.
23. Debnath AK. Quantitative structure-activity relationship (QSAR) paradigm--Hansch era to new millennium. *Mini reviews in medicinal chemistry*. 2001;1(2):187-95.
24. Khan AU. Descriptors and their selection methods in QSAR analysis: paradigm for drug design. *Drug discovery today*. 2016;21(8):1291-302.
25. Shahlai M. Descriptor selection methods in quantitative structure-activity relationship studies: a review study. *Chemical reviews*. 2013;113(10):8093-103.
26. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*. 2001;46(1-3):3-26.
27. Bajorath J. Selected concepts and investigations in compound classification, molecular descriptor analysis, and virtual screening. *Journal of chemical information and computer sciences*. 2001;41(2):233-45.
28. Puzyn T, Leszczynski J, Cronin MT. *Recent advances in QSAR studies: methods and applications*: Springer Science & Business Media; 2010.
29. Puzyn T, Suzuki N, Haranczyk M, Rak J. Calculation of quantum-mechanical descriptors for QSPR at the DFT level: is it necessary? *Journal of chemical information and modeling*. 2008;48(6):1174-80.
30. Zhou Z, Parr RG. Activation hardness: new index for describing the orientation of electrophilic aromatic substitution. *Journal of the American Chemical Society*. 1990;112(15):5720-4.
31. Huuskonen J, Livingstone D, Manallack D. Prediction of drug solubility from molecular structure using a drug-like training set. *SAR and QSAR in Environmental Research*. 2008;19(3-4):191-212.
32. Matsson P, Doak BC, Over B, Kihlberg J. Cell permeability beyond the rule of 5. *Advanced drug delivery reviews*. 2016;101:42-61.
33. Planey SL, Kumar R. Lipophilicity indices for drug development. *Journal of Applied Biopharmaceutics and Pharmacokinetics*. 2013;1(1):31-6.
34. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. *Analytica chimica acta*. 1986;185:1-17.

35. Latan H, Ramli NA. The results of partial least squares-structural equation modelling analyses (PLS-SEM). 2013.
36. Hoffman BT, Kopajtic T, Katz JL, Newman AH. 2D QSAR modeling and preliminary database searching for dopamine transporter inhibitors using genetic algorithm variable selection of Molconn Z descriptors. *Journal of medicinal chemistry*. 2000;43(22):4151-9.
37. Ajmani S, Jadhav K, Kulkarni SA. Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. *Journal of chemical information and modeling*. 2006;46(1):24-31.
38. Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*. 1988;110(18):5959-67.
39. Kim KH, Greco G, Novellino E. A critical review of recent CoMFA applications. *Perspectives in drug discovery and design*. 1998;12:257-315.
40. Norinder U. Recent progress in CoMFA methodology and related techniques. *Perspectives in drug discovery and design*. 1998;12:25-39.
41. Klebe G, Abraham U, Mietzner T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *Journal of medicinal chemistry*. 1994;37(24):4130-46.
42. Polanski J, Walczak B. The comparative molecular surface analysis (COMSA): a novel tool for molecular design. *Computers & chemistry*. 2000;24(5):615-25.
43. Gohlke H, Klebe G. DrugScore meets CoMFA: adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein. *Journal of medicinal chemistry*. 2002;45(19):4153-70.
44. Silber K, Heidler P, Kurz T, Klebe G. AFMoC enhances predictivity of 3D QSAR: a case study with DOXP-reductoisomerase. *Journal of medicinal chemistry*. 2005;48(10):3547-63.
45. Duca JS, Hopfinger AJ. Estimation of molecular similarity based on 4D-QSAR analysis: formalism and validation. *Journal of chemical information and computer sciences*. 2001;41(5):1367-87.
46. Vedani A, Dobler M. 5D-QSAR: the key for simulating induced fit? *Journal of medicinal chemistry*. 2002;45(11):2139-49.
47. Shao J. Linear model selection by cross-validation. *Journal of the American statistical Association*. 1993;88(422):486-94.
48. Golbraikh A, Shen M, Xiao Z, Xiao Y-D, Lee K-H, Tropsha A. Rational selection of training and test sets for the development of validated QSAR models. *Journal of computer-aided molecular design*. 2003;17(2-4):241-53.
49. Collantes ER, Xing L, Miller PC, Welsh WJ, Profeta S. Comparative molecular field analysis as a tool to evaluate mode of action of chemical hybridization agents. *Journal of agricultural and food chemistry*. 1999;47(12):5245-51.
50. Clark RD, Fox PC. Statistical variation in progressive scrambling. *Journal of computer-aided molecular design*. 2004;18(7-9):563-76.

51. Wold S, Eriksson L, Clementi S. Statistical validation of QSAR results. *Chemometric methods in molecular design*. 1995;309-38.
52. Rohini K, Srikumar P. Therapeutic role of coumarins and coumarin-related compounds. *Journal of Thermodynamics & Catalysis*. 2014;5(2):1.
53. Kostova I. Studying plant-derived coumarins for their pharmacological and therapeutic properties as potential anticancer drugs. *Expert opinion on drug discovery*. 2007;2(12):1605-18.
54. 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*. 2005;12(8):887-916.
55. Ojala T. Biological screening of plant coumarins. 2001.
56. Venugopala KN, Rashmi V, Odhav B. Review on natural coumarin lead compounds for their pharmacological activity. *BioMed research international*. 2013;2013.
57. Wu L, Wang X, Xu W, Farzaneh F, Xu R. The structure and pharmacological functions of coumarins and their derivatives. *Current medicinal chemistry*. 2009;16(32):4236-60.
58. Detsi A, Kontogiorgis C, Hadjipavlou-Litina D. Coumarin derivatives: an updated patent review (2015-2016). *Expert opinion on therapeutic patents*. 2017;27(11):1201-26.
59. Jain P, Joshi H. Coumarin: Chemical and pharmacological profile. 2012.
60. Murray RD. The naturally occurring coumarins. *Fortschritte der Chemie organischer Naturstoffe/Progress in the Chemistry of Organic Natural Products*: Springer; 2002. p. 1-619.
61. Bourgaud F, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, et al. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochemistry Reviews*. 2006;5(2-3):293-308.
62. 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*. 2002;12:53-4.
63. Asif M. Pharmacological activities and phytochemistry of various plant containing coumarin derivatives. *Current Science*. 2015;1(3):77-90.
64. 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*. 2010;17(13):1325-38.
65. Tosun A, Özkal N, Baba M, Okuyama T. Pyranocoumarins from *Seseligummiferum* subsp. *corymbosum* growing in Turkey. *Turkish Journal of Chemistry*. 2005;29(3):327-34.
66. Galm U, Heller S, Shapiro S, Page M, Li S-M, Heide L. Antimicrobial and DNA gyrase-inhibitory activities of novel clorobiocin derivatives produced by mutasynthesis. *Antimicrobial agents and chemotherapy*. 2004;48(4):1307-12.
67. Hishmat O, Miky J, Farrag A, Fadl-Allah E. Synthesis of some coumarin derivatives and their antimicrobial activity. *Archives of Pharmacal Research*. 1989;12(3):181-5.

68. Gleye C, Lewin G, Laurens A, Jullian J-C, Loiseau P, Bories C, et al. Acaricidal Activity of Tonka Bean Extracts. Synthesis and Structure– Activity Relationships of Bioactive Derivatives. *Journal of natural products*. 2003;66(5):690-2.
69. Kontogiorgis C, Detsi A, Hadjipavlou-Litina D. Coumarin-based drugs: a patent review (2008–present). *Expert opinion on therapeutic patents*. 2012;22(4):437-54.
70. Anand P, Singh B, Singh N. A review on coumarins as acetylcholinesterase inhibitors for Alzheimer’s disease. *Bioorganic & medicinal chemistry*. 2012;20(3):1175-80.
71. Verhoef TI, Redekop WK, Daly AK, Schie RM, Boer A, Maitland-van der Zee AH. Pharmacogenetic-guided dosing of coumarin anticoagulants: algorithms for warfarin, acenocoumarol and phenprocoumon. *British journal of clinical pharmacology*. 2014;77(4):626-41.
72. Dawood DH, Batran RZ, Farghaly TA, Khedr MA, Abdulla MM. New Coumarin Derivatives as Potent Selective COX-2 Inhibitors: Synthesis, Anti-Inflammatory, QSAR, and Molecular Modeling Studies. *Archiv der Pharmazie*. 2015;348(12):875-88.
73. Modh RP, Kumar SP, Jasrai YT, Chikhahia KH. Design, Synthesis, Biological Evaluation, and Molecular Modeling of Coumarin–Piperazine Derivatives as Acetylcholinesterase Inhibitors. *Archiv der Pharmazie*. 2013;346(11):793-804.
74. Latif NAA, Batran RZ, Khedr MA, Abdalla MM. 3-Substituted-4-hydroxycoumarin as a new scaffold with potent CDK inhibition and promising anticancer effect: Synthesis, molecular modeling and QSAR studies. *Bioorganic chemistry*. 2016;67:116-29.
75. Mladenović M, Mihailović M, Bogojević D, Matić S, Nićiforović N, Mihailović V, et al. In vitro antioxidant activity of selected 4-hydroxy-chromene-2-one derivatives—SAR, QSAR and DFT studies. *International journal of molecular sciences*. 2011;12(5):2822-41.
76. Worachartcheewan A, Suvannang N, Prachayasittikul S, Prachayasittikul V, Nantasenamat C. Probing the origins of aromatase inhibitory activity of disubstituted coumarins via QSAR and molecular docking. *EXCLI journal*. 2014;13:1259.
77. Kontogiorgis C, Nicolotti O, Mangiatordi GF, Tognolini M, Karalaki F, Giorgio C, et al. Studies on the antiplatelet and antithrombotic profile of anti-inflammatory coumarin derivatives. *Journal of enzyme inhibition and medicinal chemistry*. 2015;30(6):925-33.
78. Ambre PK, Pissurlenkar RR, Wavhale RD, Shaikh MS, Khedkar VM, Wan B, et al. Design, synthesis, and evaluation of 4-(substituted) phenyl-2-thioxo-3, 4-dihydro-1H-chromino [4, 3-d] pyrimidin-5-one and 4-(substituted) phenyl-3, 4-dihydro-1H-chromino [4, 3-d] pyrimidine-2, 5-dione analogs as antitubercular agents. *Medicinal Chemistry Research*. 2014;23(5):2564-75.
79. Kabeya LM, da Silva CH, Kanashiro A, Campos JM, Azzolini AEC, Polizello ACM, et al. Inhibition of immune complex-mediated neutrophil oxidative metabolism: a pharmacophore model for 3-phenylcoumarin derivatives using GRIND-based 3D-QSAR and 2D-QSAR procedures. *European journal of medicinal chemistry*. 2008;43(5):996-1007.
80. De Araújo RS, Guerra FQ, de O Lima E, De Simone CA, Tavares JF, Scotti L, et al. Synthesis, structure-activity relationships (SAR) and in silico studies of coumarin derivatives with antifungal activity. *International journal of molecular sciences*. 2013;14(1):1293-309.

81. Budzisz E, Brzezinska E, Krajewska U, Rozalski M. Cytotoxic effects, alkylating properties and molecular modelling of coumarin derivatives and their phosphonic analogues. *European journal of medicinal chemistry*. 2003;38(6):597-603.
82. Han S, Zhang F-F, Qian H-Y, Chen L-L, Pu J-B, Xie X, et al. Design, syntheses, structure–activity relationships and docking studies of coumarin derivatives as novel selective ligands for the CB2 receptor. *European journal of medicinal chemistry*. 2015;93:16-32.
83. Gomez JE, McKinney JD.M. tuberculosis persistence, latency, and drug tolerance. *Tuberculosis*. 2004;84(1):29-44.
84. Manvar A, Malde A, Verma J, Virsodia V, Mishra A, Upadhyay K, et al. Synthesis, anti-tubercular activity and 3D-QSAR study of coumarin-4-acetic acid benzylidene hydrazides. *European journal of medicinal chemistry*. 2008;43(11):2395-403.
85. Jha T, Chakraborty P, Adhikari N, Halder AK, Maiti M. QSAR study on coumarins as antimeningoencephalitic agents. *Internet Electron J Mol Des*. 2009;8:1-13.
86. Mobley H, Hausinger R. Microbial ureases: significance, regulation, and molecular characterization. *Microbiological reviews*. 1989;53(1):85-108.
87. Lodhi MA, Nawaz SA, Iqbal S, Khan KM, Rode BM, Choudhary MI. 3D-QSAR CoMFA studies on bis-coumarine analogues as urease inhibitors: A strategic design in anti-urease agents. *Bioorganic & medicinal chemistry*. 2008;16(6):3456-61.
88. Li BJ, Chiang CC, Hsu LY. QSAR Studies of 3,3'-(Substituted-Benzylidene)-Bis-4-Hydroxycoumarin, Potential HIV-1 Integrase Inhibitor. *Journal of the Chinese Chemical Society*. 2010;57(4A):742-9.
89. Anand A, Naik RJ, Revankar HM, Kulkarni MV, Dixit SR, Joshi SD. A click chemistry approach for the synthesis of mono and bis aryloxy linked coumarinyl triazoles as anti-tubercular agents. *European journal of medicinal chemistry*. 2015;105:194-207.

# **ANNEXES**

## Annexe 1

## L'approche "read-across"

Hypothèse la plus récente est également connue sous le nom de "read-across", ce qui signifie que la propriété inconnue d'un composé est la même que celle connue d'un autre composé si les deux composés sont suffisamment similaires.

## Annexe 2

## Les règles de Lipinski

C'est un ensemble des règles prédisent que l'absorption sera probablement faible quand

1. Le poids moléculaire est plus grand que 500.
2. Le nombre de liaisons hydrogène donneur est plus grand que 5.
3. Le nombre de liaisons hydrogène accepteur est plus grand que 10.
4. Le coefficient de partage [évalué par  $\log(P)$ ] est plus grand que 5.

## Annexe 3 : Structures chimiques des coumarines di substituées 1-34

- 1  $R^1 = \text{OCH}_2\text{C}_6\text{H}_5$ ,  $R^2 = \text{H}$
- 2  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-CH}_3\text{-}m$ ,  $R^2 = \text{H}$
- 3  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-F-}m$ ,  $R^2 = \text{H}$
- 4  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-Cl-}m$ ,  $R^2 = \text{H}$
- 5  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-CF}_3\text{-}m$ ,  $R^2 = \text{H}$
- 6  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-OCF}_3\text{-}m$ ,  $R^2 = \text{H}$
- 7  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-NO}_2\text{-}m$ ,  $R^2 = \text{H}$
- 8  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-F-}p$ ,  $R^2 = \text{H}$
- 9  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-Cl-}p$ ,  $R^2 = \text{H}$
- 10  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-OCH}_3\text{-}p$ ,  $R^2 = \text{H}$
- 11  $R^1 = \text{OCH}_2\text{C}_6\text{H}_4\text{-OCF}_3\text{-}p$ ,  $R^2 = \text{H}$
- 12  $R^1 = \text{OCH}_2\text{C}_6\text{H}_3\text{-3,5 di F}$ ,  $R^2 = \text{H}$
- 13  $R^1 = \text{OCH}_2\text{C}_6\text{H}_3\text{-3,4 di F}$ ,  $R^2 = \text{H}$
- 14  $R^1 = \text{OC}_6\text{H}_5$ ,  $R^2 = \text{H}$
- 15  $R^1 = \text{OC}_6\text{H}_4\text{-F-}m$ ,  $R^2 = \text{H}$
- 16  $R^1 = \text{OC}_6\text{H}_4\text{-Cl-}m$ ,  $R^2 = \text{H}$
- 17  $R^1 = \text{OC}_6\text{H}_4\text{-OCH}_3\text{-}m$ ,  $R^2 = \text{H}$
- 18  $R^1 = \text{OC}_6\text{H}_4\text{-CH}_3\text{-}p$ ,  $R^2 = \text{H}$
- 19  $R^1 = \text{OC}_6\text{H}_4\text{-Cl-}p$ ,  $R^2 = \text{H}$
- 20  $R^1 = \text{OC}_6\text{H}_4\text{-CN-}p$ ,  $R^2 = \text{H}$
- 21  $R^1 = \text{OC}_6\text{H}_4\text{-COCH}_3\text{-}p$ ,  $R^2 = \text{H}$
- 22  $R^1 = \text{OC}_6\text{H}_4\text{-N(CH}_3)_2\text{-}p$ ,  $R^2 = \text{H}$



- 23  $R^1 = \text{OC}_6\text{H}_3\text{-3,5 di F}$ ,  $R^2 = \text{H}$
- 24  $R^1 = \text{OC}_6\text{H}_3\text{-3,4 di F}$ ,  $R^2 = \text{H}$
- 25  $R^1 = \text{NHC}_6\text{H}_5$ ,  $R^2 = \text{H}$
- 26  $R^1 = \text{OH}$ ,  $R^2 = \text{H}$
- 27  $R^1 = \text{OCH}_3$ ,  $R^2 = \text{H}$
- 28  $R^1 = \text{OCH}_3$ ,  $R^2 = \text{C}_6\text{H}_5$
- 29  $R^1 = \text{OC}_6\text{H}_5$ ,  $R^2 = \text{C}_6\text{H}_5$
- 30  $R^1 = \text{OC}_6\text{H}_3\text{-3,4 di F}$ ,  $R^2 = \text{C}_6\text{H}_5$
- 31  $R^1 = \text{OC}_6\text{H}_5$ ,  $R^2 = \text{C}_6\text{H}_4\text{-Cl-}p$
- 32  $R^1 = \text{OC}_6\text{H}_5$ ,  $R^2 = \text{C}_6\text{H}_4\text{-CN-}p$
- 33  $R^1 = \text{OCH}_3$ ,  $R^2 = \text{H}$
- 34  $R^1 = \text{OC}_6\text{H}_5$ ,  $R^2 = \text{C}_6\text{H}_5$

### Annexe 4: Structures chimiques et descripteurs moléculaires correspondants des dérivés de la coumarine 1-18

|    | Structures  | $Clog P^*$ | $E_{LUMO}^{\dagger}$ |      |
|----|---|------------|----------------------|------|
| 1  | <br>(Umbelliferone)   | 1.62       | 0.07                 |      |
| 2  | <br>(Umbelliferone)   | 1.37       | 0.08                 |      |
|    |   |            |                      |      |
| 3  | $R_1$<br>-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> - | $R_2$      | 2.60                 | 0.08 |
| 4  | CH <sub>3</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub>             | H          | 3.06                 | 0.08 |
| 5  | $\begin{matrix} CH_3 \\   \\ CH-CH_2 \end{matrix}$  | H          | 2.40                 | 0.08 |
| 6  |   | H          | 3.04                 | 0.08 |
| 7  | CH <sub>3</sub> CH <sub>2</sub>   | H          | 1.47                 | 0.08 |
| 8  | -CH <sub>2</sub> -CH <sub>2</sub> -NH-CH <sub>2</sub> -CH <sub>2</sub> -                        |            | 1.31                 | 0.08 |
| 9  | -CH <sub>2</sub> -CH <sub>2</sub> -O-CH <sub>2</sub> -CH <sub>2</sub> -                         |            | 1.32                 | 0.08 |
| 10 | NH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub>   | H          | 0.42                 | 0.08 |
| 11 | H <sub>2</sub> N-CH <sub>2</sub> -CH <sub>2</sub> -CH <sub>2</sub>                              | H          | 0.76                 | 0.08 |
| 12 |   |            | 1.91                 | 0.06 |
| 13 |   | 1.13       | 0.03                 |      |
| 14 |   | 1.62       | 0.05                 |      |
| 15 |   | 1.51       | 0.05                 |      |
| 16 |   | 1.91       | 0.06                 |      |
| 17 |   | 1.13       | 0.03                 |      |
| 18 |   | 1.0        | 0.05                 |      |

### Annexe 5 : Structures chimiques et activité antifongique in vitro des composés 1-24- contre *A. fumigatus* et *A. flavus*

| Compound | Chemical Structures | <i>A. fumigatus</i>    |         |            |            | <i>A. flavus</i> |            |        |       |
|----------|---------------------|------------------------|---------|------------|------------|------------------|------------|--------|-------|
|          |                     | ATCC 16913 (log1/cygc) | IPP 210 | ATCC 46913 | ATCC 40640 | LM 247           | ATCC 16013 | LM 210 | LM 26 |
| 1        |                     | 1024 (2.15)            | 1024    | 1024       | 1024       | 1024             | 1024       | 1024   | 1024  |
| 2        |                     | 64 (3.40)              | 64      | 64         | 64         | 128              | 128        | 128    | 128   |
| 3        |                     | ≥2048 (1.89)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 4        |                     | ≥2048 (1.89)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 5        |                     | ≥2048 (1.89)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 6        |                     | ≥2048 (1.93)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 7        |                     | 1024 (2.26)            | 1024    | 1024       | 1024       | ≥2048            | 1024       | 1024   | 1024  |
| 8        |                     | ≥2048 (2.21)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 9        |                     | ≥2048 (1.97)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 10       |                     | ≥2048 (2.00)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 11       |                     | 64 (3.48)              | 64      | 64         | 64         | 256              | 64         | 64     | 64    |
| 12       |                     | 64 (3.50)              | 64      | 64         | 64         | 128              | 64         | 64     | 64    |
| 13       |                     | ≥2048 (2.16)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 14       |                     | 256 (2.89)             | 256     | 256        | 256        | 256              | 256        | 256    | 256   |
| 15       |                     | 16 (4.10)              | 16      | 16         | 16         | 16               | 16         | 16     | 16    |
| 16       |                     | 128 (3.25)             | 128     | 128        | 128        | 1024             | 1024       | 1024   | 1024  |
| 17       |                     | ≥2048 (2.05)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 18       |                     | ≥2048 (2.16)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 19       |                     | ≥2048 (1.99)           | ≥2048   | ≥2048      | ≥2048      | ≥2048            | ≥2048      | ≥2048  | ≥2048 |
| 20       |                     | 32 (3.80)              | 64      | 32         | 32         | 1024             | 32         | 32     | 64    |
| 21       |                     | 512 (2.68)             | 512     | 512        | 512        | 512              | 512        | 512    | 512   |
| 22       |                     | 16 (4.11)              | 16      | 16         | 16         | 16               | 16         | 16     | 16    |
| 23       |                     | 512 (2.69)             | 512     | 512        | 512        | 512              | 512        | 512    | 512   |
| 24       |                     | 1024 (2.30)            | 1024    | 1024       | 1024       | 1024             | 1024       | 1024   | 1024  |
| AmpB *   |                     | 2                      | 2       | 2          | 2          | 8                | 2          | 2      | 512   |

## Annexe 6 : Activités agonistes des composés 13e33 dans des cellules CHO exprimant le récepteur CB1 ou CB2, déterminées par des tests de mobilisation du calcium.

| No.       | R <sub>1</sub> | R <sub>2</sub> | MW <sup>a</sup> | ClogP <sup>b</sup> | PSA <sup>c</sup> , Å <sup>2</sup> | CB1 EC <sub>50</sub> , μM | CB2 EC <sub>50</sub> , μM (CB2 agonism, %) <sup>d</sup> | SI <sup>e</sup> |
|-----------|----------------|----------------|-----------------|--------------------|-----------------------------------|---------------------------|---|-----------------|
| 13        | Ethyl          |                | 315             | 3.47               | 73.66                             | >10                       | >10   | –               |
| 14        | n-Propyl       |                | 329             | 4.00               | 73.10                             | >10                       | >10   | –               |
| 15        | n-Butyl        |                | 343             | 4.53               | 73.40                             | >10                       | 0.144   | >69.4           |
| 16        | n-Pentyl       |                | 357             | 5.06               | 72.87                             | >10                       | 0.245   | >40.8           |
| 17        | n-Hexyl        |                | 371             | 5.58               | 72.81                             | >10                       | (36%)   | –               |
| 18        |                |                | 341             | 3.91               | 72.93                             | >10                       | (42%)   | –               |
| 19        | Bn             |                | 377             | 4.71               | 70.33                             | >10                       | (33%)   | –               |
| 20        | n-Butyl        |                | 351             | 4.46               | 80.51                             | >10                       | >10   | –               |
| 21        | n-Butyl        |                | 329             | 3.97               | 73.59                             | >10                       | 0.793   | >12.6           |
| 22        | n-Butyl        |                | 301             | 3.08               | 81.46                             | >10                       | (34%)   | –               |
| 23        | n-Butyl        |                | 315             | 3.47               | 77.99                             | >10                       | >10   | –               |
| 24        | n-Butyl        |                | 317             | 4.08               | 79.54                             | >10                       | >10   | –               |
| 25        | n-Butyl        |                | 317             | 3.73               | 61.62                             | >10                       | 0.972   | >10.3           |
| 26        | n-Butyl        |                | 395             | 5.15               | 62.52                             | >10                       | >10   | –               |
| 27        | n-Butyl        |                | 331             | 4.26               | 65.85                             | >10                       | (41%)   | –               |
| 28        | n-Butyl        |                | 331             | 4.39               | 68.84                             | >10                       | 0.130   | >76.9           |
| 29        | n-Butyl        |                | 331             | 4.39               | 70.71                             | >10                       | 0.321   | >31.2           |
| 30        | n-Butyl        |                | 303             | 3.33               | 71.38                             | >10                       | 0.103   | >97.1           |
| 31        | n-Butyl        |                | 344             | 3.74               | 78.63                             | >10                       | (37%)   | –               |
| 32        | n-Butyl        |                | 346             | 2.62               | 94.81                             | >10                       | >10   | –               |
| 33        |                |                | 329             | 3.04               | 65.16                             | >10                       | >10   | –               |
| GW842166X |                |                | 449             | 3.53               | 96.79                             | >10                       | 0.431   | >23.2           |
| CP55940   |                |                | 376             | 5.82               | 136.85                            | 0.007                     | 0.026   | 0.3             |

<sup>a</sup> MW: Molecular weight.

<sup>b</sup> ClogP: calculated by sybyl x1.3.

<sup>c</sup> PSA: polar surface area, calculated by sybyl x1.3.

<sup>d</sup> CB2 agonism (% basal) at 10 μM ligand concentration.

<sup>e</sup> SI: selectivity index for the CB2 receptor, SI = EC<sub>50</sub>(CB1)/EC<sub>50</sub>(CB2).

**Annexe 7 : Activité antituberculeuse in vitro des composés 6 (a-u) contre  
Mycobactérium tuberculosis H37Rv**

| Compounds     | R                  | R'                    | MIC (µg/mL) |
|---------------|--------------------|-----------------------|-------------|
| 6a            | 4-Br               | 7-CH <sub>3</sub>     | 50          |
| 6b            | 4-Br               | 6-CH <sub>3</sub>     | 50          |
| 6c            | 4-Br               | 6-Br                  | 50          |
| 6d            | 4-Br               | 7-Cl                  | 50          |
| 6e            | 4-OCH <sub>3</sub> | 7-CH <sub>3</sub>     | 50          |
| 6f            | 4-Cl               | 6-Br                  | 50          |
| 6g            | 4-OCH <sub>3</sub> | 7-Cl                  | 50          |
| 6h            | 4-Cl               | 7-CH <sub>3</sub>     | 50          |
| 6i            | 4-OCH <sub>3</sub> | 6-Br                  | 50          |
| 6j            | 4-Br               | 6-OCH <sub>3</sub>    | 50          |
| 6k            | 4-OCH <sub>3</sub> | 6-OCH <sub>3</sub>    | 50          |
| 6l            | 4-Cl               | 7-Cl                  | 50          |
| 6m            | 4-Cl               | 5,6-Benzo             | 50          |
| 6n            | 4-Br               | 7,8-Benzo             | 100         |
| 6o            | 4-OCH <sub>3</sub> | 5,6-Benzo             | 50          |
| 6p            | 4-Cl               | 6-CH <sub>3</sub>     | 50          |
| 6q            | 4-OCH <sub>3</sub> | 7-OCH <sub>3</sub>    | 50          |
| 6r            | 4-OCH <sub>3</sub> | 7,8-Benzo             | 50          |
| 6s            | 4-OCH <sub>3</sub> | 5,7-DiCH <sub>3</sub> | 50          |
| 6t            | 4-Br               | 5,6-Benzo             | 50          |
| 6u            | 4-OCH <sub>3</sub> | 6-CH <sub>3</sub>     | 50          |
| Pyrazinamide  | -                  | -                     | 3.125       |
| Streptomycin  | -                  | -                     | 6.25        |
| Ciprofloxacin | -                  | -                     | 3.125       |

| Compounds     | R                      | MIC (µg/mL) |
|---------------|------------------------|-------------|
| 9a            | 7-CH <sub>3</sub>      | 6.25        |
| 9b*           | 6-CH <sub>3</sub>      | 0.2         |
| 9c            | 7,8 Di-CH <sub>3</sub> | 3.125       |
| 9d            | 5,7 Di-CH <sub>3</sub> | 6.25        |
| 9e            | 7,8 Benzo              | 12.5        |
| 9f            | 7-OCH <sub>3</sub>     | 6.25        |
| Pyrazinamide  | -                      | 3.125       |
| Streptomycin  | -                      | 6.25        |
| Ciprofloxacin | -                      | 3.125       |

\* Compound with highest activity.

## Résumé

L'étude de la relation structure activité quantitative et qualitative (Q) SAR tente corrélér les descripteurs liés à la structure de la molécule avec l'activité biologique à l'aide des approches statistiques et mathématiques (étude in silico). Dans ce travail, nous essayons de résumer le statut de (Q)SAR et de résumer plusieurs recherches variées sur les coumarines en profitent de leur diversité structurelle comme une source inépuisable des molécules. Dans le but d'étudier cette relation structure activité existante entre la structure moléculaire des coumarines et leurs activités biologiques nous avons rassemblés plusieurs recherches publiées qui étudient plusieurs activités intéressantes en utilisant différents méthodes et techniques (Q)SAR. Utilisant des méthodes statistiques variées à savoir, le grand pouvoir prédictif de ces modèles, Les différentes informations structurelles sur l'activité étudiée qui pourrait également être utilisée pour concevoir de nouvelles séries d'agents thérapeutiques. Finalement, les résultats obtenus suggèrent l'importance de l'étude de la relation structure activité comme un axe principal utilisée en plusieurs disciplines notamment la conception des médicaments, ainsi que les dérivés de la coumarine comme des composés prometteurs et servant dans l'avenir comme des références afin d'approfondir leurs applications dans le futur.

**Mot clé :** SAR – QSAR – coumarines – descripteur moléculaire – validation de modèle – méthodes statistiques

## Abstract

The study of the relationship structure quantitative and qualitative activity (Q) SAR attempts to correlate the descriptors related to the structure of the molecule with biological activity using statistical and mathematical approaches (in silico study). In this work, we try to summarize the (Q) SAR status and summarize several varied researches on coumarins taking advantage of their structural diversity as an inexhaustible source of molecules. In order to investigate this relationship structure existing activity between the molecular structure of coumarins and their biological activities we have collected several published researches that study several interesting activities using different methods and techniques (Q) SAR. Using various statistical methods namely, the great predictive power of these models, the different structural information on the activity studied that could also be used to design new sets of therapeutic agents. Finally, the results obtained suggest the importance of studying the relationship structure activity as a main axis used in several disciplines including drug design, as well as coumarin derivatives as promising compounds and serving in the future as references to deepen their applications in the future.

**Keyword:** SAR - QSAR - coumarins - molecular descriptor - model validation - statistical methods

## ملخص

دراسة العلاقة الكمية والنوعية بين البنية والنشاط تعمل على ربط الخصائص البنوية وتأثيرها على نشاط الجزيء باستعمال طرق رياضية واحصائية. في عملنا المقدم نحاول تلخيص حالة وتلخيص العديد من الأبحاث المتنوعة حول الكومارين مع الاستفادة من تنوعها البنوي كمصدر لا ينضب من الجزيئات. من أجل دراسة هذه العلاقة التنشيطية القائمة بين البنية الجزيئية للكومارين وأنشطتها البيولوجية، قمنا بجمع العديد من الأبحاث المنشورة التي تدرس العديد من الأنشطة المثيرة للاهتمام باستخدام أساليب وتقنيات مختلفة. باستخدام مختلف الأساليب الإحصائية والقوة التنبؤية الكبيرة لهذه النماذج، فإن المعلومات الهيكلية المختلفة عن النشاط المدروس يمكن استخدامها أيضاً لتصميم مجموعات جديدة من العوامل العلاجية. وأخيراً، تشير النتائج التي تم الحصول عليها إلى أهمية دراسة نشاط الأدوية، بالإضافة إلى مشتقات الكومارين كمركبات بنية العلاقة كمحور رئيسي يستخدم في العديد من التخصصات بما في ذلك تصميم واعدة وتقديمها في المستقبل. مراجع لتعميق تطبيقاتهم في المستقبل.

**الكلمة الرئيسية:** - الكومارين - واصف جزيئي - التحقق من صحة النموذج - الأساليب الإحصائية