

REPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE  
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Université Frères Mentouri Constantine 1  
Faculté des Sciences de la Nature et de la Vie  
Département de Biochimie et Biologie cellulaire et moléculaire



### MEMOIRE DE FIN D'ETUDES

Mémoire présenté en vue de l'obtention du diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité : Biochimie appliquée

Intitulé :

**Etude comparative de l'efficacité prédictive des modèles 2D-QSAR  
et 3D-QSAR pour la conception de molécules anti-malaria.**

Présenté et soutenu par :

Le 20/06/2023

DABOUZ Radhia

BOUHEMHEM Sarra

Membres d'évaluation :

<b>Président</b>	Mr. MOKRANI ELHASSEN	MCA.UFM Constantine 1.
<b>Encadrant</b>	Mr. BOUHEDJAR Khalid	MRA.CRBT Constantine.
<b>Examineur</b>	Mm. CHERIET Mouna	MRB. CRBT Constantine.

Année universitaire : 2022/2023





## *Remerciement*

*On remercie Dieu le tout puissant de nous avoir donné la santé,  
la volonté et la patience d'entamer et de terminer ce mémoire.*

### *A notre encadrant*

*Nous souhaitons remercier vivement notre encadrant **Mr. Khalid BOUHEDJAR** chercheur au centre de recherche en biotechnologie (CRBt), pour sa patience, sa disponibilité, sa gentillesse et surtout ses bons conseils afin d'élaborer ce mémoire dans de bonnes conditions.*

### *A l'équipe du (CRBt)*

*Nous tenons à montrer notre gratitude à toutes les personnes qu'on a croisées au (CRBt), ce qui a contribué à créer une si bonne atmosphère au cours de notre séjour dans le laboratoire 18.*

### *Aux examinateurs :*

*Nous tenons également à remercier **Mr. MOKRANI ELHASSEN**. Maître de conférence A à l'université Frères Mentouri- Constantine1, pour avoir accepté de présider la soutenance de notre mémoire.*

*Nous exprimons nos vifs remerciements au **Mm. CHERIET Mouna**. Maitresse de recherche au centre de recherche en biotechnologie (CRBt) D'avoir pris le temps d'examiner notre travail.*

*Nous adressons nos sincères remerciements à tous les enseignants qui ont assuré notre formation durant les cinq ans d'étude et toutes les personnes qui ont contribué de près ou de loin durant tous notre cycle d'étude.*



## *Dédicace*

*Nous nous engageons sur ce travail*

*Dédié à nos parents, merci*

*Pour tous leurs sacrifices, leur amour,*

*Leur tendresse, Leur soutien et leurs prières*

*Tout au long de notre processus d'apprentissage,*

*Pour tout Leur aide et leur présence dans nos vies.*

*Merci à nos frères et sœurs pour leurs encouragements*

*Soutien permanent et moral*

*A tous nos amis et familles.*

### ملخص

تهدف هذه الاطروحة الى انشاء نماذج 2D-QDAR و 3D-QSAR موثوقة ومستقرة وتنبؤية.

استعملت منهجية QSAR في نمذجة النشاط المثبط ل PfENR الذي يمثل ل 61 من مشتقات Triclosan وذلك من خلال مقارنة نتائج البرنامجين (Sybyl و DRAGON-Mobydigs).

نتائج DRAGON-Mobydigs ( $Q^2 = 72.45\%/R^2 = 77.55\%$ ) حيث ان النشاط المثبط كان مرتبطا بأربع واصفات (DECC ,BEHe4 ,L2U ,C-026). وقد أثبتت هذه النتائج التي تم إنشاؤها بواسطة طريقة 2D-QSAR عمليات التحقق الداخلية والخارجية التي تؤكد استقرارها وقوتها وقدراتها التنبؤية.

أظهرت النتائج التي تم الحصول عليها بواسطة Sybyl لنموذج CoMFA ( $R^2 = 94.40\%$ ) ( $Q^2 = 64.33\%$ ), وبالنسبة ل CoMSIA ( $R^2 = 95.57\%$ ) ( $Q^2 = 53.51\%$ ), مما يشير ان النماذج موثوقة وقادرة على التنبؤ بالنشاط المدروس للمركبات الجديدة.

الكلمات الدالة : Triclosan, ENR, QSAR, CoMFA, CoMSIA.

### Résumé

Ce mémoire a pour objectifs d'établir des modèles 2D-QSAR et 3D-QSAR fiables, stables et prédictifs.

La méthodologie QSAR dont deux approches (DRAGON-Mobydigs et Sybyl) sont utilisées pour modéliser l'activité inhibitrice de l'enoyl acyl réductase représenté par  $pIC_{50}$  de 61 dérivés du Triclosan, en comparant leurs résultats, ainsi choisir le model le plus performant.

Pour le modèle DRAGON-Mobydigs ( $R^2 = 77.55\%$ ,  $Q^2 = 72.45\%$ ) dont l'activité inhibitrice a été corrélés avec 4 descripteurs (DECC, BEHe4, L2U, C-026). Ces résultats établis par la méthode 2D-QSAR ont prouvé leurs validations internes et externes confirmant leurs stabilités, leurs robustesses et leurs capacités prédictives.

Les résultats obtenus par sybyl pour le modèle CoMFA ( $R^2 = 94.40\%$ ,  $Q^2 = 64.33\%$ ) et pour le modèle CoMSIA ( $R^2 = 95.57\%$ ,  $Q^2 = 53.51\%$ ) qui indique que les modèles sont fiables et ils sont capables de prédire l'activité étudiée de nouveaux composés.

**Mots clés :** Triclosan, ENR, QSAR, CoMFA, CoMSIA.

### Abstrat

This study aims to develop reliable, stable and predictive 2D-QSAR and 3D-QSAR models.

The QSAR methodology, of which two approaches (DRAGON-Mobydigs and Sybyl) are used to model the inhibitory activity of enoyl acyl reductase represented by  $pIC_{50}$  of 61 Triclosan derivatives, by comparing their results, thus choosing the most suitable model perform.

For the DRAGON-Mobydigs model ( $R^2 = 77.55\%$ ,  $Q^2 = 72.45\%$ ) whose inhibitory activity was correlated with 4 descriptors DECC, BEHe4, L2U, C-026. These results established by the 2D-QSAR method have proven their internal and external validations confirming their stability, their robustness and their predictive capacities.

The results obtained by sybyl for the CoMFA model ( $R^2=94.40\%$  ,  $Q^2 = 64.33\%$ ) and for the CoMSIA model ( $R^2=95.57\%$  ,  $Q^2= 53.51\%$ ) which indicates that the models are reliable and they are able to predict the studied activity of new compounds.

Keywords: Triclosan, ENR, QSAR, CoMFA, CoMSIA.

## TABLE DES MATIERES

## LISTE DES TABLEAUX

## LISTE DES FIGURES

## LISTE DES SYMBOLES ET ABREVIATIONS

## INTRODUCTION GENERALE.....1

## PARTIE BIBLIOGRAPHIQUE

### Chapitre 1 : Généralités sur le paludisme

1. Introduction.....	4
2. Généralités sur le paludisme.....	5
2.1. Définition.....	5
2.2. Agents pathogène.....	5
2.3. Manifestations cliniques du paludisme.....	6
2.4. Cycle parasitaire du plasmodium.....	7
2.5. Principales méthodes de diagnostic du Paludisme .....	8
2.6. La vaccination .....	9
3. La conception de médicaments ou Drug Design.....	10
3.1. Définition d'un médicament .....	10
3.2. Les étapes de développement du médicament .....	10
3.2.1. La recherche.....	10
3.2.2. Les études précliniques.....	11
3.2.3. Les études cliniques.....	11
3.3. Les approches de drug design .....	12
3.3.1. Approche « Structure-based ».....	12
3.3.2. Approche « Ligand-based ».....	13
4. Classification et mode d'action des médicaments antipaludiques.....	13
4.1. Les quinoléines.....	13
4.2. Artémisinine et ses dérivés.....	13
4.3. Les antibiotiques.....	14



4.4. Les antifolates.....	14
4.5. L'atovaquone.....	14
5. La résistance aux médicaments antipaludiques.....	14
5.1. P. falciparum chloroquine resistance transporter (Pfcr1).....	14
5.2. P. falciparum multi-drug resistance 1 (Pfmdr1).....	15

### Chapitre 02 : La méthodologie QSAR

1. Généralités sur la relation quantitative structure activité (QSAR).....	16
1.1. Introduction à la modélisation QSAR.....	16
1.2. Aperçu historique.....	17
1.3. Objectifs du QSAR.....	18
2. La relation quantitative structure activité (2D-QSAR).....	18
2.1. Principe.....	18
2.2. Les descripteurs moléculaires.....	19
2.3. Construction d'un modèle 2D-QSAR.....	21
3. La relation quantitative structure activité (3D-QSAR).....	24
3.1. Définition.....	24
3.2. Principe.....	25
3.3. Étapes pour la procédure 3D-QSAR.....	25
3.4. Méthodes d'alignement moléculaires.....	26
4. Méthodes statistiques des modèles.....	28
4.1. La régression linéaire multiple.....	28
4.2. La régression de moindres carrés partiels.....	29
4.3. Paramètres d'évaluation des modèles.....	29
4.4. Interprétation et validation des modèles.....	30

## ETUDE EXPERIMENTALE

1. Matériels et méthodes.....	32
1.1. Ensembles de données expérimentales.....	32
1.2. Etude 2D-QSAR.....	32
1.3. Etude 3D-QSAR.....	34
2. Résultats et discussions.....	37
2.1. Construction d'un modèle 2D-QSAR.....	37
2.1.1. Analyse de la régression.....	38
2.1.2. Qualité de l'ajustement.....	40
2.1.3. Validation.....	41
2.1.4. Domaine d'applicabilité.....	42
2.1.5. Conclusion.....	43
2.2. Construction d'un modèle 3D-QSAR.....	43
2.2.1. Conclusion.....	46

## CONCLUSION GENERALE

## ANNEXES

## REFERENCES BIBLIOGRAPHIQUES

**Listes des tableaux**

<b>Tableau</b>	<b>Titre</b>	<b>Page</b>
Tableau 1	Les descripteurs classés en fonction de la dimension chimique.	19
Tableau 2	Les descripteurs intervenant dans les modèles.	37
Tableau 3	Paramètres statistiques du modèle 1.	38
Tableau 4	Paramètre statistiques du modèle 2.	39
Tableau 5	Paramètres statistiques du modèle 3.	39
Tableau 6	Paramètres statistiques du modèle 4.	40
Tableau 7	Les résultats pour le modèles PLS de CoMFA et CoMSIA	44

## Liste des figures

Figure	Titre	Page
Figure 1	Carte du monde représentant les zones où la population a des risques de contracter la malaria.	4
Figure 2	Structure de Plasmodium.	5
Figure 3	Cycle du plasmodium chez l'anophèle et chez l'homme.	7
Figure 4	Le vaccin antipaludique RTS,S.	9
Figure 5	Phases de découverte et du développement d'un médicament.	10
Figure 6	Classification des médicaments antipaludiques.	14
Figure 7	Structure prédite et haplotypes représentatifs du transporteur multi-drug resistance.	15
Figure 8	Modèle de l'étude de relation structure activité.	19
Figure 9	Construction d'un modèle 2D-QSAR.	23
Figure 10	Le processus de modélisation 3D-QSAR.	25
Figure 11	Workflow FB d'un modèle 3D-QSAR.	26
Figure 12	Structure centrale commune (échafaudage) du triclosan.	32
Figure 13	Droite d'ajustement pour le modèle 4 des valeurs expérimentales et prédits de la notation $pIC_{50}$ pour le modèle de QSAR.	41
Figure 14	Test de randomisation du modèle.	42
Figure 15	Diagramme de Williams.	42
Figure 16	Valeurs prédites de $pIC_{50}$ et valeurs expérimentales de $pIC_{50}$ pour les 61 molécules d'ensemble entraînement obtenues par analyse PLS en utilisant les modèles CoMFA et CoMSIA.	43
Figure 17	Composés du jeu de données alignés sur le triclosan.	44
Figure 18	Cartes de contours de (A) CoMFA et (B) CoMSIA en combinaison avec le composé 52.	45

## LISTE DES SYMBOLES ET ABREVIATIONS

### Listes de symboles et abréviations

Abreviation	Signification
2D	Structure deux dimensions.
3D	Structure tridimensionnelles.
ADME/Tox	Absorption, distribution, métabolisme et excrétion/Toxicité.
ADN	Acide désoxyribonucléique.
ABC	ATP binding cassette.
AMM	Autorisation de mise sur le marché.
AQ	Amodiaquine.
CoMFA	Comparative Molecular Field Analysis.
CoMSIA	Comparative Molecular Similarity Indices Analysis.
ENR	Enoyl Acyl réductase.
ESS	somme des carrés expliqués.
HUMO	orbitale moléculaire occupée de plus haute énergie.
G6PD	Glucose6-phosphate déshydrogénase.
IC <sub>50</sub>	Concentration inhibitrice.
LUMO	orbitale moléculaire non –occupée de plus basse énergie.
MLR	Régression linéaire multiple.
MM+	Mécanique Moléculaire (+).
OCDE	Organisation de coopération et de développement économiques.
OMS	Organisation mondial de la santé.
P	Coefficient de partage octanol-eau.
PCR	Polymerase Chain Reaction.
Pf	Plasmodium falciparum.
pIC <sub>50</sub>	Concentration inhibitrice médiane.
Pk	Plasmodium knowlesi.
pLDH	lactate déshydrogénase spécifique du parasite.

## LISTE DES SYMBOLES ET ABREVIATIONS

PLS	Partial Least Squares.
Pm	Plasmodium malariae.
PM3	Parametrization Method 3.
PNLP	Programme National de Lutte contre le Paludisme.
Po	Plasmodium ovale.
Pv	Plasmodium vivax.
Q	Cross-validated correlation coefficient.
Q <sup>2</sup>	Coefficient de validation externe.
QSAR	Relation quantitative-structure-activité.
QSPR	Relation quantitative-structure-propriété.
R&D	recherche et développement expérimental.
R	Coefficient de corrélation.
R <sup>2</sup>	Coefficient de détermination
R <sup>2</sup> <sub>adj</sub>	Coefficient de détermination ajusté.
RMN	Résonance Magnétique Nucléaire.
RMSE	Résonance Magnétique Nucléaire.
RSS	Somme des carrée résiduels
S	Standard error of estimation.
TCS	Triclosan.
TDR	Test de diagnostic rapide.
TSS	Somme des carrés totaux.

***Introduction  
générale***

## Introduction générale

Le paludisme est l'une des principales maladies infectieuses le plus répandue dans le monde, prédominant principalement dans les pays tropicaux et subtropicaux [1], l'émergence et la propagation de la résistance aux médicaments constituent une menace constante. La chloroquine (CQ) est depuis longtemps le médicament de choix pour le traitement du paludisme ; cependant, des parasites résistants à la CQ sont maintenant présents dans la plupart des régions où le paludisme est endémique [2], il existe donc un besoin urgent d'identifier ou de concevoir de nouveaux composés actifs avec de nouvelles activités inhibitrices [3], il peut être judicieux d'apprendre des essais biologiques antiplasmodiaux précédents et de prédire la bioactivité de ces produits naturels avant les essais biologiques expérimentaux [4].

L'énoyl acyl transporteur protéine réductase (ENR) est une enzyme cruciale dans la voie de synthèse des acides gras de type II de nombreux agents pathogènes tels que *Plasmodium falciparum*, l'agent étiologique de la forme la plus grave du paludisme. En raison de sa fonction essentielle de réduction des doubles liaisons des acides gras et de l'absence d'homologue humain, *PfENR* est une cible médicamenteuse intéressante. L'un des inhibiteurs de *PfENR* les plus étudiés est un médicament appelé triclosan (TCL) [5].

En effet, la mise sur le marché d'un médicament efficace représente un énorme investissement en temps et en argent. Plus précisément, il est estimé qu'il faut environ 2,6 milliards de dollars pour développer et mettre un médicament sur le marché, et cela prend en moyenne de 10 à 15 ans. De plus, le taux d'échec est élevé, atteignant près de 90 %, ce qui indique que près de 70 % des fonds sont investis dans des médicaments qui échouent aux essais cliniques en raison d'une efficacité insuffisante ou d'effets indésirables [6].

À l'heure actuelle, les progrès de la découverte de médicaments assistée par ordinateur (CADD) ont permis d'évaluer, d'améliorer et de cribler rapidement des milliers voire des millions de molécules, permettant ainsi de sélectionner rapidement un candidat médicament ciblant une protéine spécifique. Cette approche permet d'économiser considérablement de l'argent et du temps [6].



La connaissance des propriétés et des activités est d'une importance capitale pour pouvoir classer et utiliser les composés chimiques. La caractérisation expérimentale complète est difficile, voire impossible, pour des raisons de temps, de coût, de dangerosité de certains essais ou d'éthique (limitations des essais sur les animaux). L'utilisation des méthodes alternatives à l'expérience est devenue plus qu'indispensable. Parmi ces méthodes, on trouve les méthodes de modélisation moléculaire qui permettent de justifier les données expérimentales disponibles et prédire les propriétés/activités pour des composés nouveaux ou des composés pour lesquels les données expérimentales ne sont pas disponibles. Parmi ces méthodes de modélisation les plus utilisées, on peut citer les méthodes QSPR (Quantitative Structure-Property Relationships) et QSAR (Quantitative Structure-Activity Relationships) [7].

La relation quantitative structure activité (QSAR) est une méthode CADD qui implique la génération d'un modèle mathématique, qui relie les activités biologiques d'un ensemble de molécules à leurs descripteurs moléculaires (propriétés physicochimiques). Ces propriétés ont un impact clé sur l'activité du médicament [6].

Pour établir un modèle QSAR, trois éléments sont nécessaires. Le premier concerne l'activité biologique (par exemple la toxicité) mesurée pour un ensemble de molécules. La seconde concerne les descripteurs. Enfin, le troisième doit être un apprentissage statistique méthode utilisée pour connecter les deux premiers éléments [8].

Notre étude vise à développer un modèle robuste et fiable en se basant sur l'utilisation des méthodes 2D-QSAR et 3D-QSAR dans le but d'expliquer et prédire l'activité antiplasmodiale d'une série de 61 dérivés de triclosan inhibiteurs de l'enzyme acyl carrier réductase de *Plasmodium falciparum*.

Le manuscrit de ce mémoire est articulé sur deux grandes parties précédées d'une introduction générale :

- une partie bibliographique qui contient deux chapitres :
  - le 1<sup>er</sup> chapitre est intitulé «Généralités sur le paludisme» ou nous avons abordés dans un aperçu général sur le paludisme (L'agent pathogène, méthode de diagnostic, la vaccination, etc...), la conception de médicaments (Drug Design), les médicaments antipaludiques et les différentes formes de résistance envers les médicaments antipaludiques .

## INTRODUCTION GENERALE

---

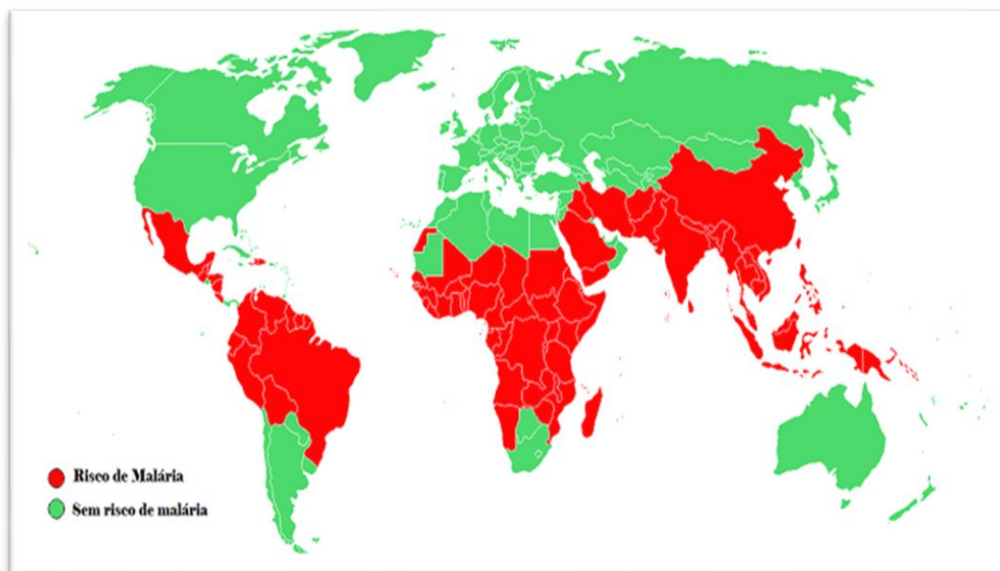
- Le 2<sup>ème</sup> chapitre nommé « La méthodologie QSAR » qui représente un rappelle sur la méthodologie QSAR ( un aperçu historique, principe générale, les descripteurs, les méthodes 2D-QSAR et 3D-QSAR, les méthodes statistiques ect...).
  - Une deuxième partie expérimentale où nous avons présenté, discuté et comparé les résultats obtenus par l'approche 2D-QSAR en utilisant les logiciels Chem-Draw, Hyper-Chem, Dragon, MobyDigs, ainsi que les résultats obtenus par la méthode 3D-QSAR en utilisant le logiciel Sybyl, que nous les avons traité tout au long de notre séjour au centre de recherche en biotechnologie de Constantine (CRBt).
- Enfin, nous clôturons par une conclusion générale suivie d'un annexe.

***Chapitre 1 : Généralités  
sur le paludisme***

## 1. Introduction

Le paludisme c'est une maladie parasitaire qui met gravement en danger la santé humaine et a également un impact négatif sur les économies et les sociétés dans lesquelles elles sévissent. *Plasmodium falciparum* est responsable de la majorité des décès dus au paludisme. Avec le sida et la tuberculose, il a été répertorié comme l'une des trois principales menaces mondiales pour la santé publique par l'Organisation mondiale de la santé (OMS) [9].

Selon le Rapport mondial sur le paludisme de l'OMS, le nombre de cas de paludisme est estimé à 247 millions en 2021 dans 84 pays d'endémie palustre, la plupart des cas supplémentaires sont estimés dans la région Afrique de l'OMS [10], la figure 1 montre les pays les plus menacés par le paludisme.



**Figure 1** : carte du monde représentant les zones où la population a des risques de contracter la malaria (en rouge la zone la plus à risque).

*Plasmodium falciparum*, le parasite le plus mortel responsable du paludisme [11], le développement d'une résistance de *P. falciparum* aux médicaments antipaludiques pose de sérieux problèmes pour le contrôle de la maladie [12], Cela met en évidence un besoin urgent de médicaments nouveaux et améliorés avec de nouveaux mécanismes d'action pour éviter la résistance [13].

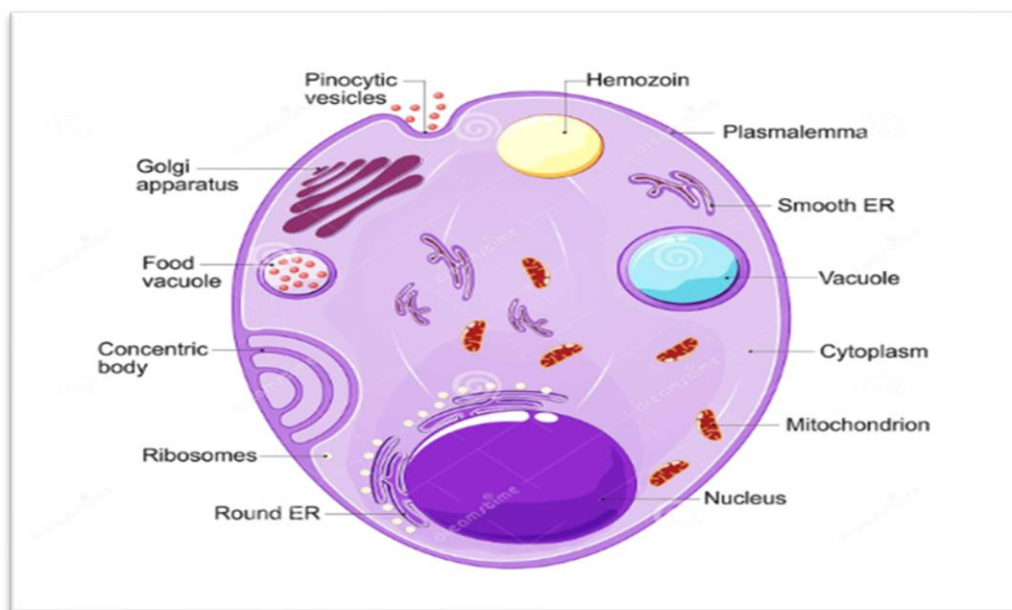
## 2. Généralités sur le paludisme

### 2.1. Définition

Le paludisme c'est une maladie infectieuse mortelle [14]. Causé par un parasite protozoaire appartenant à la genre Plasmodium [15]. Cette maladie parasitaire est transmise à l'homme par un moustique anophèle femelle lors d'un repas de sang [16].

### 2.2. Agent pathogène

Le Plasmodium est le parasite responsable du paludisme dans le monde (figure 2) . Ils'agit d'un protozoaire unicellulaire appartenant à l'embranchement Apicomplexa [17].



**Figure 2 :** Structure du Plasmodium.

Il existe cinq espèces de Plasmodium qui causent le paludisme chez l'homme, Plasmodium falciparum (Pf), Plasmodium vivax (Pv), Plasmodium ovale (Po), Plasmodium malariae (Pm), Plasmodium knowlesi (Pk) [18].

#### ➤ **P.faciparum**

Pf est l'espèce la plus répandue, en particulier dans les régions tropicales et subtropicales, présent en Afrique subsaharienne, en Asie du Sud-Est, en Amérique latine [19]. Pf est responsable de la « fièvre tierce maligne ». Sa période d'incubation est de 7 à 12 jours. Pf est responsable des formes cliniques graves, notamment du neuropaludisme [20].

➤ **P.vivax**

Pv est le parasite du paludisme le plus courant causant des maladies cliniques en dehors de l'Afrique [21], responsable de la « fièvre tierce bénigne ». Aujourd'hui, ce terme « bénin » n'est plus adapté, car de nombreuses études ont montré que P. vivax pourrait être l'origine d'un accès grave et compliqué, voire mortel [22]. Sa période d'incubation dure de 11 à 13 jours [22].

➤ **P.ovale**

Il sévit notamment en Afrique intertropicale et, à un moindre degré, en Asie du Sud-Est et dans certaines régions du Pacifique, provoque une « fièvre tierce bénigne ». Sa durée d'incubation est de 2 semaines au minimum. Comme chez Pv, des rechutes tardives jusqu'à 5 ans peuvent être observées [23].

➤ **P.malariae**

Il sévit dans toutes les zones d'endémie, mais de manière beaucoup plus sporadique. Il se caractérise par sa durée d'incubation beaucoup plus longue que les autres espèces plasmodiales humaines, de 15 à 21 jours [24]. Les patients ont de la fièvre tous les 72 heures lors d'une infection en raison de la durée plus longue cycle de vie du parasite [25]. Ainsi, Pm est considéré comme causant un paludisme chronique, ils peuvent durer des décennies. L'infection est bénigne mais Pm peut parfois entraîner des complications rénales [26].

➤ **P.knowlesi**

Il sévit en Asie du Sud-Est car la transmission de cette espèce plasmodiale est étroitement liée à la répartition géographique des singes macaques. La maladie se présente comme les autres paludismes avec de la fièvre, des frissons et des maux de tête avec des caractéristiques peu communes comme des nausées, vomissements, des myalgies, arthralgies, des symptômes des voies respiratoires supérieures et une jaunisse [27].

### **2.3. Manifestations clinique du paludisme**

➤ **Accès palustre non compliqué**

Les symptômes d'une infection aiguë par *Plasmodium* sont non spécifiques et comportent généralement céphalées, asthénie, douleurs musculaires, troubles digestifs, fièvre périodique [28].

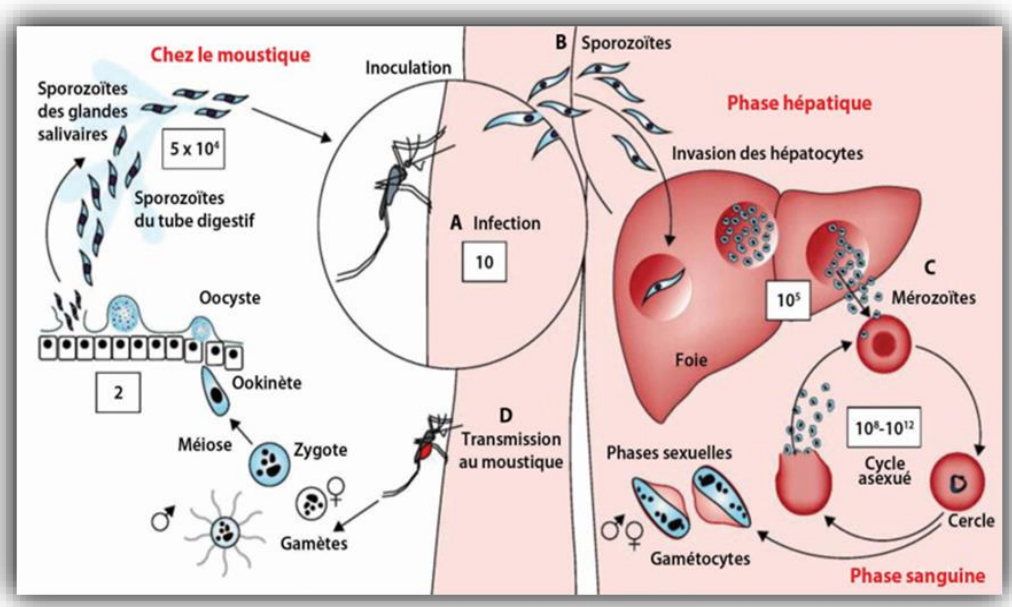
➤ **Accès palustre grave et compliqué**

Le paludisme grave est une urgence médicale, c'est une cause majeure de mortalité infantile dans les pays tropicaux, les complications, majoritairement dues à P. falciparum mais existant

aussi avec *P. vivax* [29], sont : atteinte multiviscérale avec ou sans atteinte neurologique (neuropaludisme), coma, convulsions, Anémie sévère, ictère ou hyperbilirubinémie, insuffisance rénale, acidose métabolique, syndrome de détresse respiratoire aiguë, hyperparasitémie, fausse-couche ou un retard de croissance dans l’utérus [30].

**2.4. Cycle parasitaire du Plasmodium**

Le cycle de développement parasitaire est pratiquement identique pour toutes les espèces de Plasmodium qui infectent l’homme, le cycle de vie se déroule chez deux hôtes : l’hôte intermédiaire (l’homme), où se déroule la reproduction asexuée (schizogonie) et l’hôte définitif (femelle de moustique), (Figure 3) [31].



**Figure 3 :** Cycle du plasmodium chez l’anophèle et chez l’homme.

**2.4.1. Chez l’homme**

➤ **Schizogonie pré-érythrocytaire**

le sporozoïte, est injectée dans la peau d’un vertébré par la femelle puis va pénétrer dans la circulation sanguine ou lymphatique et circulent pour infecter le foie, certains sporozoïtes parviennent à gagner le foie et à se multiplier dans les cellules hépatiques. Après multiplication, les parasites se transforment en schizontes pré-érythrocytaires, les cellules hépatiques vont éclater et libérer des milliers de mérozoïtes uninucléés dans le sang [31]. Chez certaines espèces plasmodiales comme *Pv*, *Po* restent en phase dormante dans le foie (hypnozoïtes) [32].

➤ **Schizogonie érythrocytaire**

les mérozoïtes libérés pénètrent dans la circulation sanguine et envahissent rapidement les érythrocytes [33]. Le cycle érythrocytaire se caractérise par une schizogonie, qui permet le passage du stade ring en trophozoïte, puis en schizonte [31]. À l'intérieur d'un érythrocyte, le parasite se nourrit de sa cellule hôte, puis se multiplie pour former davantage de mérozoïtes qui sortent et envahissent de nouveaux érythrocytes [32].

**2.4.2. Chez le moustique**

La poursuite du cycle de vie dépend maintenant de l'introduction des gamétocytes dans l'intestin d'un moustique femelle qui se nourrit, où les deux types de gamétocytes s'échappent de leurs cellules hôtes. Les gamétocytes mâles se divisent rapidement en un certain nombre de microgamètes flagellés mobiles dont chacun peut féconder un macrogamète femelle pour former un zygote. Le parasite devient alors un ookinète mobile, pénétrant la paroi intestinale du moustique puis sous forme d'oocyste arrondi. Le parasite se multiplie de manière asexuée à l'intérieur de celui-ci pour former plusieurs centaines de sporozoïtes mobiles (sporogonie) [17].

**2.5. Principales méthodes de diagnostic du Paludisme**

Le diagnostic du paludisme consiste en l'identification des parasites ou des antigènes du paludisme dans le sang de l'hôte [34].

➤ **Examen microscopique**

Il s'agit de la méthode standard pour diagnostiquer le paludisme, il se targue de pouvoir distinguer toutes les principales espèces de Plasmodium, la sensibilité et la spécificité de la microscopie optique dépendent des éléments suivants : la qualité de la lame colorée, le temps disponible pour lire un film sanguin et la compétence de la personne qui manipule le microscope [35].

➤ **Méthode moléculaire**

La méthode PCR amplifie l'acide désoxyribonucléique (ADN) du parasite, la PCR utilise un thermocycleur pour amplifier l'ADN en utilisant différentes températures pour la dénaturation et l'hybridation [36]. Les principales limites de la PCR incluent la



complexité de son application en milieu clinique, le coût élevé impliqué et le long délai d'exécution [37].

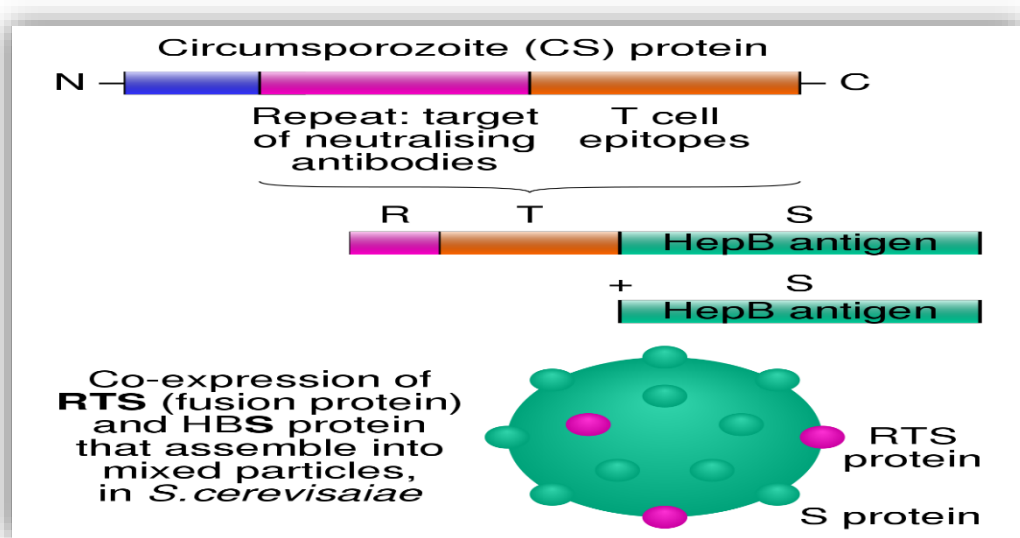
➤ **TDR**

Les TDR sont des méthodes immunochromatographiques qui utilisent des anticorps monoclonaux sur une bandelette de test, qui sont dirigés contre des enzymes ou des antigènes spécifiques du parasite présents dans le sang des individus infectés produisant un changement de couleur visible [38]. Les cibles actuelles des TDR sont Pf HRP 2 ou des enzymes de la voie glycolytique du parasite pLDH ou l'aldolase de Plasmodium [39], [40].

La technique implique l'application d'un échantillon de sang sur la carte de test suivi d'un réactif tampon (3 à 5 gouttes). La présence de bandes spécifiques dans la carte de test (après 15 à 30 min) indique la présence de plasmodium [41].

**2.6. La vaccination**

En juillet 2015, le vaccin antipaludique le plus avancé, le RTS,S fabriqué par GlaxoSmithKline (GSK), a reçu pour la première fois un avis positif des régulateurs européens [42], le nouveau vaccin a été désigné RTS,S pour indiquer la présence de la région de répétition CSP (R), épitopes de cellules T (T) fusionnés au HBsAg (S) et assemblés avec des copies non fusionnées de HbsAg (Figure 4) [43]. L'antigène vaccinal candidat RTS,S contre le paludisme cible la protéine circumsporozoïte Pf (CSP), une protéine de 412 acides aminés présente à la surface des sporozoïtes [44].



**Figure 4 :** Le vaccin antipaludique RTS,S.

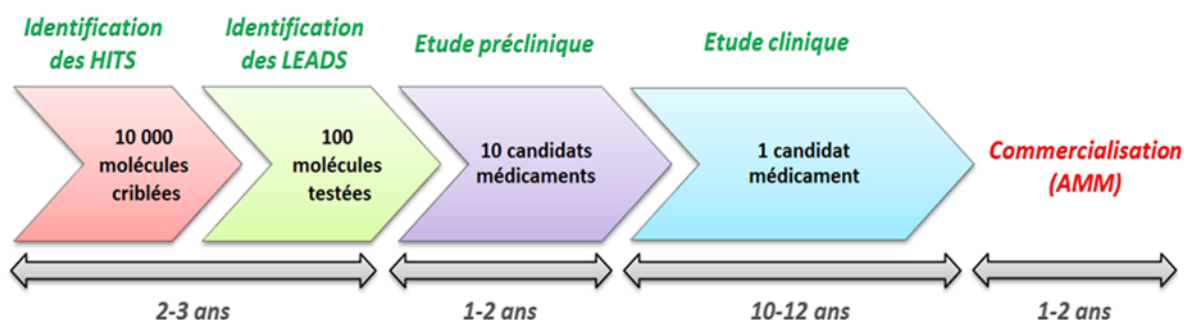
### 3. La conception de médicaments ou drug design

#### 3.1. Définition d'un médicament

Toute substance ou composition présentée comme possédant des propriétés curatives ou préventives à l'égard des maladies humaines ou animales. Un médicament est composé d'un principe actif, qui est la substance lui conférant ses propriétés thérapeutiques, associé à des excipients, qui sont des substances inactives d'un point de vue pharmacologique. Les excipients peuvent jouer un rôle dans la stabilité et l'absorption du principe actif ou encore conditionner l'aspect, la couleur ou le goût du médicament [44].

#### 3.2. Les étapes de développement de médicament

Le cycle de vie d'un médicament (Figure 5), est encadré par de nombreuses réglementations marquants ainsi les différentes étapes portant du développement de la molécule active jusqu'à la distribution du médicament sur le marché. Sur 10 000 molécules testées, une seule sera commercialisée. C'est donc un processus extrêmement long et coûteux : il faut compter minimum 15 ans de recherche pour un coût total avoisinant le milliard d'euros [45].



**Figure 5** : Phases de découverte et du développement d'un médicament .

##### 3.2.1. La recherche

Elle vise à la détermination d'une cible biologique, va permettre de découvrir et optimiser des molécules actives sur une cible biologique [46], cette étape comprend :

###### ➤ Identification de la cible biologique

Il s'agit d'une entité biologique qui peut être une protéine, un gène ou un ARN. La liaison de la cible à la molécule médicamenteuse provoque une réponse biologique qui peut être mesurée à la fois in vitro et in vivo [47].

### ➤ Identification des hits

Une fois la cible identifiée et validée, l'étape suivante est logiquement de tenter de moduler ses effets pour pouvoir agir favorablement sur le processus pathologique, et ceci par interaction avec des molécules. Des composés capables de tels effets sont appelés des touches ou « hits ». La recherche de hits se fait par criblage de larges banques de composés aussi appelées « chimiothèques », soit *in vitro*, le plus souvent grâce aux techniques d'HTS, soit *in silico* par criblage virtuel sur ordinateur. Les processus actuels de R&D utilisent très fréquemment une combinaison de ces deux techniques [48].

### ➤ Génération et optimisation des leads

Cette étape nécessite l'intervention des chimistes médicaux et de chimoinformaticiens qui vont tenter, à partir des hits précédemment identifiés, d'obtenir de nouvelles molécules que l'on qualifiera de têtes de séries ou « leads ». Les leads idéaux sont des composés qui sont plus actifs et plus sélectifs que les hits dont ils sont issus, tout en présentant des propriétés pharmacocinétiques optimales [47]. Pour obtenir ces leads, de vastes études de relations structure-activité (RSA) sont menées en faisant varier les structures chimiques des hits par modification des groupements fonctionnels tout en gardant leurs squelettes de bases. Lors de ces études, les informations (activité, sélectivité, propriétés physico-chimiques,...) des composés nouvellement synthétisés sont comparées à celles des hits et permettent de guider la sélection de nouveaux leads [48].

### 3.2.2. Les études précliniques

La deuxième phase concerne les tests précliniques. Les molécules sélectionnées sont testées *in vitro* et sur des modèles animaux afin d'évaluer leurs activités biologiques, leurs toxicologies et leurs pharmacocinétiques. La faisabilité technique du médicament est aussi évaluée, c'est-à-dire les différentes méthodes nécessaires pour tester ces molécules sont déterminées et les moyens financiers à mettre en œuvre sont définis. A la fin de cette étape, seules 10 molécules candidates sont retenues. Enfin, ces molécules rentrent en phase de développement clinique [49].

### 3.2.3. Les études cliniques

Les essais cliniques constituent l'étape critique de tout processus de R&D mais aussi l'étape la plus longue et la plus coûteuse. L'efficacité du candidat médicament chez l'homme, ainsi que sa pharmacocinétique et sa sécurité d'emploi sont évaluées et consignées en vue de la demande d'autorisation sur le marché (AMM) [50]. L'administration se faisant chez l'homme, l'encadrement de ces essais au niveau législatif et réglementaire est très strict [48].

➤ **Phase I «phase d'innocuité»**

Sont généralement effectuées sur des volontaires humains en bonne santé et sur un petit groupe de personnes [45].

➤ **Phase II « tests de biodisponibilité ou d'efficacité »**

Réalisée chez quelques centaines de patients, elle poursuit les études pharmacocinétiques et pharmacodynamiques. Elle permet de vérifier ou non l'hypothèse d'efficacité pharmacologique du produit, et de préciser la posologie optimale à utiliser lors de la phase III [51].

➤ **Phase III « double aveugle »**

Celle-ci consiste en une confirmation du potentiel thérapeutique du candidat médicament sur une très large population de malades puisque plusieurs milliers de patients peuvent être inclus. Lors de cette phase, une confirmation de l'efficacité est recherchée et peut être comparée avec celle du traitement de référence et le profil de tolérance (interactions, effets indésirables...) est établi [51].

➤ **Phase IV « post marketing » ou de « pharmacovigilance »**

C'est après l'autorisation de mise sur le marché (AMM) après que le médicament a été lancé sur le marché que viendra cette phase. La société poursuit sa surveillance du médicament. Le raisonnement derrière cette phase est de vérifier toute nouvelle réaction indésirable grave ou qui n'était pas détectée dans les phases antérieures et qui peut être observé dans cette phase. S'il se trouve qu'une réaction indésirable grave est observée, la société peut retirer le médicament du marché [49].

### **3.3. Les approches de drug design**

Il existe deux grands types de conception de médicaments. les approches fondées sur la connaissance des ligands (ligand-based) et les approches fondées sur la structure du récepteur (structure-based, comme par exemple le docking). Si les secondes approches ne peuvent être employées que lorsque la structure 3D du récepteur est connue, les méthodes basées sur les ligands sont particulièrement pertinentes quand les informations sur le récepteur sont peu abondantes [52].

#### **3.3.1. Approche « Structure-based »**

Les approches structure-based s'appuient sur la connaissance de la structure de la cible visée pour calculer les énergies d'interaction entre la protéine cible et les composés testés. Ces

approches nécessitent d'avoir identifié la cible biologique impliquée dans la pathologie et de disposer de données structurales de la cible. Idéalement, la structure tridimensionnelle disposant d'une haute résolution est déterminée à l'aide de la cristallographie aux rayons X ou encore de la RMN. Cependant, il arrive parfois que la structure de la cible visée ne soit pas encore élucidée. Il est alors possible de construire la structure 3D de la macromolécule étudiée à l'aide des approches par homologie. Cette structure tridimensionnelle peut être employée lors d'un amarrage moléculaire (docking) [53].

### **3.3.2. Approche « Ligand-based »**

Les approches ligand-based sont généralement préférées lorsqu'il n'y a pas ou peu de données sur la structure de la cible. Ces approches utilisent les recherches existantes sur les ligands connus comme étant actifs ou inactifs par rapport à la cible étudiée, elles font l'hypothèse que des structures similaires du point de vue moléculaire peuvent exercer une activité biologique équivalente. Ainsi, des structures chimiques d'interactions moléculaires comparables vis-à-vis de la cible vont avoir des activités similaires. Cependant, deux molécules d'activité équivalente peuvent avoir des structures chimiques totalement différentes. Cette approche a deux modèles très importants qui sont la modélisation de pharmacophore et Quantitative Structure Activity Relationship (QSAR) [54].

## **4. Classification et mode d'action des médicaments antipaludiques**

### **4.1. Les quinoléines**

Les médicaments dérivés des quinoléines agissent principalement sur les stades sanguins du parasite, mais certains agents antipaludiques agissent au niveau hépatique. Ces médicaments inhibent en formant un complexe avec l'hème dans la vacuole alimentaire du parasite. Ainsi, la polymérisation de l'hème est bloquée. En conséquence, l'hème libéré lors de la dégradation de l'hémoglobine s'accumule à des niveaux toxiques, tuant ainsi le parasite avec ses déchets toxiques [55].

### **4.2. Artémisinine et ses dérivés**

L'artémisinine inhibe toutes les phases parasitaires au sein des érythrocytes, en particulier au début de leur développement. Elle inhibe également la transmission des gamétocytes des humains aux moustiques [56], les dérivés de l'artémisinine sont recommandés pour le traitement en association avec d'autres agents antipaludiques, il s'agit de la thérapie combinée à base d'artémisinine (ACT) [57].

**4.3. Antibiotiques**

Les antibiotiques agissent sur le parasite en bloquant certaines fonctions de l'apicoplaste tels que la transcription des apicoplastes, la réplication du génome de l'apicoplaste ou le blocage des systèmes de traduction procaryotes (doxycycline, spiramycine, clindamycine) [58].

**4.4. Les antifolates**

Les antifolates sont des agents antipaludéens qui inhibent la synthèse de l'acide folique, qui est cruciale pour la synthèse des nucléotides et des acides aminés [59].

**4.5. L'atovaquone**

C'est le premier agent antipaludéen approuvé qui cible les mitochondries de Plasmodium . L'atovaquone inhibe le transport des électrons en bloquant les parties du cytochrome b du complexe du cytochrome bc1, en agissant comme un analogue de l'ubiquinone. Lorsqu'elle est associée au proguanil, l'atovaquone est sûre et efficace pour les femmes enceintes et les enfants [60].

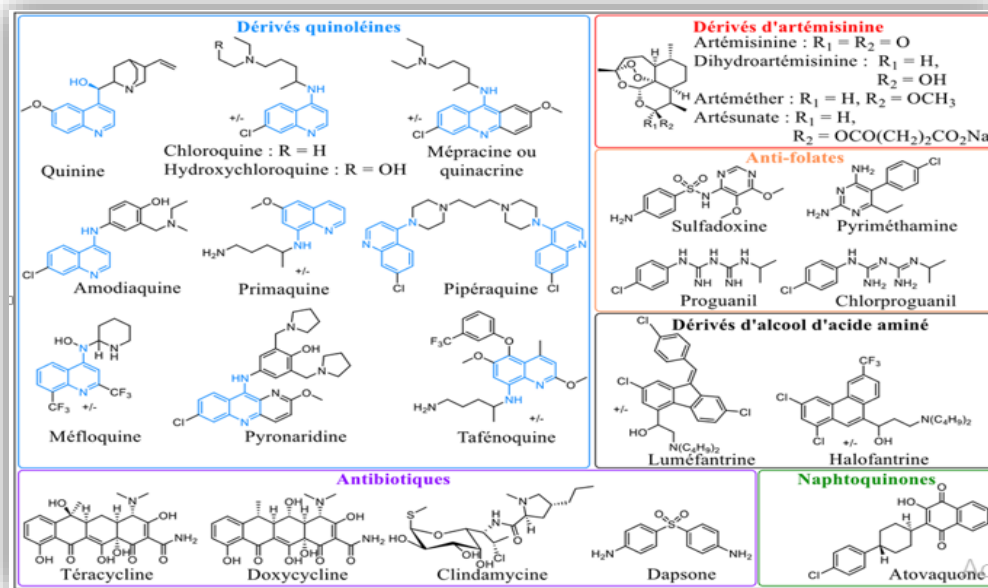


Figure 6 : Classification des médicaments antipaludiques.

**5. La résistance aux médicaments antipaludiques**

**5.1. P. falciparum chloroquine transporter ( PfCRT )**

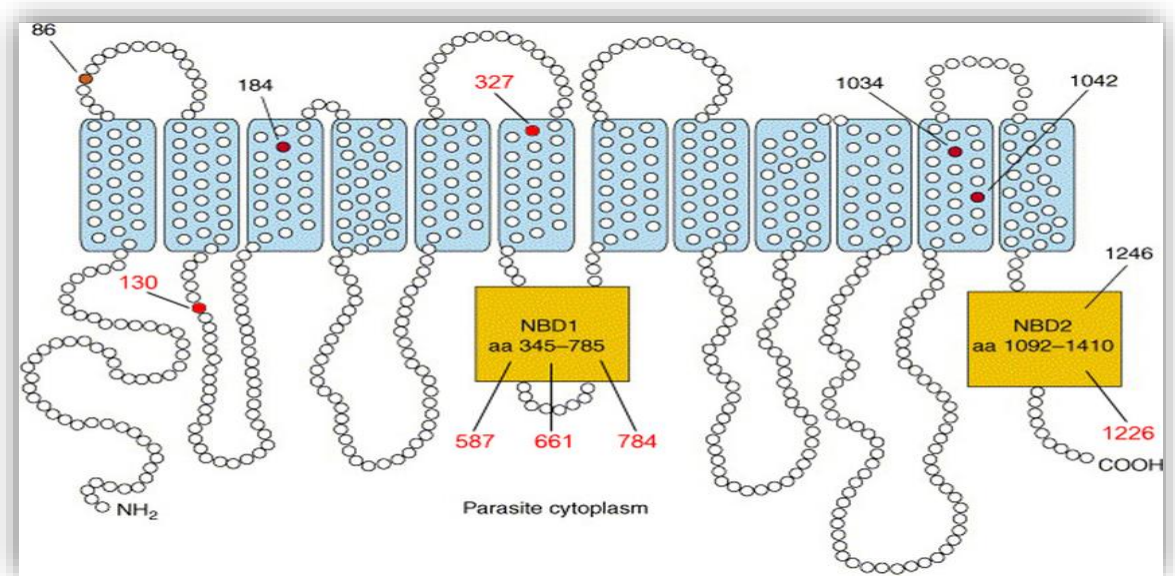
Ce gène situé sur le chromosome 7 code pour un transporteur membranaire de la vacuole digestive alimentaire impliquée dans le transport du médicament vers l'intérieur de la cellule. La mutation sur le codon 76 (K→T) [61], permet au parasite de limiter l'accumulation de chloroquine dans sa vacuole digestive (figure 8), où elle exerce son action inhibitrice [62]. PfCRT

est également impliqué dans la baisse de sensibilité du parasite à l'amodiaquine et à la quinine. [63], [64].

**5.2. P. falciparum multi-drug resistance 1 ( Pfmdr-1 )**

Situé sur le chromosome 5 [61], code pour un transporteur de type ABC ( *ATP binding cassette* ) [65]. Les mécanismes de résistance sont liés à :

- ✚ un nombre accru de copies de Pfmdr1 entraînant une augmentation de l'expression de la protéine et une résistance à la méfloquine, à la luméfántrine, à la quinine et aux artémisinines [66].
- ✚ l'apparition de mutations au niveau des codons 86 (N→Y), 184 (Y→F), 1034 (S→C), 1042 (N→D) et 1246 (D→Y) comme le montre dans la Figure 7 , entraînant une altération de sensibilité des parasites à certains antipaludiques comme les amino-4-quinoléines. Il existe un effet antagoniste entre la sensibilité à la chloroquine et à la méfloquine : la mutation N86Y diminue la sensibilité des parasites à la chloroquine, mais augmente celle de la méfloquine. De même, l'augmentation du nombre de copies du gène (86N) augmente la résistance à la méfloquine et à l'inverse accroît la sensibilité



à la chloroquine [66].

**Figure 7 :** Structure prédite et haplotypes représentatifs du transporteur multi-drug resistance.

*Chapitre 2 :*  
*La méthodologie QSAR*



## 1. Généralités sur la relation quantitative structure activité

### 1.1. Introduction à la modélisation QSAR

La modélisation de la relation quantitative structure activité (QSAR) est une approche computationnelle bien établie pour l'analyse des données chimiques. Les modèles QSAR sont développés en établissant des Relations linéaires ou non linéaires entre les valeurs des descripteurs chimiques calculées à partir de la structure moléculaire et les propriétés ou bioactivités mesurées expérimentalement de ces molécules, suivies de l'application de ces modèles pour prédire ou concevoir de nouveaux produits chimiques avec les propriétés souhaitées [67].

De nombreuses publications ont fait progresser les domaines traditionnels de la modélisation QSAR, tels que la prédiction des activités biologiques et des propriétés ADME/Tox, en s'appuyant sur l'utilisation réussie de la modélisation QSAR dans les industries chimiques, agrochimiques, pharmaceutiques et cosmétiques. Ainsi, les modèles sont devenus une partie intégrante du processus de découverte de médicaments, fournissant des conseils importants dans la planification des expériences [67].

Dans les molécules de chimioinformatique, sont représentés par des descripteurs mathématiques qui encodent des structures et des propriétés moléculaires. Des méthodes statistiques multivariées sont utilisées pour établir des relations entre les descripteurs et une propriété cible, telle que la bioactivité moléculaire. Les données sur la bioactivité chimique utilisées dans le développement de modèles QSAR, un ensemble de composés qui partagent généralement une structure centrale commune (échafaudage) et portent différents substituants (groupes R) sur un ou plusieurs sites [67].

La modélisation QSAR comporte généralement trois étapes [68] :

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

### 1.2. Aperçu historique

Il ya a plus d'un siècle, Crum-Brown et Fraser a exprimé l'idée que l'action physiologique d'une substance était en fonction de sa composition chimique [69].

Quelques décennies plus tard, en 1893, Richet a montré que la cytotoxicité de certains composés organiques était inversement proportionnelle à leur solubilité dans l'eau [70].

Au début du 20<sup>ème</sup> siècle, Meyer et Overton, ont suggé indépendamment que l'action narcotique (dépressive) d'un groupe de composés organiques correspondait à leurs coefficients de partage « huile-eau » [71].

En 1962, Hansch et Muir ont publié leur brillante étude sur la structure-activité relations entre les régulateurs de croissance des plantes et leur dépendance aux constantes de Hammett et hydrophobie [72].

$$\pi_x = \log P_x - \log P_H \quad (\text{eq 1})$$

$P_x$  et  $P_H$  représentent les coefficients de partage d'un dérivé et de la molécule mère respectivement.

L'année 1964 est considérée comme le début des méthodes QSAR modernes. Hansch et Fujita ont établi les premières corrélations entre les propriétés physico-chimiques ( $\log P$ , pKa, paramètres stériques et électroniques) et l'activité biologique (activité enzymatique, pharmacologique), Ces méthodes seront appelées par la suite l'analyse de Hansch et l'analyse de Free Wilson [73], [74].

$$\log \frac{1}{c} = a\sigma + b\pi + ck \quad (\text{eq 2})$$

$$\log \frac{1}{c} = a \cdot \log P - b(\log P)^2 + c\sigma + k \quad (\text{eq 3})$$

$$BA = \sum a_i x_i + u \quad (\text{eq 4})$$

Les limites de cette approche ont conduit à l'équation Fujita-Ban plus sophistiquée qui utilisait le logarithme de l'activité, ce qui a aligné le paramètre d'activité sur d'autres termes liés à l'énergie libre [75].

$$\log BA = \sum G_i X_i + u \quad (\text{eq 5})$$

Développement de QSAR : QSAR comparatifs (C-QSAR) [76], Hologram QSAR (HQSAR) qui sert à former l'hologramme moléculaire, ces halogrammes sont utilisées comme

descripteurs QSAR, encodant les informations chimiques et topologiques des molécules [77], Inverse QSAR [78], QSAR binaire [79].

Maintenant, des méthodes 3D comme l'étude CoMFA (Comparative Molecular Field Analysis) [80] et CoMSIA (Comparative Molecular Similarity Indices Analysis) permettent de traiter les relations structure-activité en trois dimensions, 3D-QSAR/QSPR [81].

### 1.3. Objectifs du QSAR

La plupart des méthodes QSAR se concentrent sur les objectifs suivants:

- Corréler quantitativement et récapituler les relations entre les tendances des altérations de la structure chimique et les changements respectifs des paramètres biologiques pour comprendre quelles propriétés chimiques sont les déterminants les plus probables de leurs activités biologiques.
- Optimiser les leads existants pour améliorer leurs activités biologiques.
- Prédire les activités biologiques de composés non testés et parfois encore indisponibles. [82].

## 2. La relation quantitative structure activité (2D-QSAR)

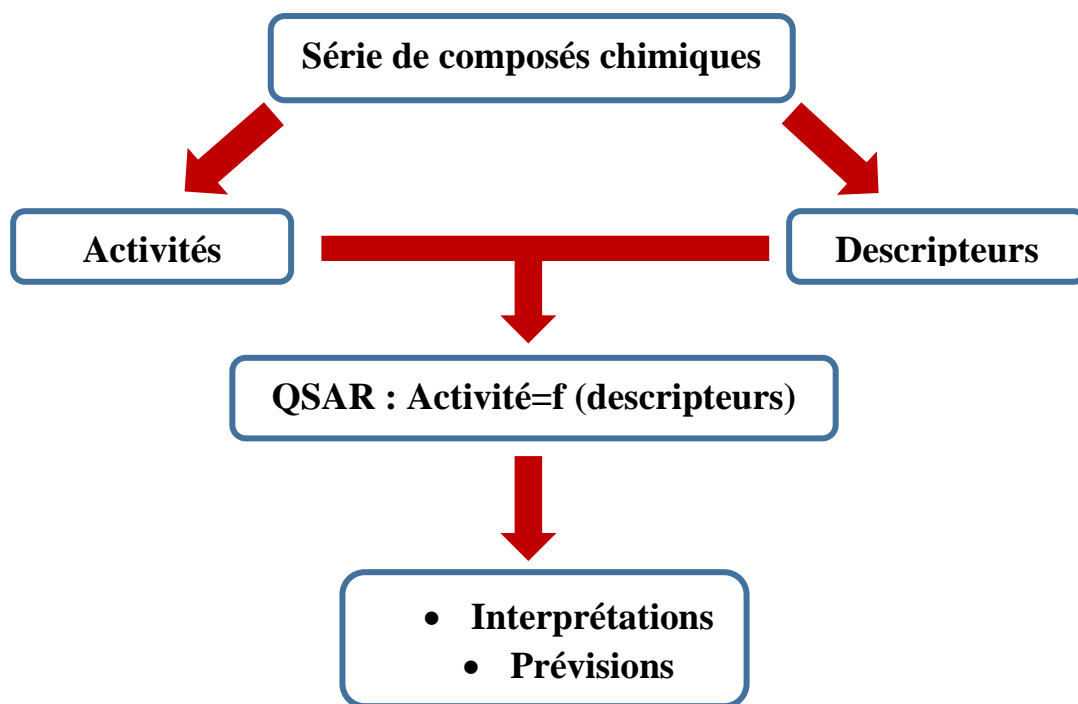
### 2.1. Principe

Le principe des méthodes de calcul 2D-QSAR est de mettre en œuvre une relation mathématique liant quantitativement des descripteurs moléculaires à une observable macroscopique (propriété physicochimique ou activité biologique) pour une série de composés chimiques similaires à l'aide de méthodes d'analyse de données statistiques. [83], cette méthode explique les changements dans les propriétés physiques et chimiques du composé causés par des changements structurels et comment l'activité biologique du composé est altérée par ces changements [84].

La forme mathématique la plus générale de 2D-QSAR est [83] :

$$[\text{Activité biologique}] = f [\text{Descripteurs}]$$

Dans l'équation, l'activité biologique est normalement exprimée comme  $\log [1/C]$ , où C est généralement la concentration minimum requise pour causer une réponse biologique définie [85], 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 [86].



**Figure 8 :** Modèle de l'étude de relation structure activité.

## 2.2. Descripteurs moléculaires

Les descripteurs sont des valeurs numériques qui encodent les caractéristiques physicochimiques des molécules à partir de leurs représentations structurales. La génération des descripteurs constitue l'étape cruciale lors de l'analyse de QSAR et la qualité d'un modèle dépendra essentiellement de la nature des descripteurs générés [87].

Le Tableau 1 présente une des manières de catégoriser les descripteurs classés en fonction de la dimension chimique utilisée pour les définir.

**Tableau 1 :** les descripteurs classés en fonction de la dimension chimique.

Dimension (D)	Niveau de description	Exemples de descripteurs
0-D	Formule chimique, propriété	Nombre d'atomes, poids moléculaire
1-D	Formule chimique développée	Dénombrement de fragments
2-D	Graphe moléculaire	Indices topologiques
3-D	Représentation 3-D	Surface moléculaire, surface polaire
4-D	Conformation des molécules	Energie potentielle des conformations

### 2.2.1. Les descripteurs 1-D

Sont accessibles à partir de la formule brute de la molécule (par exemple  $C_6H_6O$  pour le phénol), et décrivent des propriétés globales du composé. Il s'agit par exemple de sa composition, c'est-à-dire les atomes qui le constituent, ou de sa masse molaire. On peut remarquer que ces descripteurs ne permettent pas de distinguer les isomères de constitution [88]

### 2.2.2. Les descripteurs 2-D

Les descripteurs 2D sont obtenus à partir de la structure plane de la molécule. Ils sont classés en deux catégories, topologiques et constitutionnelles [89].

#### ➤ Les indices 2D constitutionnels

Qui caractérisent les différents composants de la molécule. Il s'agit par exemple du nombre de liaisons simples ou multiples, du nombre de cycles,...etc [90].

#### ➤ les indices 2D topologiques

Ce sont des quantificateurs numériques de topologie moléculaire qui sont mathématiquement dérivés de manière directe et sans ambiguïté du graphe structurel d'une molécule. Ils peuvent être sensibles à une ou plusieurs caractéristiques structurales de la molécule telles que la taille, la forme, la symétrie, la ramification et la cyclicité [91].

Les indices topologiques sont principalement basés sur les distances entre les atomes calculées par le nombre de liaisons intermédiaires et sont donc considérés comme des indices de liaison traversante [91].

### 2.2.3. Les descripteurs 3-D

Les descripteurs 3D sont évalués à partir des positions relatives des atomes dans l'espace, et décrivent des caractéristiques plus complexes. Leurs calculs nécessitent donc de connaître la géométrie 3D de la molécule [92].

#### ➤ Les descripteurs géométriques

Les plus importants sont le volume moléculaire, la surface accessible au solvant,...ect [93]. Ces descripteurs peuvent être obtenus expérimentalement ou par modélisation moléculaire, empirique. Ils sont basés sur l'arrangement spatial des atomes constituant la molécule et sont définis par les coordonnées des noyaux atomiques et la grosseur de la molécule représentée. Ces descripteurs incluent des informations sur la surface moléculaire obtenue par les aires de Van Der Waals et leur superposition [94].

➤ **Les descripteurs physicochimiques**

Certains d'entre eux reflètent la composition moléculaire du composé (le nombre et le type d'atomes et de liaisons présents dans la molécule, le nombre de cycle, les propriétés donneur/accepteur de liaison H, cation, anion, etc...). D'autres représentent le caractère hydrophile ou lipophile de la molécule généralement évalué à partir du coefficient de partage Octanol/eau représenté par le log P [95].

➤ **Les descripteurs quantiques/électroniques**

Ces descripteurs caractérisent la distribution des charges des molécules (polarité des molécules) et les paramètres de la chimie quantique : le moment dipolaire, les énergies HOMO (orbitale moléculaire occupée de plus haute énergie) et LUMO (orbitale moléculaire non occupée de plus basse énergie) [96].

### 2.3. Construction d'un modèle 2D-QSAR

Généralement dans le workflow pour la construction d'un modèle QSAR, les étapes à suivre sont comme suites (Figure 9):

1. Constituer une base de données à partir des mesures expérimentales fiables de la propriété ou de l'activité de chaque composé [97], les composés doivent appartenir à une série congénérique, doivent avoir le même mécanisme d'action et un mode de liaison identique/comparable, les données d'activité pour tous les composés doivent être dans les mêmes unités de mesure [82].
2. Sélectionner les descripteurs en relation avec la propriété ou l'activité étudiée [97].
3. Diviser cette base de données, aléatoirement, en une série d'apprentissage (*training set*) qui contient généralement les 2/3 de la base de données et une série de test (*test set*) constituée par le 1/3 restant [97].
4. Etablir des modèles mathématiques en utilisant la série d'apprentissage [97].
5. Caractériser les modèles élaborés par leurs indices de validation internes et vérifier leur robustesse par un test de hasardisation (*Randomization*) de la variable dépendante Y (réponse) [97].

✚ Validation par le test d'hasardisation (Y-randomisation test) :

Pour tester la robustesse des modèles obtenus, le test de randomisation est souvent utilisé qui permet d'affirmer que les bonnes corrélations entre les descripteurs et l'activité ne sont pas dues au hasard. Pour ce faire, les observations sont désorganisées au hasard dix fois, c'est-à-dire que la

colonne de la réponse pIC50 sera modifiée au hasard, mais les colonnes des descripteurs restent inchangées, Ces derniers doivent avoir des performances  $R^2$  et  $R_{cv}^2$  très faibles [84].

6. Valider les modèles élaborés en utilisant la série de test et calculer leurs paramètres statistiques de validation externe [97].
7. Elaborer le domaine d'applicabilité du modèle retenu.

✚ Le domaine d'applicabilité :

Le domaine d'applicabilité est la région de l'espace chimique définie par les molécules de L'ensemble d'apprentissage du modèle, permet de définir la zone dans laquelle un composé pourra être prédit avec confiance [98]. C'est un outil qui élimine les molécules d'un ensemble de tests situés en dehors de l'espace chimique de l'ensemble d'apprentissage, ce qui est requis dans les processus de validation mis en œuvre au niveau de l'OCDE [99].

Il existe plusieurs méthodes pour la détermination du domaine d'applicabilité d'un modèle QSAR, parmi ces méthodes on trouve la méthode de leviers [84].

Le domaine d'applicabilité sera discuté à l'aide du diagramme de Williams qui représente les résidus de prédiction standardisés en fonction des valeurs des leviers. [99]

8. Explorer et exploiter les modèles validés pour comprendre les mécanismes et les modes d'action [97].

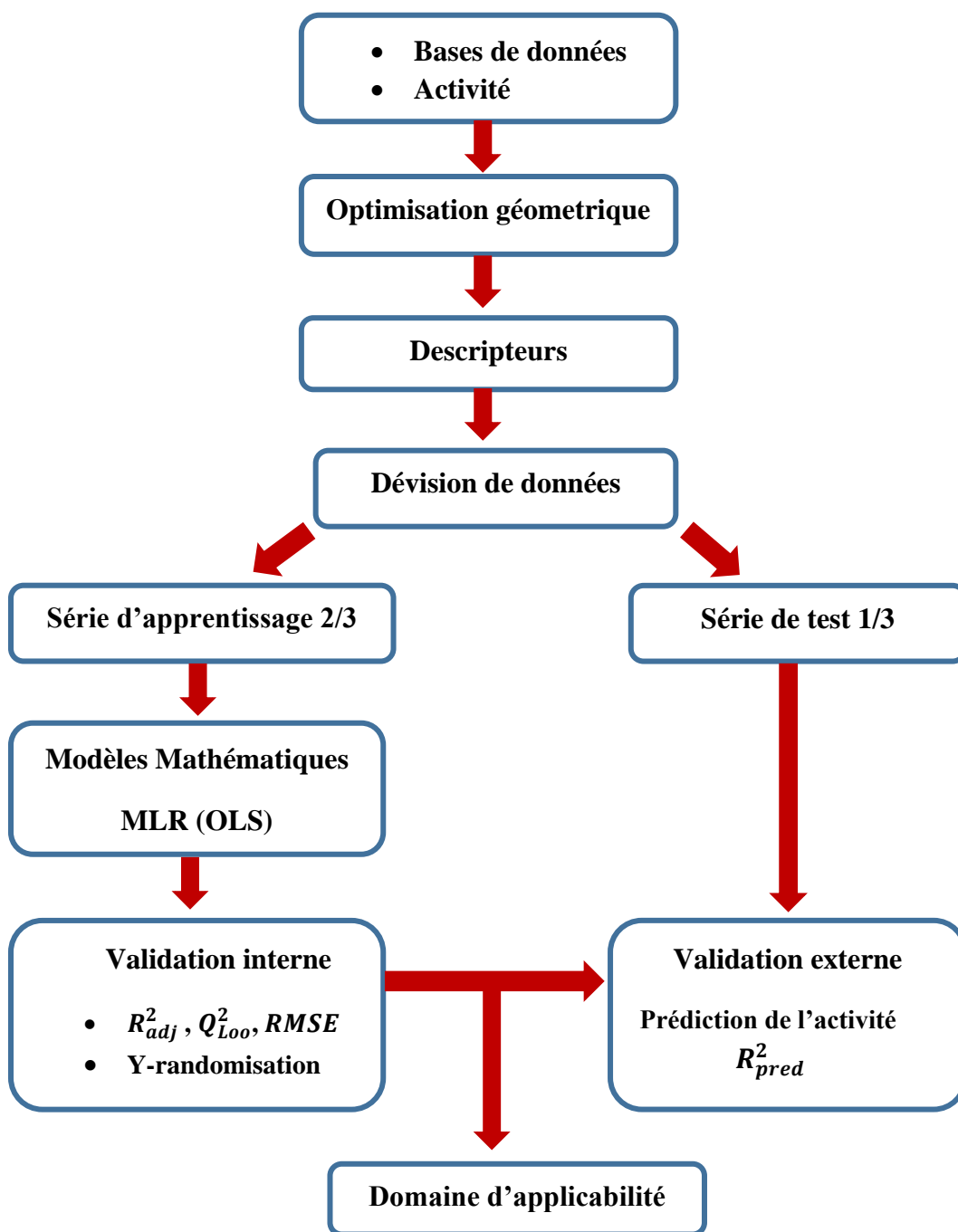


Figure 9 : Construction d'un modèle 2D-QSAR



### 3. La relation quantitative structure activité (3D-QSAR)

Lorsqu'il est question de la QSAR en 3D, les chimistes computationnels partent généralement du principe que l'analyse QSAR prend en compte une structure tridimensionnelle du composé dans sa conformation d'énergie minimale et construit un modèle QSAR basé sur divers champs tridimensionnels générés [100], [101].

#### 3.1. Définition

La relation quantitative structure-activité tridimensionnelle (3D-QSAR) est la méthode la plus couramment utilisée dans la conception de médicaments, qui comprend l'analyse comparative du champ moléculaire (CoMFA) et l'analyse comparative des indices de similarité moléculaire (CoMSIA) [102].

Des modèles 3D-QSAR peuvent être mise en œuvre lorsque la structure 3D de la cible biologique est résolue. Les approches 3D-QSAR ont été développées pour corrélérer l'activité biologique d'une série de composés actifs de référence avec l'arrangement spatial de nombreuses propriétés de la molécule telles que les propriétés stériques, lipophiliques et électronique et pour fournir des indications pour l'optimisation par pharmacomodulation et la conception de nouveaux composés avec des profils d'activité améliorés [103].

- **CoMFA (Comparative Molecular Field Analysis)**

CoMFA est une méthode qui corrèle les valeurs des champs de structure avec les activités biologiques. CoMFA génère une équation corrélant l'activité biologique avec la contribution des champs d'énergie d'interaction à chaque point de grille. La technique conventionnelle CoMFA a été utilisée pour calculer les propriétés stériques et électrostatiques sur la base des monopôles centrés sur l'atome, des potentiels de Coulomb et de Lennard-Jones. Cette méthode a été développée en 1988 et reste l'une des plus populaires pour la modélisation 3D QSAR [104].

- **CoMSIA (Comparative Molecular Similarity Indices Analysis)**

Le modèle CoMSIA fournit des informations sur les champs stériques, électrostatiques, hydrophobes, donneurs de liaisons hydrogène et accepteurs de liaisons hydrogène, CoMSIA capable de donner des modèles plus stables que CoMFA dans les études 3D-QSAR [103].

Lors de la génération des modèles CoMFA et CoMSIA, les molécules alignées ont été importées dans un réseau cubique 3D généré automatiquement avec un pas de grille de 2 Å. Dans les études CoMFA, les champs stériques et électrostatiques ont été pris en compte [105].

### 3.2. Principe

Pour la méthode 3D-QSAR, l'objectif est d'établir la relation entre les descripteurs de champ moléculaire 3D (champ stérique calculé à l'aide du potentiel de Lennard Jones et champs électrostatiques évalués par le potentiel de Coulomb) et l'efficacité de conversion de puissance.

La Figure 10 montre le processus de modélisation 3D-QSAR corrélant l'espace des descripteurs de champ moléculaire (c'est-à-dire les vecteurs de  $S_1, S_2, S_3, S_4 \dots$ ) et l'efficacité de conversion de puissance (PCE). Les symboles de  $LV_1, LV_2, \dots, LV_j$  sont des variables latentes mutuellement indépendantes, qui n'ont pas de signification physique définie. Les paramètres de  $a_j$  (ie,  $a_1, a_2 \dots a_j$ ) et  $b_{j,n}$  (c'est-à-dire,  $b_{1,n}, b_{2,n} \dots b_{j,n}$ ) sont des coefficients dans le modèle QSAR 3D résultant. Il convient de noter que ce sont les variables [44].

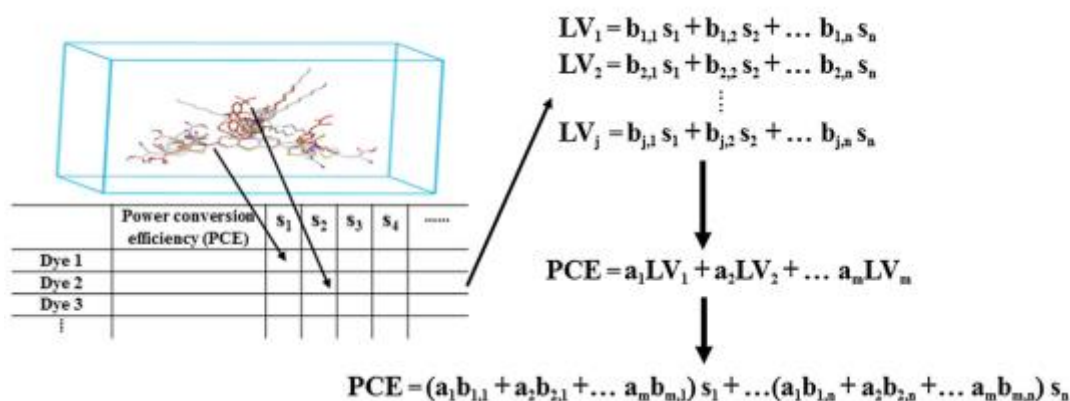
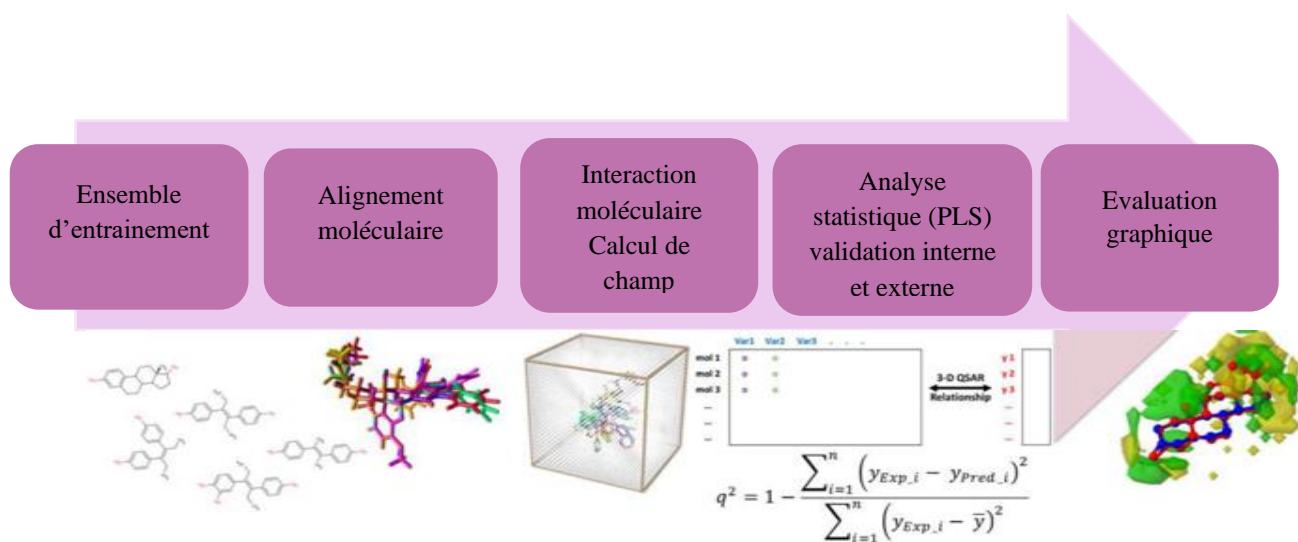


Figure 10 : le processus de modélisation 3D-QSAR.

### 3.3. Étapes pour la procédure 3D QSAR

- 1. Préparation des données :** Collecte des données expérimentales sur l'activité biologique des composés d'intérêt. Les données doivent être complètes et de haute qualité.
- 2. Prétraitement des structures :** Génération des structures 3D des composés à l'aide d'un logiciel de modélisation moléculaire. Les structures doivent être optimisées pour obtenir des conformations stables.
- 3. Alignement moléculaire :** Alignement des structures moléculaires des composés sur une référence commune. Cela permet de superposer les structures et de les aligner pour faciliter la comparaison et l'analyse.
- 4. Génération des champs moléculaires :** Calcul des champs moléculaires tels que les champs stériques, électrostatiques, hydrophobes, etc., pour chaque composé à l'aide de méthodes appropriées.

5. **Sélection des descripteurs** : Identification et sélection des descripteurs moléculaires pertinents qui représentent les propriétés structurales et physico-chimiques importantes des composés.
6. **Construction du modèle QSAR** : Utilisation d'une méthode statistique telle que la régression linéaire, la régression partielle des moindres carrés (PLS), etc., pour construire un modèle QSAR en corrélant les descripteurs moléculaires aux activités biologiques mesurées.
7. **Validation du modèle** : Évaluation de la qualité et de la validité du modèle QSAR en utilisant des méthodes de validation croisée telles que la validation croisée leave-one-out (LOO) et la validation croisée externe. Cela permet de vérifier la capacité prédictive du modèle.
8. **Interprétation des résultats** : Analyse des relations structure-activité mises en évidence par le modèle QSAR. Cela permet de comprendre les caractéristiques structurales importantes qui influencent l'activité biologique des composés.
9. **Utilisation du modèle QSAR** : Application du modèle QSAR pour prédire l'activité biologique de nouveaux composés qui n'ont pas été inclus dans l'ensemble d'apprentissage. Cela peut être utilisé pour guider la conception et l'optimisation de nouveaux composés ayant une activité souhaitée.
10. **Validation externe** : Validation du modèle QSAR sur des ensembles de test externes indépendants pour évaluer sa robustesse et sa généralité.



**Figure 11** : Workflow Field-Based d'un modèle 3D-QSAR.

Ces étapes constituent une procédure générale pour la réalisation d'une analyse 3D-QSAR avec la méthode FB (Field-Based). Cependant, il est important de noter que les détails spécifiques de la méthode peuvent varier en fonction du logiciel et des approches utilisés.

### 3.4. Les méthodes d'alignement moléculaires

L'un des problèmes les plus cruciaux dans la plupart des méthodes 3D-QSAR basées sur l'alignement est que leurs résultats sont très sensibles à la manière dont les conformations bioactives de toutes les molécules se superposent l'autre. Dans les cas où toutes les molécules d'un ensemble de données ont une structure centrale rigide commune, les molécules peuvent être alignées facilement en utilisant la procédure d'ajustement des moindres carrés. Cependant, en cas d'hétérogénéité structurelle dans l'ensemble de données, l'alignement de molécules hautement flexibles devient assez difficile et chronophage. Plusieurs approches ont été proposées pour superposer les molécules le plus fidèlement possible, certaines dont sont les suivants :

- **Superposition basée sur le chevauchement d'atomes**

Cette méthode implique un appariement atome à atome correspondant entre les molécules. On l'appelle aussi l'approche pharmacophore et est la méthode la plus populaire, car elle donne la meilleure correspondance des positions des atomes présélectionnés. Il est bénéfique dans identifier la dissemblance entre des molécules similaires, mais ne peut pas être appliqué à des molécules de types structurels différents où les atomes correspondants sont difficiles à sélectionner.

- **Superposition basée sur les sites de liaison**

Dans cette méthode moléculaire est obtenu en superposant les sites actifs des récepteurs ou les résidus des récepteurs qui interagissent avec le ligands. Cette approche est considérée comme plus concevable, malgré les problèmes d'analyse conformationnelle dus à l'amélioration degrés de liberté.

- **Superposition basée sur les champs/pseudo-champs**

Cette méthode effectuer une superposition en comparant les similitudes dans les champs d'énergie d'interaction calculés entre les molécules. Similitude électrostatique et similitude de surface moléculaire des indices ont également été utilisés par les chercheurs pour alignement.

- **Superposition à base de pharmacophores :**

Cette méthode utilise un pharmacophore hypothétique comme une commune utile modèle cible. Chaque molécule est conformationnellement dirigée prendre la forme obligatoire pour ses caractéristiques sous-moléculaires correspondre soit à un pharmacophore connu, soit à celui qui est généré lors de l'analyse conformationnelle.

- **Superposition basée sur plusieurs conformères :**

Cette méthode est particulièrement utile dans les cas où les ligands peuvent se lier à un récepteur de multiples façons, ou lorsque la bonne liaison mode est inconnu et les ligands ont un bon degré de flexibilité conformationnelle. Par exemple, le 3D-QSAR méthode COMPASS (décrite dans une section ultérieure) de manière itérative détermine et sélectionne la meilleure conformation bioactive et alignement optimal à partir d'un ensemble de poses initiale [82].

#### 4. Méthodes statistiques des modèles

L'objectif de l'analyse statistique est justement de « démêler » ces descripteurs et d'identifier ceux qui sont corrélés à la variable cible, qui produisent du signal, de ceux qui ne le sont pas, qui produisent du bruit. L'analyse statistique permet également d'identifier les descripteurs qui sont corrélés entre eux pour ne garder que les principaux et réduire ainsi la redondance d'informations [99].

L'analyse statistique détermine et quantifie les corrélations entre les descripteurs et la variable cible. Elle indique également la contribution relative de chaque descripteur dans l'explication globale de l'activité. Le modèle statistique est une équation donnant la valeur de la variable cible en fonction de la somme des valeurs pondérées des descripteurs [99].

Les principaux outils statistiques pour obtenir un modèle sont :

- ❖ la régression linéaire multi variée (*Multivariate Linear Regression* - MLR).
- ❖ la régression des moindres carrés partiels (*Partial Least Squares* - PLS).

##### 4.1. La régression linéaire multiple

La régression linéaire multiple (MLR) est la plus simple méthode statistique de modélisation et la plus appliquée dans les études de la relation structure-activité. La méthode a été popularisée par *Hansch* en reliant l'activité biologique aux propriétés expérimentales lipophiliques, électroniques et stériques pour des séries de composés [106].

La MLR utilisée pour modéliser une relation linéaire entre une variable dépendante Y (pIC<sub>50</sub>) et des Variables indépendantes X (descripteurs moléculaires).

La MLR est basée sur la technique des moindres carrés (ordinary least squares) : le modèle est ajusté de sorte que la somme des carrés entre les différentes valeurs réelles et prédites soit minimisée. MLR estime les coefficients de régression ( $R^2$ ) en appliquant l'ajustement des moindres carrés [107].

L'expression de régression prend la forme :

$$Y = a_0 + \sum_{i=1}^n a_i x_i \quad (\text{eq 6})$$

Avec :

- $Y$  : la variable dépendante (pIC50) ;
- $x_i$  : les variables indépendantes (descripteur) ;
- $n$  : le nombre de variables ;
- $a_0$  : la constante de l'équation du modèle ;
- $a_i$  : les coefficients de descripteurs dans l'équation du modèle.

#### 4.2. La régression des moindres partiels

La méthode des moindres carrés partiels (PLS), une extension de l'analyse de régression multiple, a été appliquée pour construire la relation entre la structure et les activités biologiques. [108], la qualité de l'analyse PLS est évaluée par deux paramètres clés : Le coefficient de corrélation de validation croisée ( $Q^2$ ) et le coefficient de détermination ( $R^2$ ) [109].

Est une procédure 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 [20]. PLS donne une solution statistiquement robuste même lorsque les variables indépendantes sont étroitement liées entre elles, ou lorsque les variables indépendantes dépassent le nombre d'observations. Ainsi, PLS est capable d'analyser des données complexes de structure-activité de manière plus réaliste et d'interpréter efficacement l'influence de la structure moléculaire sur l'activité biologique [16].

La méthode PLS a été utilisée pour la génération du modèle 3D-QSAR, qui associait également les descripteurs CoMFA et CoMSIA à l'activité biologique [110].

#### 4.3. Paramètres d'évaluation des modèles

##### ➤ Coefficient de détermination $R^2$

C'est le coefficient de corrélation de Bravais Pearson entre  $y_i$  et  $\hat{y}_i$ , c'est à dire entre valeurs observées et prédites par le modèle de régression, il est noté  $R$ , sa valeur varie entre 0 et 1 [51], il est défini mathématiquement comme :

$$R = \sqrt{1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2}} \quad (\text{eq 7})$$

➤ **Coefficient de détermination ajusté  $R_{adj}^2$**

C'est une modification du coefficient de corrélation de Pearson qui affine le nombre de descripteurs utilisés dans le modèle de régression multilinéaire qui sera toujours inférieur ou égal au coefficient de corrélation de Pearson tel que défini ci-dessous :

$$R_{Adjusted}^2 = 1 - (1 - r^2) \frac{n-1}{n-p-1} \quad (\text{eq 8})$$

où N correspond au nombre de composés dans l'ensemble d'apprentissage sous forme de points de données et p est le nombre de descripteurs dans le modèle construit [98].

➤ **Erreur quadratique moyenne (RMSE)**

Une bonne mesure pour évaluer les performances de prédiction du modèle généré qui est proportionnelle au score moyen observé [111] tel que défini ci-dessous.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_{pred} - y_{obs})^2}{n}} \quad (\text{eq 9})$$

➤ **l'erreur type résiduel « s »**

L'écart type (s) est un autre paramètre habituellement rapporté; il indique dans quelle mesure la fonction de régression prédit les données observées [112], ce paramètre est donné par :

$$s = \sqrt{\frac{\sum_{i=1}^N (y_{pred} - y_{obs})^2}{n-p-1}} \quad (\text{eq 10})$$

Dont p est le nombre de variables indépendantes.

➤ **Le Test de Fisher-Snedecor (F)**

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

$$F = \frac{ESS/p}{RSS/(n-p-1)} \quad (\text{eq 11})$$

Où n, et (p-1) sont des degrés de liberté associés à ESS et RSS respectivement [113].

#### 4.4. Interprétation et validation des modèles

La validation du modèle est une manière essentielle de modéliser la relation quantitative structure activité (QSAR). Ceci est fait pour tester la stabilité interne et la capacité prédictive des modèles QSAR [107].

➤ **Validation interne**

La validation interne a utilisé la validation croisée leave One-out (LOO). Cela se fait en excluant le(s) point(s) des données de l'ensemble d'apprentissage, puis en construisant le modèle basé sur les activités de données restantes et enfin, en utilisant ce modèle pour tester les données exclues. Ce processus a été répété jusqu'à ce que les activités de l'ensemble d'apprentissage aient été prédites. Le coefficient de validation croisée  $R_{cv}^2$  ou  $Q_{LOO}^2$  a été calculé pour l'ensemble d'apprentissage, c'est une mesure du pouvoir prédictif du modèle de régression [114], il est défini comme :

$$Q_{LOO}^2 = 1 - \frac{\sum_{i=1}^N (y_{pred,i} - y_{obs,i})^2}{\sum_{i=1}^N (y_{obs,i} - \bar{y}_{obs,i})^2} = 1 - \frac{PRESS}{TSS} \quad (\text{eq 12})$$

où  $y_{obs,i}$  et  $y_{pred,i}$  font référence aux scores d'activité observés et prédits. Plus le score de  $Q_{LOO}^2$  est proche de 1 plus le pouvoir prédictif du modèle est mieux.

➤ **Validation externe**

La fiabilité et les performances prédictives des modèles ont également été évaluées à l'aide des paramètres de validation externes pertinents [98]. Cette méthode consiste à prédire la propriété/activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modèle. Une fois l'ensemble de validation mis en place, il suffit alors d'appliquer le modèle QSAR aux molécules qui le composent et de déterminer la corrélation existant entre les activités calculées et expérimentales. Plus cette corrélation est importante, plus le modèle est capable de prédire les activités pour des molécules hors l'ensemble d'apprentissage, cette validation est caractérisée par le paramètre  $R_{pred}^2$  [107], il est défini mathématiquement comme suit :

$$R_{pred}^2 = \frac{\sum (y_{pred(test)} - y_{test})^2}{\sum (y_{train} - y_{mean(train)})^2} \quad (\text{eq 13})$$

$y_{pred(test)}$  et  $y_{test}$  indiquent les valeurs d'activité prédites et observées pour l'ensemble des test et  $y_{train}$  indique la valeur d'activité moyenne de l'ensemble d'apprentissage. Pour le modèle prédictif QSAR, la valeur de  $R_{pred}^2$  doit être supérieure à 0,5 [54].



*Partie II : Etude  
Expérimentale*

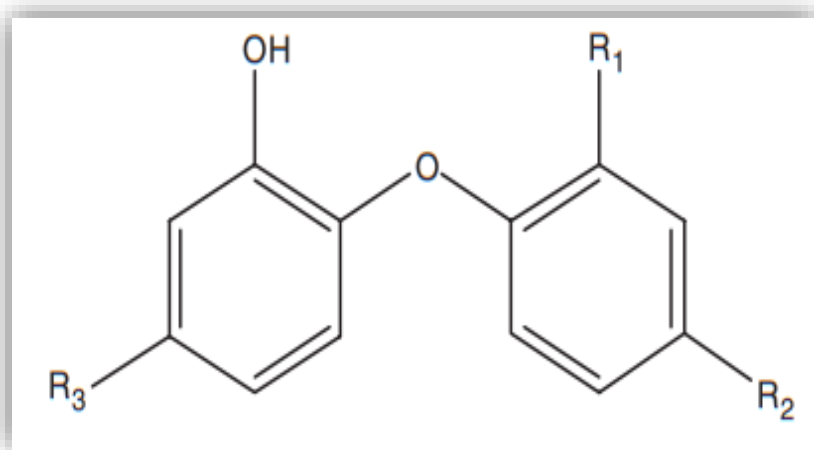
# *Matériels et méthodes*

## 1. Matériels et méthodes

### 1.1. Ensemble de données expérimentales

Pour étudier les relations quantitatives structure-activité (QSAR), une série de dérivés de triclosan (Annexe 01), inhibiteurs de la réductase d'acyl enoyl porteur de *Plasmodium falciparum* a été utilisée pour développer des modèles 2D-QSAR et 3D-QSAR, les structures et l'activité biologique expérimentale de l'IC<sub>50</sub> de ces composés ont été converties en pIC<sub>50</sub> (-log IC<sub>50</sub>) et utilisée comme variable dépendante dans l'analyse CoMFA et CoMSIA.

Les résultats biologiques sont exploités par la méthode de régression linéaire multiple (MLR = multiple linear regression) dans le but d'établir un modèle QSAR.



**Figure 12 :** Structure centrale commune (échafaudage) du triclosan.

### 1.2. Etude 2D-QSAR

#### 1.2.1. Traitements de données

Le développement et l'étude des Relations Quantitatives Structure-Activité (2D-QSAR) passe par trois étapes :

- a. Dessin des molécules :** Pour le dessin des structures des molécules on a utilisé le logiciel « ChemDraw » .
- b. Optimisation :** tous les calculs d'optimisation des structures des molécules ont été effectués en utilisant le logiciel HyperChem.
- c. Modélisation :** la modélisation passe par deux étapes :
  - ✓ Calcul de descripteurs : Les géométries ainsi optimisées ont été transférées dans le logiciel Dragon pour le calcul de plus descripteurs.

- ✓ Conception du modèle : Dans le présent travail, la sélection du modèle a été réalisée par algorithme génétique, en utilisant le logiciel MOBY DIGS [115].

### **ChemDraw**

Est un outil complet destiné aux chimistes et biologistes, intégrant toute une gamme d'outils intelligents permettant de faciliter les travaux des chercheurs, il est développé en 1985 par David A. Evans et Stewart Rubenstein. C'est un outil essentiel et préféré pour illustrer les concepts chimiques et biologiques et il est imposé comme une référence pour le dessin de structures moléculaires. Puisqu'il est simple à utiliser, puissant et permet de dessiner de manière intuitive et efficace en deux et trois dimensions [116], [117].

### **HyperChem**

Est un environnement de modélisation moléculaire sophistiqué, connu pour sa qualité, sa flexibilité et sa facilité d'utilisation. Il propose différentes méthodes d'optimisation (PM3, MM+, AM1, etc...) En optimisant la molécule, les descripteurs moléculaires peuvent être calculés telle que le volume moléculaire, les énergies HUMO et LUMO...etc. Les descripteurs moléculaires théoriques ont été calculés par le processus suivant : les structures moléculaires ont été pré-optimisées par le champ de force de la mécanique moléculaire MM+ issu de ce logiciel de modélisation moléculaire [118], [119].

### **DRAGON**

Est une application de calcul de descripteurs moléculaires, ces descripteurs peuvent être utilisés pour évaluer les molécules et la relations structure-activité ou structure -propriété, ainsi que pour l'analyse de similarité et criblage à haut débit des bases de données moléculaires. Les versions de DRAGON ont évolué avec le temps afin de faire avancer la recherche dans le QSAR. Sachant que le DRAGON n'a pas été conçu comme un logiciel QSAR, il ne fournit que des descripteurs moléculaires et ne peut pas effectuer une analyse QSAR (au contraire le MobyDigs). Cependant, DRAGON permet la fusion des descripteurs moléculaires calculés et des propriétés définies par l'utilisateur pour un ensemble de molécules, fournissant un fichier de sortie complet qui est facilement chargé [120].

### **MobyDigs**

C'est un logiciel qui permet de calculer un model QSAR avec des régressions linéaires multiples [115].

### 1.2.2. Répartition de l'échantillon

Pour la sélection des composés de calibrage et de validation, nous avons utilisé le choix aléatoire est la technique la plus simple. elle suppose qu'un ensemble d'échantillons tirés aléatoirement d'un lot suffisamment grand suit la distribution statistique du lot entier [121].

### 1.3. Etude 3D-QSAR

La présente étude met en avant l'importance de l'analyse 3D-QSAR, en se concentrant sur différents aspects tels que l'ensemble de données utilisé, l'optimisation et l'alignement des structures, ainsi que les études CoMFA et CoMISA. De plus, une analyse des moindres carrés partiels (PLS) a été réalisée pour approfondir l'analyse des données.

A l'aide du logiciel SYBYL-X 2.1.1, la conception, l'optimisation moléculaire et l'alignement ont été mis en œuvre pour le 3D-QSAR [122]. Ce programme utilise l'optimisation géométrique des Tripos ForceField [123], [124].

#### 1.3.1. Optimisation et alignements moléculaires

Toutes les structures 3D ont été construites à l'aide du logiciel Sybyl 2.1.1 sketcher (TRIPOS Associates Inc.). Les géométries des structures 3D ont été optimisées en utilisant la méthode tripos MMFF94 avec la méthode de Powel. Le nombre maximal d'itérations pour la minimisation a été fixé à  $10 \times 10^3$ , et la convergence de l'énergie a été atteinte lorsque le critère de gradient d'énergie de 0,05 Kcal/mole a été atteint, permettant ainsi d'obtenir des conformations stables. Les charges électrostatiques ont été attribuées à l'aide de la méthode de Gasteiger-Hucker, qui a ensuite été utilisée pour les études 3D-QSAR [125].

L'alignement moléculaire est une étape cruciale dans l'étude 3D-QSAR, car la qualité et la capacité prédictive du modèle dépendent directement de cet alignement. Tous les composés ont été alignés en utilisant la méthode d'alignement de la base de données. Comme aucune information sur la nature des complexes ligand-récepteur n'est fournie, le composé 52 avec  $IC_{50} = 0.038$ , qui était l'un des composés les plus actifs contre *Plasmodium falciparum*, a été utilisé comme modèle pour aligner les autres composés de l'ensemble de données sur lui, en se basant sur la sous-structure commune à deux cycles illustrée dans la Figure [126].

### 1.3.2. Étude CoMFA et CoMSIA

Tous les composés alignés ont été placés dans une boîte en treillis avec un espacement de grille de 2 Å, généré automatiquement par le logiciel. CoMFA repose sur le principe selon lequel les variations d'activité biologique sont corrélées aux variations des composants stériques et électrostatiques des composés. Les champs CoMFA ont été générés en utilisant un atome de carbone hybridé sp<sup>3</sup> comme sonde stérique (rayon de van der Waals de 1,53 Å) et une sonde électrostatique (charge positive simple) à chaque point de grille. La valeur de coupure a été fixée à 30 kcal/mol. Les descripteurs d'indices de similarité CoMSIA, qui incluent les composantes stériques, électrostatiques, hydrophobes, accepteurs d'hydrogène et donneurs d'hydrogène, ont été calculés en utilisant la même boîte en treillis utilisée dans les calculs CoMFA [127].

### 1.3.3. Analyse des moindres carrés partiels (PLS)

La méthode des moindres carrés partiels (PLS) a été utilisée pour établir une corrélation linéaire entre les descripteurs CoMFA et CoMSIA en tant que variables indépendantes et les valeurs de pIC<sub>50</sub> en tant que variables dépendantes.

Le nombre optimal de composantes a été déterminé en utilisant la procédure Leave-One-Out (LOO). Le modèle final (analyse non validée en interne) a été développé pour produire le coefficient de corrélation R<sup>2</sup> non validé en utilisant le nombre optimal de composantes dérivées du modèle présentant le plus haut coefficient de validation croisée (Q<sup>2</sup>) [128].

***Résultats et  
discussions***

## 2. Résultat et discussions

### 2.1. Construction d'un modèle 2D-QSAR

La construction du modèle QSAR a été déterminée en prenant pour variable dépendante le logarithme de l'inverse de  $IC_{50}$  (en  $\mu mol / litre$ ) pour les inhibiteurs de la protéine porteuse énoyl acyl réductase chez *Plasmodium falciparum* (PfENR) et pour les variables explicatives on a choisis entre 4 modèles :

- ❖ Modèle un descripteur L2u ;
- ❖ Modèle deux descripteurs R6m , R8v ;
- ❖ Modèle trois descripteurs RDF010e , Mor03e , P2u ;
- ❖ Modèle quatre descripteurs DECC , BEHe4 , L2u , C-026.

Tableau regroupe les définitions de différents descripteurs.

**Tableau 2:** Les descripteurs intervenant dans les modèles.

Descripteurs	Classe	Signification
<b>L2u</b>	WHIM descripteur	Indice WHIM directionnel de la taille du 2 <sup>ème</sup> composant.
<b>R6m</b>	GETAWAY descripteur	R autocorrélation du décalage 6/ pondéré par les masses atomiques.
<b>R8v</b>	GETAWAY descripteur	R autocorrélation du décalage 8/ pondéré par les volumes atomiques de van der Waals.
<b>Mor03e</b>	3D-MoRSE descripteur	Signal 3D-MoRSE 03/ pondéré par les électronégativités atomiques de Sanderson.
<b>P2u</b>	WHIM descripteur	Indice WHIM directionnel de la forme de la 2 <sup>ème</sup> composante / non pondéré.



<b>DECC</b>	descripteurs topologiques	Excentrique.
<b>BEHe4</b>	valeurs propres de Burden.	valeur propre la plus élevée n. 4 de la matrice de charge / pondérée par les électronégativités atomiques de Sanderson .
<b>C-026</b>	Fragments centrés sur l'atome.	R—CX—R.

Ces 61 composés ont été divisés en un ensemble d'apprentissage (49 composés) et un ensemble de test (12 composés) en utilisant la randomisation ainsi que la diversité chimique et biologique.

### 2.1.1. Analyse de la régression

L'activité inhibitrice ENR représenté par le cologarithme décimal d'IC<sub>50</sub> (pIC<sub>50</sub>) de 61 dérivés de Triclosan. Parmi les modèles sélectionnés par l'algorithme génétique nous avons retenu les modèles dont les équations sont :

#### 2.1.1.1. Modèle pour 1 Descripteur L2u

##### Équation du modèle :

$$\text{Log (1/IC50)} = 1.77 (\pm 0.22) - 0.59 (\pm 0.066) \text{ L2u} \quad (\text{eq 14})$$

Le meilleur modèle pour 1 dimension pour 2D QSAR a été obtenu avec L2u indice WHIM directionnel de la taille du 2<sup>ème</sup> composant. les paramètres statistiques sont regroupé dans le tableau 4.

**Tableau 3 :** Paramètres statistiques du modèle 1.

ID	MODEL	R <sup>2</sup>	Q <sup>2</sup> <sub>Lo0</sub>	Q <sup>2</sup> <sub>LMO</sub> 10%	Q <sup>2</sup> <sub>LMO</sub> 20%	Q <sup>2</sup> <sub>LMO</sub> 30%	Q <sup>2</sup> <sub>ext</sub>	Q <sup>2</sup> <sub>F1</sub>	Q <sup>2</sup> <sub>F2</sub>	Q <sup>2</sup> <sub>F3</sub>	KΔ	F	S
1	L2u	62.26	59.75	50.94	49.47	49.46	81.38	76.38	76.88	76.88	73.15	77.55	0.582

**2.1.1.2. Modèle pour 2 Descripteurs R6m R8v**

**Équation du modèle :**

$$\text{Log (1/IC}_{50}) = -0.45 (\pm 0.36) + 4.18 (\pm 0.67816) \mathbf{R6m} - 7.05604 (\pm 0.77428) \mathbf{R8v} \quad (\text{eq 15})$$

Le meilleur modèle pour 2 dimensions a été obtenu grâce à l'utilisation de l'algorithme génétique avec R6m qui représente R autocorrélation du décalage 6/ pondéré par les masses atomiques et R8v qui montre R autocorrélation du décalage 8/ pondéré par les volumes atomiques de van der Waals. Les paramètres statistiques sont regroupés dans le tableau 5.

**Tableau 4 : Paramètre statistiques du modèle 2.**

ID	MODEL	R <sup>2</sup>	Q <sup>2</sup> <sub>LOO</sub>	Q <sup>2</sup> <sub>LMO</sub> 10%	Q <sup>2</sup> <sub>LMO</sub> 20%	Q <sup>2</sup> <sub>LMO</sub> 30%	Q <sup>2</sup> <sub>ext</sub>	Q <sup>2</sup> <sub>F1</sub>	Q <sup>2</sup> <sub>F2</sub>	Q <sup>2</sup> <sub>F3</sub>	KΔ	F	S
2	<b>R6m R8v</b>	63.85	59.25	58.69	58.39	58.44	56.82	54.67	47.95	66.88	11.85	59.58	0.474

**2.1.1.3. Model pour 3 Descripteurs RDF010e Mor03e P2u**

**Équation du modèle :**

$$\text{Log (1/IC}_{50}) = 3.162 (\pm 0.29656) - 0.56068 (\pm 0.6556) \mathbf{RDF010e} - 0.32423 (\pm 0.05732) \mathbf{Mor03e} - 6.47124 (\pm 0.93977) \mathbf{P2u} \quad (\text{eq 16})$$

Le meilleur modèle pour 3 dimensions a été obtenu grâce à utilisation de l'algorithme génétique avec RDF010e qui représente Fonction de distribution radiale -1.0/ Pondérée par les électronégativités atomiques de Sanderson et Mor03e qui montre Signal 3D-MoRSE 03/ pondéré par les électronégativités atomiques de Sanderson et P2u qui représente Indice WHIM directionnel de la forme de la 2<sup>ème</sup> composante / non pondéré. Les paramètres statistiques sont regroupés dans le tableau 6.

**Tableau 5 : Paramètres statistiques du modèle 3.**

ID	MODEL	R <sup>2</sup>	Q <sup>2</sup> <sub>LOO</sub>	Q <sup>2</sup> <sub>LMO</sub> 10%	Q <sup>2</sup> <sub>LMO</sub> 20%	Q <sup>2</sup> <sub>LMO</sub> 30%	Q <sup>2</sup> <sub>ext</sub>	Q <sup>2</sup> <sub>F1</sub>	Q <sup>2</sup> <sub>F2</sub>	Q <sup>2</sup> <sub>F3</sub>	KΔ	F	S
3	<b>RDF010e Mor03e P2u</b>	68.77	64.04	65.22	63.23	62.70	57.45	92.95	72.43	72.43	14.74	52.37	0.443

**2.1.1.4. Modèle pour 4 Descripteurs DECC BEHe4 L2u C-026**

**Équation du modèle :**

$$\text{Log}(1/IC_{50}) = 6.97378 (\pm 1.29271) + 1.81407 (\pm 0.32011) \text{ DECC} - 3.57247 (\pm 0.59627) \text{ BEHe4} - 0.22143 (\pm 0.08146) \text{ L2u} + 0.386 (\pm 0.1059) \text{ C-026} \quad (\text{eq 17})$$

Le meilleur modèle pour 4 dimensions a été obtenu grâce à l'utilisation de l'algorithme génétique avec DECC qui représente un excentrique, BEHe4 qui représente valeur propre la plus élevée n. 4 de la matrice de charge / pondérée par les électronégativités atomiques de Sanderson, L2u qui montre Indice WHIM directionnel de la taille du 2<sup>ème</sup> composant et C-026 qui représente R-CX-R. Les paramètres statistiques sont regroupés dans le tableau 7.

**Tableau 6 : Paramètres statistiques du modèle 4.**

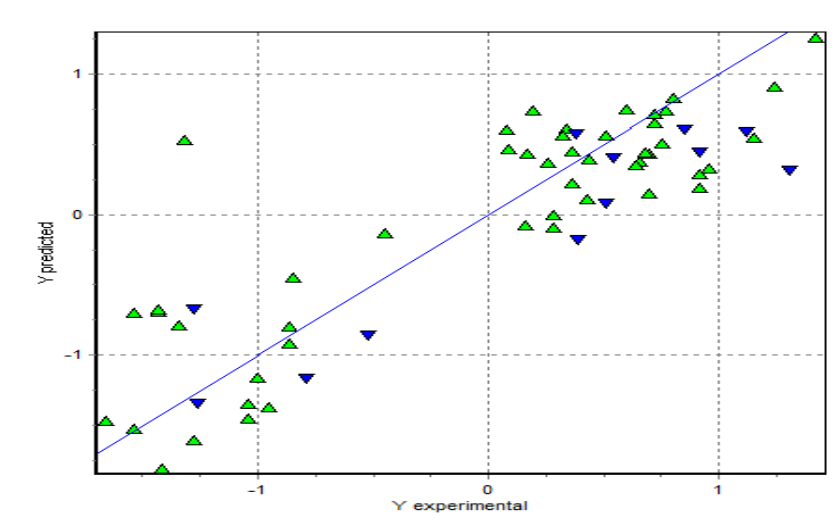
ID	MODEL	R <sup>2</sup>	Q <sup>2</sup> <sub>LOO</sub>	Q <sup>2</sup> <sub>LMO</sub> 10%	Q <sup>2</sup> <sub>LMO</sub> 20%	Q <sup>2</sup> <sub>LMO</sub> 30%	Q <sup>2</sup> <sub>ext</sub>	Q <sup>2</sup> <sub>F1</sub>	Q <sup>2</sup> <sub>F2</sub>	Q <sup>2</sup> <sub>F3</sub>	KΔ	F	S
4	DECC BEHe4 L2u C-026	77.55	72.45	72.42	71.72	71.24	81.03	77.17	77.17	73.79	14.51	39.2	0.444

D'après les résultats sélectionné (les 4 tableaux) on a choisie celui de 4 descripteurs (R<sup>2</sup> =77.55 %) (Q<sup>2</sup><sub>LOO</sub> =72.45 %) (Q<sup>2</sup><sub>ext</sub>=81.03 %).

**2.1.2. Qualité de l'ajustement**

Les valeurs des paramètres statistiques (Tableau 9) prennent en considération la corrélation entre les quatre descripteurs (DECC, BEHe4, L2u, C-026) des 61 dérivés de triclosan, les valeurs de (Q<sup>2</sup><sub>ext</sub> =81.03%) et de (Q<sup>2</sup><sub>LOO</sub>=72.45%) montrent la bonne qualité de l'ajustement et la valeur du coefficient de détermination R<sup>2</sup> signifie que 77.55% de la variabilité de pIC<sub>50</sub> est expliquée par ces descripteurs.

le tableau 7 et la figure 14, représentent la bonne qualité d'ajustement, la robustesse et le pouvoir prédictif de notre modèle (R<sup>2</sup>, Q<sup>2</sup><sub>LOO</sub>, Q<sup>2</sup><sub>ext</sub>), les valeurs des erreurs relativement faibles, et celles de la dispersion des points autour de la droite, Montrent que les valeurs prédites sont en adéquation avec les valeurs expérimentales, caractéristique d'un bon ajustement.



**Figure 13 :** Droite d'ajustement pour le modèle 4 des valeurs expérimentales et prédits de la notation  $pIC_{50}$  pour le modèle de QSAR.

### 2.1.3. Validation

Dans cette étude l'ensemble des données a été divisé en deux sous-ensemble : sous-ensemble d'entraînement (80 % = 49 molécules) pour la construction du model et la validation interne, et un sous-ensemble de Test (20 % = 12 molécules) pour la validation externe.

Le défi actuel de QSAR/QSPR n'est plus au développement d'un modèle statistiquement solide, c'est l'élaboration d'un modèle capable de prédire avec précision l'activité ou la propriété de nouveaux produits chimiques.

#### ➤ Validation interne

Le coefficient de détermination de la validation croisée du modèle  $Q_{LOO}^2 = 72.45\%$ , témoigne d'une bonne corrélation entre l'activité prédite et l'activité réelle, et reflète une précision du modèle.

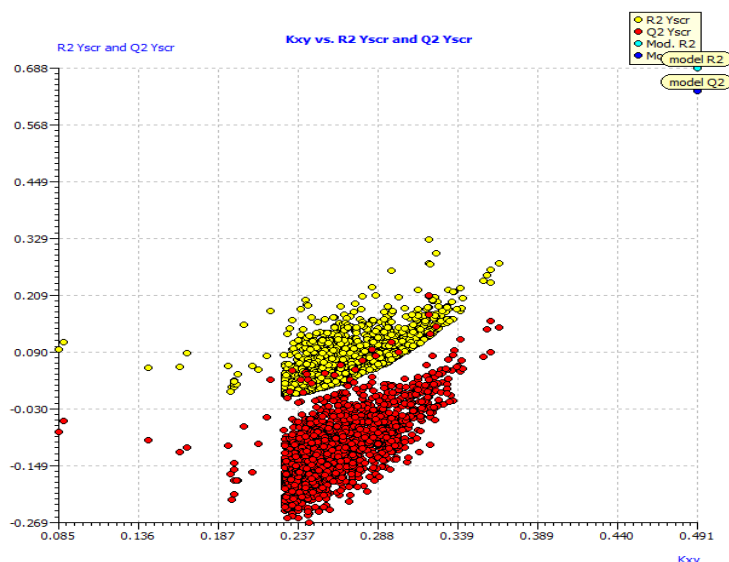
#### ➤ Validation externe

La validation interne, pour les nouveaux composés ne suffit pas pour juger la capacité prédictive du modèle développé. La validation externe doit être réalisée pour une bonne prévision des composés qui n'ont pas été utilisé lors de l'élaboration du modèle.

Les paramètres de la validation externe  $Q_{ext}^2 = 81.03\%$  confirme la bonne capacité prédictive du modèle pour les composés non impliqués dans les calculs.

## ➤ test de randomisation (Y scrambling)

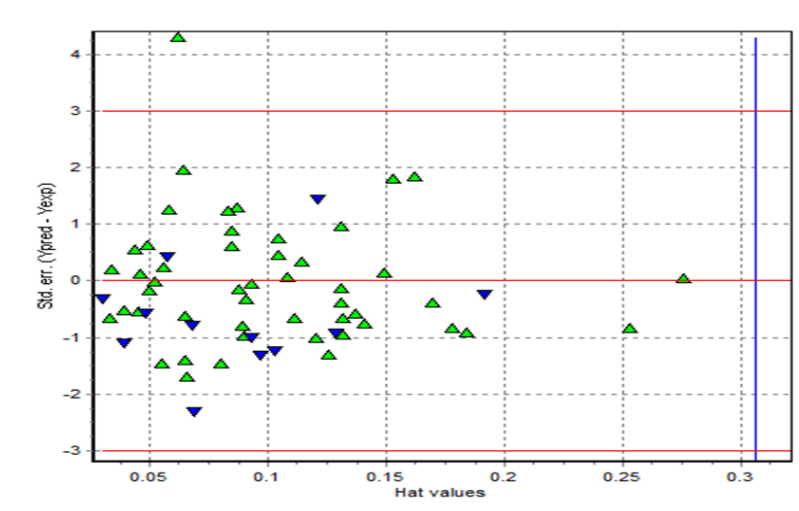
Les 61 modèles ont été développés, et la figure 15 montre les faibles valeurs de  $Q^2$  et  $R^2$  obtenues après chaque mélange, faisant ressortir que les résultats du modèle original ne sont pas aléatoires.



**Figure 14 :** test de randomisation.

### 2.1.4. Domaine d'applicabilité

Le diagramme de Williams désigne les composés qui se trouvent dans le domaine d'application de notre model (Figure 15), la plupart des composés de l'ensemble de données sont dans l'intervalle  $\pm 3$  du domaine d'application, aucune molécule n'a dépassé le seuil  $h^*=0,31$  (points levier). Ainsi, le modèle proposé pourrait être utilisé pour examiner les bases de données existantes.



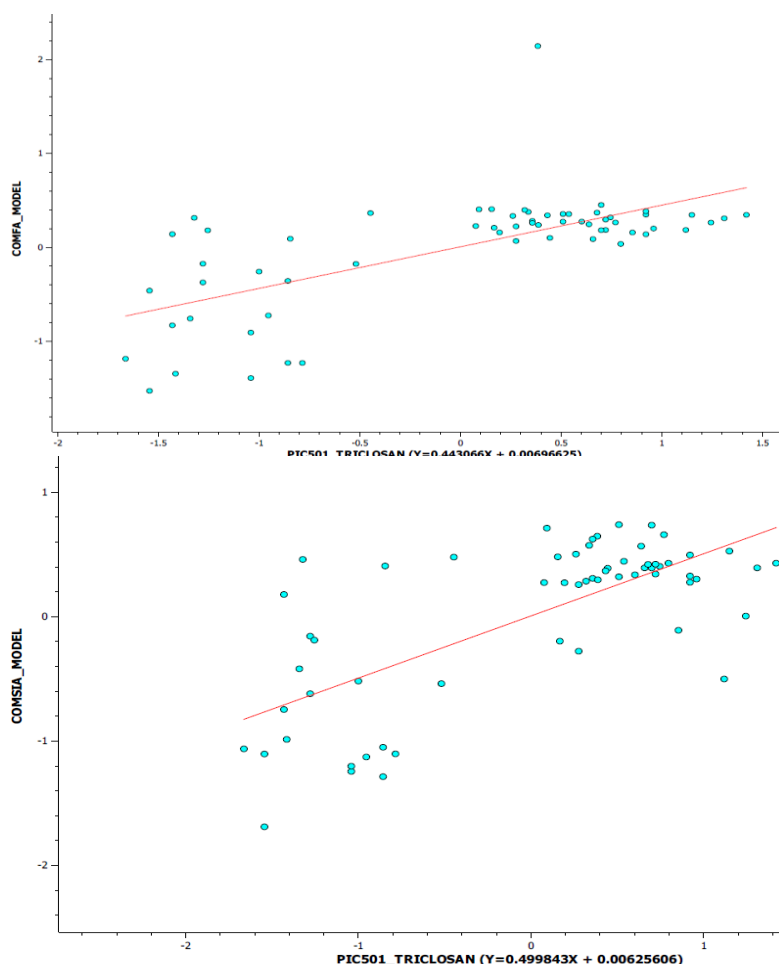
**Figure 15 :** Diagramme de Williams.

## 2.1.5. Conclusion

Nous avons développé une équation utile de QSAR qui relie des descripteurs théoriques à l'IC<sub>50</sub> expérimental de 61 molécules. Les valeurs très élevées des paramètres statistiques ( $Q_{LOO}^2$ ,  $Q_{LMO}^2$ ,  $R^2$ ,  $Q_{ext}^2$ ), montrent que le modèle proposé est robuste et possède une bonne capacité prédictive. On peut alors conclure que pour cette famille de composés, le modèle proposé pourrait prédire et donner des informations sur l'inhibition de la protéine porteuse énoyl acyl réductase chez *Plasmodium falciparum* (PfENR).

## 2.2. Construction d'un modèle 3D-QSAR

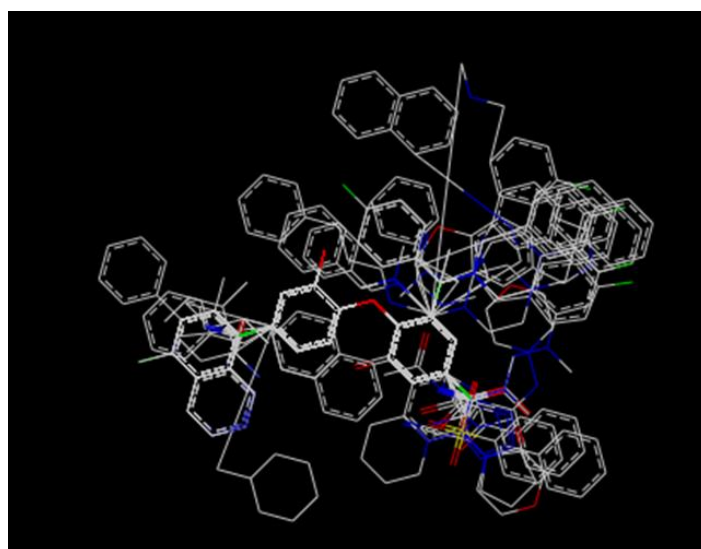
Les valeurs expérimentales et prédites de pIC<sub>50</sub> pour l'ensemble d'entraînement des modèles CoMFA et CoMSIA sont présentées dans le Tableau 7 et le graphique de pIC<sub>50</sub> expérimental par rapport à pIC<sub>50</sub> prédit pour l'ensemble d'entraînement des modèles CoMFA et CoMSIA est illustré dans la Figure 16.



**Figure 16 :** Valeurs prédites de pIC<sub>50</sub> et valeurs expérimentales de pIC<sub>50</sub> pour les 61 molécules d'ensemble entraînement obtenues par analyse PLS en utilisant les modèles CoMFA et CoMSIA.

## RESULTATS ET DISCUSSIONS

Les résultats des moindres carrés partiels (PLS) des modèles CoMFA et CoMSIA sont répertoriés dans le Tableau 7. Le coefficient de corrélation en validation croisée ( $Q^2$ ) et le coefficient de corrélation hors validation croisée ( $R^2$ ) sont considérés comme deux paramètres clés pour évaluer la qualité de l'analyse PLS, indiquant respectivement la capacité prédictive et la cohérence interne. Golbraikh et Tropsha ont rapporté que les modèles sont considérés acceptables s'ils satisfont toutes les conditions suivantes :  $Q^2 > 0,5$  et  $R^2 > 0,6$ . Les statistiques PLS des modèles CoMFA et CoMSIA indiquent que le modèle CoMFA est légèrement meilleur que le modèle CoMSIA.



**Figure 17 :** Composés du jeu de données alignés sur le triclosan.

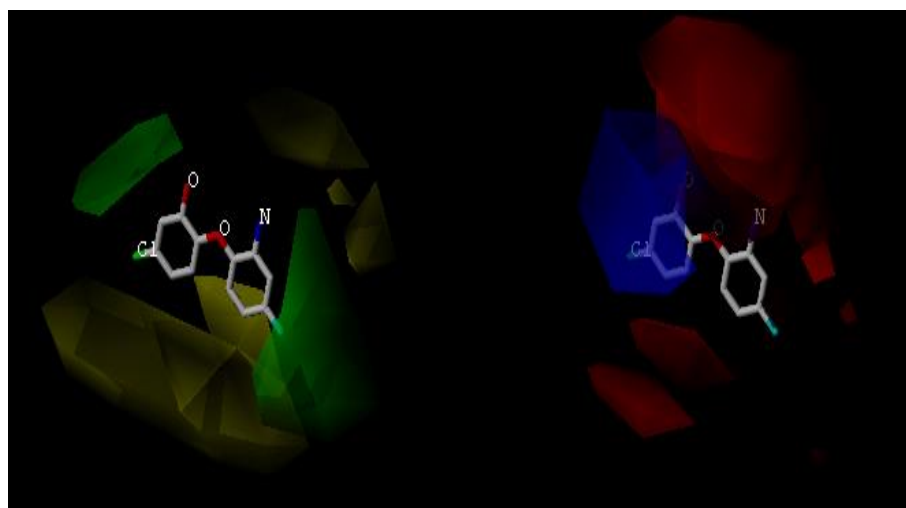
Les modèles CoMFA et CoMSIA ont donné un coefficient de corrélation acceptable en validation croisée ( $Q^2$ ) de 64.33% et 53.51% respectivement, avec un nombre de composantes de 3, ce qui est raisonnable compte tenu du nombre de composés utilisés pour générer les modèles, suggérant ainsi une bonne capacité prédictive. Des coefficients de corrélation hors validation croisée ( $R^2$ ) élevés de 94.40% pour CoMFA et 95.57% pour CoMSIA, avec une faible estimation de l'erreur standard .

**Tableau 7 :** Les résultats pour le modèles PLS de CoMFA et CoMSIA.

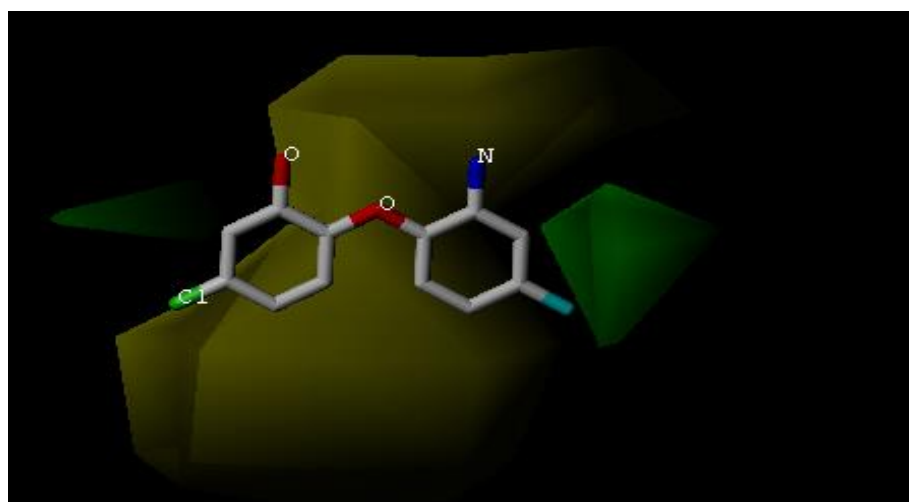
Modèle	$Q^2\%$	$R^2\%$	Standard Error
CoMFA	64.33	94.40	0.68
CoMSIA	53.51	95.57	0.65

## RESULTATS ET DISCUSSIONS

Les cartes de contours 3D-QSAR des modèles CoMFA et CoMSIA ont été représentées dans la Figure 18 en utilisant le composé 52 comme structure de référence. Le champ stérique est représenté par des cartes de contours verts et jaunes, où les contours verts représentent les régions favorables pour les groupes stériquement encombrants, tandis que les contours jaunes représentent des régions stériquement moins favorables. Le champ électrostatique est représenté en rouge et en bleu, ce qui montre les régions favorables pour les groupes riches en électrons et les régions favorables pour les groupes déficients en électrons, respectivement.



(A)



(B)

**Figure 18 :** Cartes de contours de (A) CoMFA et (B) CoMSIA en combinaison avec le composé 52.



### 2.2.1. Conclusion

Il est important de comprendre qu'il existe généralement un nombre raisonnablement élevé de conformations à basse énergie et que la sélection de la conformation ayant le score le plus élevé ne garantit pas qu'elle soit pertinente. Afin de fournir des modèles significatifs, CoMFA et CoMSIA dépendent de la connaissance de la conformation bioactive des ligands et de leur alignement structural relatif.

***Conclusion  
générale***

### CONCLUSION GENERALE

La relation entre la structure moléculaire et l'activité biologique est l'un des défis majeurs de la conception des médicaments, dont le développement de nouveaux composés antipaludique.

Notre but consiste à développer des modèles QSAR fiables sur la base d'une série de 61 dérivés Triclosan pour prédire l'activité inhibitrice de la protéine ENR.

Dans la première partie, l'utilisation des descripteurs du logiciel DRAGON-Mobydigs et aboutissant à la construction des modèles QSAR par la sélection d'approche algorithme génétique, et dont les résultats obtenus ainsi que les paramètres statistiques des modèles doivent être acceptables.

Pour ce faire, le logiciel DRAGON-Mobydigs a été utilisé ainsi qu'une analyse de régression linéaire multiple (MLR) nous permettant d'aboutir à une corrélation entre  $pIC_{50}$  et quatre descripteurs sélectionnés : (DECC, BEHe4, L2U, C-026). Dont l'objectif est de former l'équation finale de QSAR.

Les principales techniques de validation (interne, externe, la randomisation des Y...) ont été utilisées pour évaluer le modèle QSAR obtenu en accord avec les cinq principes de l'OCDE ; nous avons également vérifié les critères d'acceptabilité du modèle de Golbarikh et Tropsha.

Le model QSAR a été déterminée en prenant pour variable :

- Dépendante : le logarithme de l'inverse de  $IC_{50}$  (en mmol / litre).
- Explicatives : trois descripteurs (DECC, BEHe4, L2U, C-026).

Selon les conditions de Golbreich et Tropsha :

Le modèle trouvé :

Nous avons réalisé aussi une analyse de 3D-QSAR (CoMFA et CoMSIA) pour un ensemble de 61 molécules de triclosan. La fiabilité des modèles a été évaluée à l'aide des paramètres statistiques obtenus dans les modèles CoMFA et CoMSIA. Les modèles optimaux obtenus ont été utilisés pour prédire l'activité biologique et fournir des informations structurales clés. Les cartes de contours générées à partir de nos modèles permettent une interprétation claire et visuelle de la relation entre les structures de dérivés de triclosan et leurs activité. La plupart des résultats expérimentaux sont cohérents avec nos valeurs prévues, ce qui indique que nos

## CONCLUSION GENERALE

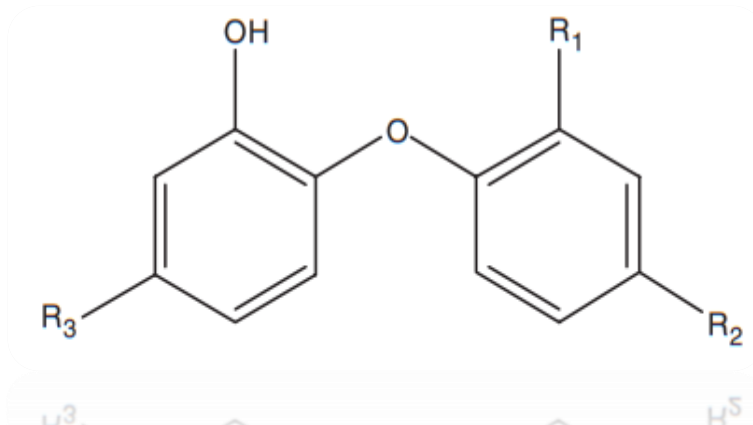
---

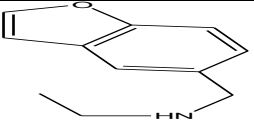
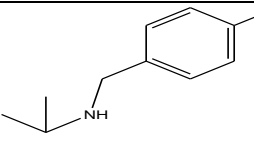
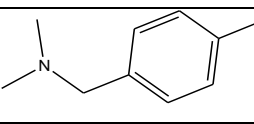
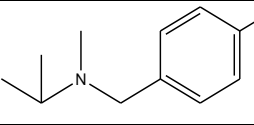
modèles sont fiables. Ces résultats très encourageants revêtent une importance cruciale pour comprendre la relation entre la structure et l'activité des dérivés de triclosan, et peuvent aider à développer des stratégies sous-jacentes pour le criblage virtuel et la conception de nouvelles molécules.

Notre objectif principal a ouvert des perspectives très prometteuses dans ce domaine, et nous envisageons de continuer à développer nos modèles en utilisant d'autres méthodes telles que le Machine Learning, le Deep Learning et les machines à graphes. Nous sommes déterminés à poursuivre notre parcours de recherche et à explorer de nouvelles approches pour améliorer encore nos résultats.

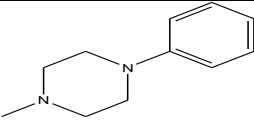
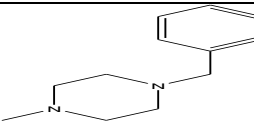
# *Annexes*

**Annexe 01 :** Valeurs expérimentales prédit et calculé par l'approche DRAGON-Mobydigs pour les 61 dérivés de Triclosan.



	$R_1$	$R_2$	$R_3$	choix	$y_{exp}$	$y_{pred}$
1	$NH_2$	$H$	$Cl$	Entrainement	-0.85	-0.45
2	$CH_2NHCH_2Ph$	$H$	$Cl$	Entrainement	-1.00	-1.16
3	$CH_2NHCH_2(p - ClC_6H_4)$	$H$	$Cl$	Entrainement	-1.54	-0.70
4	$CH_2NH(CH_2)_3Ph$	$H$	$Cl$	Entrainement	-1.28	-1.61
5	$CH_2NH(1 - Naphtyl)$	$H$	$Cl$	Entrainement	-1.04	-1.45
6	$CH_2NH(2 - Naphtyl)$	$H$	$Cl$	Entrainement	-1.66	-1.47
7	$CH_2NHCH_2Ph$	$Cl$	$Cl$	Test	-0.52	-0.86
8	$CH_2NHC(O)Ph$	$Cl$	$Cl$	Test	-1.26	-1.35
9	$CH_2NHCH_2(p - ClC_6H_4)$	$Cl$	$Cl$	Entrainement	-0.86	-0.80
10	$CH_2NHCH_2(p - PhC_6H_4)$	$Cl$	$Cl$	Entrainement	-1.54	-1.53
11	$CH_2NH(2 - Naphtyl)$	$Cl$	$Cl$	Entrainement	-1.42	-1.81
12		$Cl$	$Cl$	Entrainement	-0.95	-1.37
13		$Cl$	$Cl$	Entrainement	-0.86	-0.92
14		$Cl$	$Cl$	Entrainement	-1.34	-0.79
15		$Cl$	$Cl$	Entrainement	-1.43	-0.69
16	$CH_2NHMe$	$Cl$	$Cl$	Entrainement	-1.43	-0.68
17	$CH_2NMe_2$	$Cl$	$Cl$	Test	-1.28	-0.68

## ANNEXES

18		<i>Cl</i>	<i>Cl</i>	Test	-0.79	-1.17
19		<i>Cl</i>	<i>Cl</i>	Entrainement	-1.04	-1.35
20	<i>Cl</i>	<i>OH</i>	<i>Cl</i>	Entrainement	0.66	0.38
21	<i>Cl</i>	<i>NH<sub>2</sub></i>	<i>Cl</i>	Entrainement	0.44	0.39
22	<i>Cl</i>	<i>CN</i>	<i>Cl</i>	Entrainement	0.92	0.19
23	<i>Cl</i>	<i>CH<sub>2</sub>CN</i>	<i>Cl</i>	Entrainement	0.72	0.65
24	<i>Cl</i>	<i>COOH</i>	<i>Cl</i>	Entrainement	0.26	0.37
25	<i>Cl</i>	<i>1H - tetrazol - 5 - yl</i>	<i>Cl</i>	Test	0.38	0.57
26	<i>Cl</i>	<i>C(O)NH<sub>2</sub></i>	<i>Cl</i>	Entrainement	0.92	0.29
27	<i>Cl</i>	<i>C(O)N(H)Me</i>	<i>Cl</i>	Entrainement	0.51	0.56
28	<i>Cl</i>	<i>C(O)N(H)Bn</i>	<i>Cl</i>	Test	0.51	0.08
29	<i>Cl</i>	<i>C(O)NMe<sub>2</sub></i>	<i>Cl</i>	Entrainement	0.34	0.61
30	<i>Cl</i>	<i>C(O)N - Pyrrolidine</i>	<i>Cl</i>	Entrainement	0.09	0.47
31	<i>Cl</i>	<i>C(O)N - Piperidine</i>	<i>Cl</i>	Entrainement	0.17	0.43
32	<i>Cl</i>	<i>C(O)N - Morpholine</i>	<i>Cl</i>	Entrainement	0.70	0.43
33	<i>Cl</i>	<i>N(H)Ac</i>	<i>Cl</i>	Entrainement	1.24	0.91
34	<i>Cl</i>	<i>N(H)Bz</i>	<i>Cl</i>	Entrainement	0.77	0.74
35	<i>Cl</i>	<i>N(H)C(O)N<sub>2</sub></i>	<i>Cl</i>	Entrainement	0.80	0.83
36	<i>Cl</i>	<i>N(H)C(O)N - Morpholine</i>	<i>Cl</i>	Entrainement	0.60	0.75
37	<i>Cl</i>	<i>B(H)SO<sub>2</sub>Ph</i>	<i>Cl</i>	Entrainement	0.43	0.11
38	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>3</sub></i>	Entrainement	0.70	0.15
39	<i>Cl</i>	<i>Cl</i>	<i>Ph</i>	Test	0.85	0.60
40	<i>Cl</i>	<i>Cl</i>	<i>CN</i>	Test	1.31	0.31
41	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>CH<sub>3</sub></i>	Entrainement	0.96	0.33
42	<i>Cl</i>	<i>Cl</i>	<i>(CH<sub>2</sub>)<sub>2</sub>CH<sub>3</sub></i>	Entrainement	0.68	0.44
43	<i>Cl</i>	<i>Cl</i>	<i>(CH<sub>2</sub>)<sub>3</sub>CH<sub>3</sub></i>	Entrainement	0.32	0.56
44	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>CH(CH<sub>3</sub>)<sub>2</sub></i>	Entrainement	0.75	0.51
45	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>CH(CH<sub>3</sub>)<sub>2</sub></i>	Entrainement	0.54	0.40
46	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>CH(CH<sub>3</sub>)CH<sub>2</sub>Cl</i>	Test	0.92	0.44
47	<i>Cl</i>	<i>Cl</i>	<i>(CH<sub>2</sub>)<sub>2</sub>CH(CH<sub>3</sub>)<sub>2</sub></i>	Entrainement	0.28	-0.09

## ANNEXES

48	<i>Cl</i>	<i>Cl</i>	<i>Cyclohexylmethy</i>	Entrainement	0.36	0.45
49	<i>Cl</i>	<i>CN</i>	<i>o - CH<sub>3</sub> - Ph</i>	Test	0.39	-0.18
50	<i>Cl</i>	<i>Cl</i>	<i>o - CH<sub>3</sub> - Ph</i>	Entrainement	0.64	0.35
51	<i>Cl</i>	<i>Cl</i>	<i>m - CH<sub>3</sub> - Ph</i>	Entrainement	0.72	0.72
52	<i>Cl</i>	<i>Cl</i>	<i>p - CH<sub>3</sub> - Ph</i>	Entrainement	1.42	1.26
53	<i>Cl</i>	<i>Cl</i>	<i>p - F - Ph</i>	Entrainement	1.15	0.55
54	<i>Cl</i>	<i>Cl</i>	<i>(CH<sub>2</sub>)<sub>2</sub>Ph</i>	Test	0.12	0.59
55	<i>Cl</i>	<i>Cl</i>	<i>(CH<sub>2</sub>)<sub>3</sub>Ph</i>	Entrainement	0.36	0.22
56	<i>Cl</i>	<i>CN</i>	<i>2 - Pyridyl</i>	Entrainement	0.16	-0.08
57	<i>Cl</i>	<i>CN</i>	<i>4 - Pyridyl</i>	Entrainement	-0.45	-0.13
58	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>(2 - Pyridyl)</i>	Entrainement	0.19	0.74
59	<i>Cl</i>	<i>Cl</i>	<i>CH<sub>2</sub>(3 - Pyridyl)</i>	Entrainement	0.08	0.60
60	<i>Cl</i>	<i>CN</i>	<i>CH<sub>2</sub>(4 - Pyridyl)</i>	Entrainement	0.28	0.00
61	<i>Cl</i>	<i>Cl</i>	<i>C(O)NH<sub>2</sub></i>	Entrainement	-1.32	0.53



*Liste des références  
bibliographiques*

### Références bibliographiques

- [1] J. Sainy and R. Sharma, “QSAR analysis of thiolactone derivatives using HQSAR, CoMFA and CoMSIA,” *SAR QSAR Environ. Res.*, vol. 26, no. 10, pp. 873–892, 2015, doi: 10.1080/1062936X.2015.1095238.
- [2] Danishuddin, G. Madhukar, M. Z. Malik, and N. Subbarao, “Development and rigorous validation of antimalarial predictive models using machine learning approaches,” *SAR QSAR Environ. Res.*, vol. 30, no. 8, pp. 543–560, 2019, doi: 10.1080/1062936X.2019.1635526.
- [3] P. Shah and M. I. Siddiqi, “3D-QSAR studies on triclosan derivatives as Plasmodium falciparum enoyl acyl carrier reductase inhibitors,” *SAR QSAR Environ. Res.*, no. December 2012, pp. 37–41, 2010, doi: 10.1080/1062936X.2010.502297
- [4] S. Egieyeh, J. Syce, S. F. Malan, and A. C. Id, “Predictive classifier models built from natural products with antimalarial bioactivity using machine learning approach,” *PLoS One*, pp. 1–15, 2018.
- [5] S. Lindert and J. A. Mccammon, “Dynamics of Plasmodium falciparum enoyl-ACP reductase and implications on drug discovery,” *PROTEIN Soc.*, vol. 21, pp. 1734–1745, 2012, doi: 10.1002/pro.2155.
- [6] M. T. Ibrahim and A. Uzairu, “2D-QSAR, molecular docking, drug-likeness, and ADMET/ pharmacokinetic predictions of some non-small cell lung cancer therapeutic agents,” *J. Taibah Univ. Med. Sci.*, vol. 18, no. 2, pp. 295–309, 2023, doi: 10.1016/j.jtumed.2022.09.002.
- [7] Mme. ERRAHOUI née BELLIFA KHADIDJA, T H E S E Pour l’obtention du diplôme de Doctorat en Chimie Physique “Etude des relations quantitatives structure–toxicité des composés chimiques à l’aide des descripteurs moléculaires“. « Modélisation QSAR », UNIVERSITE ABOU BEKR BELKAÏD DE TLEMCEM, 2015.
- [8] M. Hamadache, O. Benkortbi, S. Hanini, A. Amrane, L. Khaouane, and C. Si Moussa, “A Quantitative Structure Activity Relationship for acute oral toxicity of pesticides on rats: Validation, domain of application and prediction,” *J. Hazard. Mater.*, vol. 303, pp. 28–40, 2016, doi: 10.1016/j.jhazmat.2015.09.021.

## REFERENCES BIBLIOGRAPHIQUES

---

- [9] Q. Liu, J. Deng, M. Liu, and Q. Liu, "Classification models for predicting the antimalarial activity against Plasmodium falciparum Classification models for predicting the antimalarial activity," *SAR QSAR Environ. Res.*, vol. 00, no. 00, pp. 1–12, 2020, doi: 10.1080/1062936X.2020.1740890.
- [10] *World malaria report 2022 World malaria report 2022*. 2022.
- [11] X. Hou, X. Chen, M. Zhang, and A. Yan, "QSAR study on the antimalarial activity of Plasmodium falciparum dihydroorotate dehydrogenase ( PfDHODH ) inhibitors," no. February, 2016, doi: 10.1080/1062936X.2015.1134652.
- [12] J. Sainy and R. Sharma, "QSAR analysis of thiolactone derivatives using HQSAR, CoMFA and CoMSIA," *SAR QSAR Environ. Res.*, vol. 26, no. 10, pp. 873–892, 2015, doi: 10.1080/1062936X.2015.1095238.
- [13] G. W. Ashdown *et al.*, "A machine learning approach to define antimalarial drug action from heterogeneous cell-based screens," *Sci. Adv.*, vol. 6, no. 39, pp. 2–10, 2020, doi: 10.1126/sciadv.aba9338.
- [14] M. Jongdeepaisal *et al.*, "Acceptability and feasibility of malaria prophylaxis for forest goers : findings from a qualitative study in Cambodia," *Malar. J.*, pp. 1–15, 2021, doi: 10.1186/s12936-021-03983-w.
- [15] G. A. Negatu, G. A. Abebe, and W. G. Yalew, "Prevalence of Malaria and Associated Factors among Malaria-Suspected Patients Attending Hamusit Health Center, Northwest Ethiopia: A Cross-Sectional Study," *J. Parasitol. Res.*, vol. 2022, 2022, doi: 10.1155/2022/1306049.
- [16] M. G. Corral, J. Leroux, K. A. Stubbs, and J. S. Mylne, "Herbicidal properties of antimalarial drugs," *Nat. Publ. Gr.*, no. March, pp. 1–9, 2017, doi: 10.1038/srep45871.
- [17] L. H. Bannister, "Plasmodium," no. December, 2009, doi: 10.1002/9780470015902.a0001970.pub2.
- [18] S. L. Hoffman, "The March Toward Malaria," vol. 28, no. 10, pp. 1304–1314, 2015, doi: 10.1016/j.amepre.2015.09.011.The.

## REFERENCES BIBLIOGRAPHIQUES

---

- [19] O. Ba, A. Sow, H. Ba, S. Dahdi, and B. Lo, “Seasonal transmission of malaria in the senegal river valley: Case study of the city of kaedi-mauritanie,” *Pan Afr. Med. J.*, vol. 34, pp. 1–10, 2019, doi: 10.11604/pamj.2019.34.185.20011
- [20] S. M. Diallo, “Epidémiologie et diversité génétique de Plasmodium falciparum dans le département de Kobeni , Hodh Elgharbi : Implication pour la lutte antipaludique en Mauritanie,” *l’Université Nouakchott Al Aasriya Aix-Marseille Univ.*, 2020.
- [21] D. A. Milner, “Malaria Pathogenesis,” pp. 1–11, 2018.
- [22] A. Niangaly *et al.*, “Plasmodium vivax Infections over 3 Years in Duffy Blood Group Negative Malians in Bandiagara , Mali,” vol. 97, no. 3, pp. 744–752, 2017, doi: 10.4269/ajtmh.17-0254.
- [23] O. Lupi *et al.*, “Dengue infection as a potential trigger of an imported Plasmodium ovale malaria relapse or a long incubation period in a non-endemic malaria region,” *Int. J. Infect. Dis.*, vol. 44, pp. 20–24, 2016, doi: 10.1016/j.ijid.2016.01.008.
- [24] S. Neri, D. Pulvirenti, I. Patamia, A. Zoccolo, and P. Castellino, “Acute renal failure in Plasmodium malariae infection,” *Neth. J. Med.*, vol. 66, no. 4, pp. 166–168, 2008.
- [25] W. E. Collins and G. M. Jeffery, “Plasmodium malariae: Parasite and disease,” *Clin. Microbiol. Rev.*, vol. 20, no. 4, pp. 579–592, 2007, doi: 10.1128/CMR.00027-07.
- [26] Das BS. Renal failure in malaria. *Journal of Vector Borne Diseases*. 2008 Jun; 45(2): 83-97.
- [27] M. Müller and P. Schlagenhauf, “Plasmodium knowlesi in travellers, update 2014,” *Int. J. Infect. Dis.*, vol. 22, pp. 55–64, 2014, doi: 10.1016/j.ijid.2013.12.016.
- [28] M. Douine, “Epidémiologie du paludisme chez les personnes travaillant sur des sites d’orpaillage illégal en Guyane : Quels enjeux pour la santé publique ?,” 2017, [Online]. Available: <https://tel.archives-ouvertes.fr/tel-01570446>.
- [29] C. Naing, M. A. Whittaker, V. Nyunt Wai, and J. W. Mak, “Is Plasmodium vivax Malaria a Severe Malaria?: A Systematic Review and Meta-Analysis,” *PLoS Negl. Trop. Dis.*, vol. 8, no. 8, 2014, doi: 10.1371/journal.pntd.0003071.

## REFERENCES BIBLIOGRAPHIQUES

---

- [30] N. J. White, “Severe malaria,” *Malar. J.*, vol. 21, no. 1, pp. 1–17, 2022, doi: 10.1186/s12936-022-04301-8.
- [31] F. Sow, “Métacaspases : cibles thérapeutiques contre le paludisme Fatimata Sow To cite this version : HAL Id : tel-01497431 Métacaspases : Cibles thérapeutiques contre le Paludisme,” 2017.
- [32] V. Manon, “THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L ’ UNIVERSITÉ Unité de recherche MIVEGEC Impact de l ’ utilisation des médicaments antipaludiques sur la transmission de Plasmodium sp . Présentée par Manon VILLA,” 2020.
- [33] N. J. White, “Malaria parasite clearance,” *Malar. J.*, pp. 1–14, 2017, doi: 10.1186/s12936-017-1731-1
- [34] N. Tangpukdee, C. Duangdee, P. Wilairatana, and S. Krudsood, “Malaria diagnosis: A brief review,” *Korean J. Parasitol.*, vol. 47, no. 2, pp. 93–102, 2009, doi: 10.3347/kjp.2009.47.2.93.
- [35] R. Libre, “Évaluation des performances du diagnostic microscopique du paludisme dans les établissements de santé privés en Tanzanie,” 2019.
- [36] M. Tambo, M. Mwinga, and D. R. Mumbengegwi, “Loop-mediated isothermal amplification (LAMP) and Polymerase Chain Reaction (PCR) as quality assurance tools for Rapid Diagnostic Test (RDT) malaria diagnosis in Northern Namibia,” *PLoS One*, vol. 13, no. 12, pp. 1–8, 2018, doi: 10.1371/journal.pone.0206848
- [37] H. Hopkins *et al.*, “Highly sensitive detection of malaria parasitemia in a malaria-endemic setting: Performance of a new loop-mediated isothermal amplification kit in a remote clinic in Uganda,” *J. Infect. Dis.*, vol. 208, no. 4, pp. 645–652, 2013, doi: 10.1093/infdis/jit184.
- [38] J. Odaga *et al.*, “with fever in malaria endemic settings ( Review ),” no. 4, 2014, doi: 10.1002/14651858.CD008998.pub2.www.cochranelibrary.com.
- [39] E. Pillay, S. Khodajji, B. C. Bezuidenhout, M. Litshie, and T. L. Coetzer, “Evaluation of automated malaria diagnosis using the Sysmex XN-30 analyser in a clinical setting,” *Malar. J.*, vol. 18, no. 1, pp. 1–14, 2019, doi: 10.1186/s12936-019-2655-8.

## REFERENCES BIBLIOGRAPHIQUES

---

- [40] J. C. Mouatcho and J. P. Dean Goldring, “Malaria rapid diagnostic tests: Challenges and prospects,” *J. Med. Microbiol.*, vol. 62, no. PART10, pp. 1491–1505, 2013, doi: 10.1099/jmm.0.052506-0.
- [41] O. O. Oyegoke *et al.*, “Malaria diagnostic methods with the elimination goal in view,” *Parasitol. Res.*, vol. 121, no. 7, pp. 1867–1885, 2022, doi: 10.1007/s00436-022-07512-9
- [42] L. Kurtovic *et al.*, “Multifunctional Antibodies Are Induced by the RTS , S Malaria Vaccine and Associated With Protection in a Phase 1 / 2a Trial,” vol. 224, pp. 1128–1138, 2021, doi: 10.1093/infdis/jiaa144.
- [44] A. Coppi *et al.*, “The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host,” *J. Exp. Med.*, vol. 208, no. 2, pp. 341–356, 2011, doi: 10.1084/jem.20101488.
- [43] S. L. Hoffman, “The March Toward Malaria,” vol. 28, no. 10, pp. 1304–1314, 2015, doi: 10.1016/j.amepre.2015.09.011.The.
- [44] T. VALLET, “Conception d’un outil d’évaluation de l’acceptabilité des médicaments,” p.290, 2017.
- [45] A. Rayar, “In silico drug design et chimie médicinale : développement de nouvelles molécules,” p. 401, 2017.
- [46] B. Canault, “Développement d’une plateforme de prédiction in silico des propriétés ADME-Tox,” UNIVERSITÉ D’ORLÉANS ÉCOLE, 2019.
- [47] J. P. Hughes, S. Rees, S. B. Kalindjian, and K. L. Philpott, “Principles of early drug,” pp. 1239–1249, 2011, doi: 10.1111/j.1476-5381.2010.01127.x.
- [48] N. Lagarde, “Méthodes de criblage virtuel in silico : importance de l’évaluation et application à la recherche de nouveaux inhibiteurs de l’interleukine 6,” ÉCOLE DOCTORALE Abbé Grégoire Laboratoire Génomique, Bioinformatique et Applications, 2014.

## REFERENCES BIBLIOGRAPHIQUES

---

- [49] A. Rayar, “In silico drug design et chimie médicinale : développement de nouvelles molécules,” 2017.
- [50] D. Wang and Ameet Bakhai, *Clinical Trials: A Practical Guide to Design, Analysis, and Reporting*. Remedica Medical Education and Publishing ed. 2006: Andrew Ward. 480.
- [51] Hubbard, R.E. *3D structure and the drug-discovery process*. Mol. BioSyst., 2005. p. 391-406.
- [52] C. Riva, “Application de la démarche de drug-design pour la conception de nouveaux médicaments vétérinaires contre le parasite *Varroa destructor* (Acari Varroidae),” *Univ. Normandie*, no. April, p. 191, 2017.
- [53] Cusack, K. P. et al. Design strategies to address kinetics of drug binding and residence time. *Bioorg. Med. Chem. Lett.* 25, *Bioorg Med Chem Lett.*, 2015;25(10):2019-27. doi: 10.1016/j.bmcl.2015.02.027
- [54] G. Sliwoski, S. Kothiwale, J. Meiler, and E. W. Lowe, “Computational Methods in Drug Discovery,” *Pharmacol. Rev.*, 2013 Dec 31;66(1):334-95. doi: 10.1124/pr.112.007336
- [55] A. M. Thu, A. P. Phyto, J. Landier, D. M. Parker, and F. H. Nosten, “Combating multidrug-resistant *Plasmodium falciparum* malaria,” *FEBS J.*, vol. 284, no. 16, pp. 2569–2578, 2017, doi: 10.1111/febs.14127.
- [56] J. L. Bridgford *et al.*, “Artemisinin kills malaria parasites by damaging proteins and inhibiting the proteasome,” *Nat. Commun.*, no. 2018, pp. 1–9, doi: 10.1038/s41467-018-06221-1.
- [57] S. Krishna, S. Pulcini, C. M. Moore, B. H. Teo, and H. M. Staines, “Pumped up : reflections on PfATP6 as the target for artemisinins,” *Trends Pharmacol. Sci.*, vol. 35, no. 1, pp. 4–11, 2014, doi: 10.1016/j.tips.2013.10.007.

## REFERENCES BIBLIOGRAPHIQUES

---

- [58] Ralph SA, D’Ombrain MC, McFadden GI. The apicoplast as an antimalarial drug target. *Drug Resistance Updates*. juin 2001;4(3):145- 51.doi: 10.1054/drup.2001.0205
- [59] Antony HA, Parija SC. Résistance aux médicaments antipaludiques : un aperçu *Trop Parasitol*. 2016;6(1):30.doi:10.4103/22295070.175081.
- [60] Goodman CD, Buchanan HD, McFadden GI. Le mitochondrion est-il une bonne cible antipaludique ? *Tendances Parasitol*. 2017;33 (3):185–193.doi:10.1016/j.pt.2016.10.002
- [61] O. Ba, A. Sow, H. Ba, S. Dahdi, and B. Lo, “Seasonal transmission of malaria in the senegal river valley: Case study of the city of kaedi-mauritanie,” *Pan Afr. Med. J.*, vol. 34, pp. 1–10, 2019, doi: 10.11604/pamj.2019.34.185.20011.
- [62] C. P. Sanchez, W. D. Stein, and M. Lanzer, “Is PfCRT a channel or a carrier? Two competing models explaining chloroquine resistance in *Plasmodium falciparum*,” *Trends Parasitol.*, vol. 23, no. 7, pp. 332–339, 2007, doi: 10.1016/j.pt.2007.04.013.
- [63] R. A. Cooper *et al.*, “Mutations in transmembrane domains 1, 4 and 9 of the *Plasmodium falciparum* chloroquine resistance transporter alter susceptibility to chloroquine, quinine and quinidine,” *Mol. Microbiol.*, vol. 63, no. 1, pp. 270–282, 2007, doi: 10.1111/j.1365-2958.2006.05511.x.
- [64] H. Tinto, L. Guekoun, I. Zongo, R. T. Guiguemdé, U. D’Alessandro, and J. B. Ouédraogo, “Chloroquine-resistance molecular markers (Pfcrt T76 and Pfmdr-1 Y86) and amodiaquine resistance in Burkina Faso,” *Trop. Med. Int. Heal.*, vol. 13, no. 2, pp. 238–240, 2008, doi: 10.1111/j.1365-3156.2007.01995.x.
- [65] B. Gupta *et al.*, “*Plasmodium falciparum* multidrug resistance protein 1 (pfmrp1) gene and its association with in vitro drug susceptibility of parasite isolates from north-east Myanmar,” *J. Antimicrob. Chemother.*, vol. 69, no. 8, pp. 2110–2117, 2014, doi: 10.1093/jac/dku125.
- [66] D. Menard and A. Dondorp, “Antimalarial drug resistance: a threat to malaria elimination,” *Cold Spring Harb. Perspect. Med.*, vol. 7, no. 7, pp. 1–24, 2017, doi:



## REFERENCES BIBLIOGRAPHIQUES

---

- 10.1101/cshperspect.a025619.
- [67] E. N. Muratov *et al.*, “QSAR without borders Eugene,” *Chem Soc Rev*, vol. 49, pp. 3525–3564, 2020, doi: 10.1039/d0cs00098a.
- [68] R. O. P. Erkins, H. O. N. G. F. Ang, W. E. T. Ong, and W. I. J. W. Elsh, “QUANTITATIVE STRUCTURE – ACTIVITY RELATIONSHIP METHODS : PERSPECTIVES ON DRUG DISCOVERY AND TOXICOLOGY,” *Environ. Toxicol. Chem.*, vol. 22, no. 8, pp. 1666–1679, 2003.
- [69] A. C. Brown and T. R. Fraser, “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,” *Trans. R. Soc. Edinburgh*, vol. 25, no. 1, pp. 151–203, 1867, doi: 10.1017/S0080456800028155.
- [70] C. Richet, —Noté sur le rapport entre la toxicité et les propriétés physiques des corps, *Comptes rendus des séances de la Société de biologie et de ses filiales, Paris*, 45, **1893**, 775-6.
- [71] E. Overton, —Studien über die Narkose zugleich ein Beitrag zur allgemeinen Pharmakologie, *Ed. G. Fischer, Jena*, **1901**
- [72] C. Hansch, P. P. Maloney, T. Fujita, and R. M. Muir, *Nature*, **194**, 178 (1962).
- [73] C. Hansch and T. Fujita, —p- $\sigma$ - $\pi$  Analysis. A Method for the Correlation of Biological Activity and Chemical Structure, *Journal of the American Chemical Society*, 86(8), **1964**, 1616–1626.
- [74] S.M. Free and J.W. Wilson, —A Mathematical Contribution to Structure-Activity Studies, *Journal of Medicinal Chemistry*, 7(4), **1964**, 395–399.
- [75] Fujita, T.; Ban, T. Structure-activity study of phenethylamines as substrates of biosynthetic enzymes of sympathetic transmitters. *J.Med. Chem.*, **1971**, 14, 148-152. doi: 10.1021/jm00284a016.
- [76] Hansch, C.; Gao, H. Comparative QSAR: Radical reactions of benzene derivatives in chemistry and biology. *Chem. Rev.*, **1997**, 97, 2995-3060. doi: 10.1021/cr9601021.
- [77] Lowis, D.R. HQSAR: A New, Highly Predictive QSAR Technique. In: *Tripos Technical Notes*; Tripos Inc.: USA, Vol. 1, **1997**.

## REFERENCES BIBLIOGRAPHIQUES

---

- [78] Cho, S.J.; Zheng, W.; Tropsha, A. Rational combinatorial library design. 2. Rational design of targeted combinatorial peptide libraries using chemical similarity probe and the inverse QSAR approaches. *J. Chem. Inf. Comput. Sci.*, **1998**, *38*, 259-268.
- [79] Labute, P. Binary QSAR: a new method for the determination of quantitative structure activity relationships. *Pac. Symp. Biocomput.*, **1999**, 444-455.
- [80] R.D. Cramer, D.E. Patterson, and J.D. Bunce, —Comparative molecular field analysis (CoMFA) : Effet of shape on binding of steroids to carrier proteins, *Journal of the American Chemical Society*, 110(18), **1988**, 5959-5967. doi: 10.1021/ja00226a005.
- [81] G. Klebe, U. Abraham, and T. Mietzner, —Molecular similarity indices in a comparative Analysis (CoMSIA) of drug molecules to correlate and predict their biological activity *Journal of Medicinal Chemistry*, 37(24), **1994**, 4130-4146. doi: 10.1021/jm00050a010.
- [82] J. Verma, V. Khedkar, and E. Coutinho, “3D-QSAR in Drug Design - A Review,” *Curr. Top. Med. Chem.*, vol. 10, no. 1, pp. 95–115, 2010, doi: 10.2174/156802610790232260.
- [83] R. Kasmi, E. Hadaji, O. Chedadi, A. El Aissouq, M. Bouachrine, and A. Ouammou, “2D-QSAR and docking study of a series of coumarin derivatives as inhibitors of CDK (anticancer activity) with an application of the molecular docking method,” *Heliyon*, vol. 6, no. 8, p. e04514, 2020, doi: 10.1016/j.heliyon.2020.e04514.
- [84] M. Du, D. Zhang, Y. Hou, X. Zhao, and Y. Li, “Combined 2D-QSAR , Principal Component Analysis and Sensitivity Analysis Studies on Fluoroquinolones ’ Genotoxicity,” *Int. J. Environ. Res. Public Health*, p. 14, 2019.
- [85] R. Hmamouchi and M. Bouachrine, “Tentative Pratique du Relation Quantitatives Structure-Activité / Propriété ( QSAR / QSPR ),” *Rev. Interdiscip.*, vol. 1, 2016.
- [86] F. Soualmia, Thèse de doctorat, Analyse multivariée des relations structures –activités dans des nouveaux composés organiques à utilisation médicales, Université d’Oran, (2018).
- [87] K. KRAIM, Thèse de doctorat, Etude QSAR sur activités biologiques en utilisant des produits d’origines naturels, universite de Guelma,(2009)

## REFERENCES BIBLIOGRAPHIQUES

---

- [88] Aurélie GOULON-SIGWALT-ABRAM, Thèse de doctorat, Une nouvelle méthode D'apprentissage de données structurées : applications à l'aide à la découverte de Médicaments, UNIVERSITÉ PARIS 6 Pierre et Marie Curie, (2008).
- [89] K.L. Vine, J.M. Locke, M. Ranson, S.G. Pyne, J.B. Bremner, In vitro cytotoxicity evaluation of some substituted isatin derivatives, *Bioorganic & medicinal chemistry*, 15 (2007) 931-938.doi: 10.1016/j.bmc.2006.10.035.
- [90] HADAJI EL ghali, Thèse de doctorat, Corrélation - (structure - activité anticancéreuse) par les méthodes QSAR des molécules hétérocycliques précurseurs de médicaments, Université Sidi Mohammed Ben Abdellah de Fès, (2018).
- [91] J. O. Melby, X. Li, and D. A. Mitchell, "Orchestration of Enzymatic Processing by Thiazole/Oxazole- Modified Microcin Dehydrogenases," *NIH Public Access*, vol. 23, no. 1, pp. 1–7, 2008, doi: 10.1021/bi401529y.
- [92] M. Karelson, V.S. Lobanov, A.R. Katritzky, Quantum-Chemical Descriptors in QSAR/QSPR Studies, *Chem. Rev.* 96, 1027-1044, 1996. doi: 10.1021/cr950202r.
- [93] F.Luan, X. Xu, H. Liu, Cordeiro, M.N.D.S. Prediction of the baseline toxicity of non-polar narcotic chemical mixtures by QSAR approach. *Chemosphere*, 90, 1980–1986, 2013.doi: 10.1016/j.chemosphere.2012.10.06
- [94] P. Labute, —A widely applicable set of descriptors, *Journal of Molecular Graphics and Modelling*, 18, 2000, 464–477.doi: 10.1016/s1093-3263(00)00068-
- [95] U. Sakar, J. Padmanabhan, R. Parthasarathi, V. Subramanian, and P.K. Chattaraji, Toxicity analysis of polychlorinated dibenzofurans through global and local electrophilicities, *Journal of Molecular Structure: THEOCHEM*, 758(2-3), 2006.
- [96] H. Cherroun, Mémoire de Master, Étude quantitative de la relation structure/activité (QSAR) des dérivés pipéraziniques de la phénothiazine, Université Mohamed Khider de Biskra,(2019)

## REFERENCES BIBLIOGRAPHIQUES

---

- [97] Mme. ERRAHOUI née BELLIFA KHADIDJA, T H E S E Pour l'obtention du diplôme de Doctorat en Chimie Physique "Etude des relations quantitatives structure-toxicité des composés chimiques à l'aide des descripteurs moléculaires". « Modélisation QSAR », UNIVERSITE ABOU BEKR BELKAÏD DE TLEMCEM, 2015.
- [98] M. Abdullahi, A. Uzairu, G. A. Shallangwa, P. A. Mamza, and M. T. Ibrahim, "Computational modelling studies of some 1,3 - thiazine derivatives as anti - influenza inhibitors targeting H1N1 neuraminidase via 2D - QSAR , 3D - QSAR , molecular docking , and ADMET predictions,;" *Beni-Suef Univ. J. Basic Appl. Sci.*, vol. 6, 2022, doi: 10.1186/s43088-022-00280-6.
- [99] E. Burello, "Review of (Q)SAR models for regulatory assessment of nanomaterials risks," *NanoImpact*, vol. 8, no. April, pp. 48–58, 2017, doi: 10.1016/j.impact.2017.07.002.
- [100] Kubinyi, H. (1997a). QSAR and 3D QSAR in drug design Part 1: Methodology. *Drug Discover Today*, 2(11), 457–467.
- [101] Kubinyi, H. (1997b). QSAR and 3D QSAR in drug design part 2: Applications and problems. *Drug Discovery Today*, 2(12), 538–546.
- [102] L. Huang, J. Han, J. Ran, X. Chen, Z. Wang, and F. Wu, "3D-QSAR, molecular docking and molecular dynamics simulations of oxazepane amidoacetonitrile amidoacetonitrile derivatives as novel DPPI inhibitors," *J. Mol. Struct.*, 2018, doi: 10.1016/j.molstruc.2018.05.025.
- [103] BELKAIBECH IKRAM, Mémoire de master, Etude 3D-QSAR et docking moléculaire de dérivés de pyrazoline l'activité anti-amibienne contre la souche HK1 :IMSS d'E.histolytica, UNIVERSITÉ MOHAMED BOUDIAF - M'SILA, (2021).
- [104] Z. Ul-haq, A. Khan, S. Ashraf, and A. Morales-bayuelo, "Heliyon Quantum mechanics and 3D-QSAR studies on thienopyridine analogues : inhibitors of IKK  $\beta$ ," *Heliyon*, vol. 6, no. May, p. e04125, 2020, doi: 10.1016/j.heliyon.2020.e04125.
- [105] S. Ouyang *et al.*, "Development of 3D-QSAR models for predicting the activities of chemicals to stimulate muscle growth via  $\beta$ 2-adrenoceptor," *Toxicology in Vitro*, vol.

## REFERENCES BIBLIOGRAPHIQUES

---

- 77, Dec. 2021, doi:10.1016/j.tiv.2021.105251.
- [106] E. Besalu, J.V. De Julian-Ortiz, L. Pogliani, Trends and plot methods in MLR studies. *J. Chem. Inf. Model*, 47, 751–760, **2007**. doi: 10.1021/ci6004959.
- [107] H. Doreswamy and C. . M. Vastrad, “PREDICTIVE COMPARATIVE QSAR ANALYSIS OF AS 5-NITROFURAN-2-YL DERIVATIVES MYCO BACTERIUM TUBERCULOSIS H37RV INHIBITORS,” *Heal. Informatics*, vol. 2, no. 4, 2013, doi: 10.5121/hij.2013.2404.
- [108] Q. He, Y. Zhang, Y. Hu, Y. Wang, and Z. Lin, “A selectivity study of benzenesulfonamide derivatives on human carbonic anhydrase II/IX by 3D-QSAR, Molecular Docking and Molecular Dynamics Simulation,” *Comput. Biol. Chem.*, 2019, doi: 10.1016/j.compbiolchem.2019.03.005.
- [109] I. Khan, T. Ganapathi, M.- Rehman, and M. Adil, “New indenopyrazole linked oxadiazole conjugates as anti-pancreatic cancer agents : Design , synthesis , in silico studies including 3D-QSAR analysis,” *Bioorg. Med. Chem. Lett.*, vol. 44, no. April, p. 128094, 2021, doi: 10.1016/j.bmcl.2021.128094.
- [110] M. Wang, Y. Wang, D. Kong, H. Jiang, J. Wang, and M. Cheng, “In Silico Exploration of Aryl Sulfonamide Analogs as Voltage-Gated Sodium Channel 1.7 Inhibitors by Using 3D-QSAR, Molecular Docking Study, and Molecular Dynamics Simulations Authors.,” *Comput. Biol. Chem.*, 2018, doi: 10.1016/j.compbiolchem.2018.10.009.
- [111] K.Roy, I.Mitra, S.Kar, P.K.Ojha, R.N.Das, H.Kabir, Comparative studies on some metrics for external validation of QSPR models. *J Chem Inf Model*, 52, 396–408, **2012**
- [112] P.K. Ojha, I. Mitra, R.N. Das, K. Roy, Further exploring rm2 metrics for validation of QSPR models. *Chemom. Intell. Lab. Sys.*, 107, 194–205, **2011**.
- [113] Z. Almi, Thèse de doctorat, Etude qualitative et quantitative des relations structures activités dans des hétérocycles à intérêt pharmaceutique, Université Mohamed Khider Biskra, (2015).

## REFERENCES BIBLIOGRAPHIQUES

---

- [114] G. Schuurmann, R.U. Ebert, J. Chen, B. Wang, R. Kuhne, External validation and prediction employing the predictive squared correlation coefficient-Test-set activity mean vs training set activity mean. *J Chem Inf Model*, 48, 2140–2145, **2008**.
- [115] M.KHAINNAR et O.AIMEUR, “mémoire De Fin D'études” ; Etude des relations Quantitatives Structure-Activité (QSAR) de plusieurs familles de composés sur des données biologiques, Université Frères Mentouri de Constantine, (2019).
- [116] R.MEZAACHE et S.KHENIFAR, “mémoire De Fin D'études”, *Modélisation de l'activité inhibitrice MRP1 par l'approche QSAR* , Université Frères Mentouri de Constanti (2021-2022).
- [117] Kadrilege et Datacross. CHEMDRAW 18.0, LE DESSIN MOLÉCULAIRE PAR EXCELLENCE [enligne] <https://ritme.com/fr/logiciels/chemdraw-chemoffice/> (consulté le 03/07/2019)
- [118] A. Hinchliffe, “HyperChem Release 4.5 for Windows,” *Electron. J. Theor. Chem.*, vol. 1, no. 1, pp. 62–64, Mar. 2001, doi: 10.1002/ejtc.9.
- [119] U. Hypercube, Inc., Gainesville, Florida, “HyperChem.” 1999. [Online]. Available: <http://www.hyper.com> (consulté le 25/05/2023)
- [120] Talete s.r.l. Dragon 7 available now [enligne] [http://www.talete.mi.it/products/dragon\\_description.htm](http://www.talete.mi.it/products/dragon_description.htm) (consulté le 25/05/2023)
- [121] W. Wu and D. L. Massart, “Artificial neural networks in classification of NIR spectral data: Selection of the input,” *Chemom. Intell. Lab. Syst.*, vol. 35, no. 1, pp. 127–135, Nov. 1996, doi: 10.1016/S0169-7439(96)00034-2.
- [122] Tripos Associates Inc.. SYBYL Molecular Modelling Package. St. Louis, Mo, USA: Tripos Associates Inc.; 2001.
- [123] O. Chedadi, A. El Aissouq, Y. El Ouardi, M. Bouachrine, and A. Ouammou, “3D-QSAR and molecular docking studies of 4-methyl quinazoline derivatives as PI3K $\alpha$  inhibitors,” *Journal of the Indian Chemical Society*, vol. 98, no. 11, Nov. 2021, doi: 10.1016/j.jics.2021.100183.
- [124] Gasteiger J, Saller H. Calculation of the charge distribution in conjugated systems by a quantification of the resonance concept. *Angew Chem. Int. Ed. Engl.* 1985;24: 687–9.

## REFERENCES BIBLIOGRAPHIQUES

---

- [125] Mohan Krishna K, Inturi B, Pujar G V., Purohit MN, Vijaykumar GS. Design, synthesis and 3D-QSAR studies of new diphenylamine containing 1,2,4-triazoles as potential antitubercular agents. *Eur J Med Chem.* 2014;84:516-529.
- [126] Zhu W, Liu Y, Zhai X, et al. Design, synthesis and 3D-QSAR analysis of novel 2-hydrazinyl-4-morpholinothieno[3,2-d]pyrimidine derivatives as potential antitumor agents. *Eur J Med Chem.* 2012;57 :162-175.
- [127] Zhao B, Zhao C, Hu X, et al. Design, synthesis and 3D-QSAR analysis of novel thiopyranopyrimidine derivatives as potential antitumor agents inhibiting A549 and Hela cancer cells. *Eur J Med Chem.* 2020;185:111809.
- [128] Meinguet C, Bruyère C, Frédérick R, et al. 3D-QSAR, design, synthesis and characterization of trisubstituted harmine derivatives with *in vitro* antiproliferative properties. *Eur J Med Chem.* 2015;94:45-55.

## REFERENCES BIBLIOGRAPHIQUES

---



**Année universitaire :**

2022-2023

**Présenté par :** DABOUZ Radhia

BUHEMHEM SARRA

**Etude comparative de l'efficacité prédictive des modèles 2D-QSAR et 3D-QSAR pour la conception de molécules anti-malaria.**

**Mémoire pour l'obtention du diplôme de Master en Biochimie Appliquée**

**Résumé :**

Ce mémoire a pour objectifs d'établir des modèles QSAR fiables, stables et prédictifs. La méthodologie QSAR dont deux approches (DRAGON-Mobydigs et Sybyl) sont utilisées pour modéliser l'activité inhibitrice ENR représenté par ( $pIC_{50}$ ) de 61 dérivés du Triclosan, en comparant leurs résultats, ainsi choisir le model le plus performant.

Les modèles établis de ces deux approches ont prouvé leur validation interne et externe confirmant leurs stabilités, leurs robustesses et leurs capacités prédictives.

Pour le modèle DRAGON-Mobydigs ( $R^2=77,55\%$ ,  $Q^2=72,45\%$ ) dont l'activité inhibitrice a été corrélés avec 4 descripteurs (DECC, BEHe4, L2U, C-026). Les résultats sont conformes aux principales recommandations du Golbraikh et Tropsha.

Les résultats obtenus par Sybyl :CoMFA ( $R^2 = 94.40\%$ ,  $Q^2= 64.330\%$ ), CoMSIA ( $R^2 = 95.57\%$ ,  $Q^2 = 53.51\%$ ) nous ont permis de conclure que ce dernier possède de faible paramètres prédictifs internes et externes, et un caractère très robuste et stable pouvant prédire l'activité pour de nouveaux composés.

Enfin, le model Sybyl ne peut pas être utilisé efficacement pour estimer des molécules qui n'ont pas été testées expérimentalement.

**Mots clés :** Triclosan, ENR, QSAR.

**Laboratoire de recherche :** laboratoire de Bio-informatique centre de recherche de la biotechnologie (CRBt), Constantine.

**Jury devaluation:**

- **Président:** Mr.MOKRANI ELHASSEN MCA. UFM Constantine-1
- **Encadrant:** Mr.BOUHEDJAR KHALID. MRA CRBt Constantine.
- **Examineur :** Mm.CHERIET MOUNA. MRB CRBT Constantine

**Date de dépôt :** 18-06-2022

