جامعة سيدي محمد بن عبد الله بفاس
UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH DE FES

كلية العلوم والتقنيات فاس
FACULTÉ DES SCIENCES ET TECHNIQUES DE FÈS

N° d'ordre : 73/2023

Centre d'Etudes Doctorales : Sciences et Techniques et Sciences Médicales

Formation doctorale : Sciences et Génie de la Matière de la Terre et de la Vie

Discipline : **Chimie**

Spécialité : **Chimie physique et modélisation**

Thèse pour l'obtention de DOCTORAT

Études *in silico* sur la recherche et développement (R&D) de nouveaux médicaments destinés aux maladies neurodégénératives

Préparée par : EL AISSOUQ Abdellah

Soutenue le Vendredi 20 Octobre 2023 à 16 h – Salle de Conférence de la Faculté des Sciences et Techniques de Fès devant le jury :

Nom et Prénom	Etablissement	grade	Qualité
Pr OULMEKKI Abdallah	Faculté des sciences et techniques, Université Sidi Mohamed Ben Abdellah	PES	Président
Pr MAGHAT Hamid	Faculté des sciences, Université Moulay Ismail Meknès	PES	Rapporteur
Pr AJANA Mohammed Aziz	Faculté des sciences, Université Moulay Ismail Meknès	PES	Rapporteur
Pr SAFFAJ Taoufiq	Faculté des sciences et techniques, Université Sidi Mohamed Ben Abdellah	PES	Rapporteur
Pr ZAITAN Hicham	Faculté des sciences et techniques, Université Sidi Mohamed Ben Abdellah	PES	Examinateur
Pr BOUACHRINE Mohammed	Faculté des sciences, Université Moulay Ismail Meknès	PES	Examinateur
Pr OUAMMOU Abdelkrim	Faculté des sciences Dhar El Mahraz, Université Sidi Mohamed Ben Abdellah	PES	Examinateur
Pr LAKHLIFI Tahar	Faculté des sciences, Université Moulay Ismail Meknès	PES	Invité
Pr KHALIL Fouad	Faculté des sciences et techniques, Université Sidi Mohamed Ben Abdellah	PES	Directeur de thèse

Dédicace

Je dédie ce travail à

A ma très chère mère

Quoi que je fasse ou que je dise, je ne saurai point te remercier comme il se doit. Ton affection me couvre, ta bienveillance me guide et ta présence à mes côtés a toujours été ma source de force pour affronter les différents obstacles.

A mon très cher père

Tu as toujours été à mes côtés pour me soutenir et m'encourager.

Que ce travail ma gratitude et mon affection.

A mes très chères sœurs : Fatima, Fatiha et Khadija

A mes très chers frères : Mohamed, Miloud, Hamid et Ahmed

Qui m'ont énormément aidée et soutenue et à qui je témoigne mon affection. Je ne peux exprimer à travers ces lignes tous mes sentiments d'amour et de tendresse envers vous. Puisse l'amour et la fraternité nous unissent à jamais.

A mes très chers amis

J'admirerai toujours votre gentillesse et votre humour. J'espère que notre amitié sera éternelle.

A tous ceux qui me sont chers.

Remerciements

Je remercie tout d'abord **Dieu** tout puissant de m'avoir donné la force, la connaissance, la capacité et l'opportunité d'entreprendre cette étude et de la terminer de manière satisfaisante.

Je tiens à remercier et à témoigner toute ma reconnaissance à mon cher Directeur de Thèse Professeur **FOUAD KHALIL** de la Faculté des Sciences et Techniques (FST), Université Sidi Mohammed Ben Abdellah (USMBA) de Fès pour la confiance qu'il m'a accordée dès mon arrivée et pour l'excellent encadrement dont j'ai bénéficié durant cette thèse. Un grand merci pour avoir été à l'écoute de mes besoins. Je vous remercie pour m'avoir guidé dans mes recherches dès la première expérience et pour vos conseils tous très perspicaces. Merci pour votre patience, votre calme, votre gentillesse, votre persévérance et votre bonne humeur, je suis honoré de vous avoir comme directeur de thèse.

Mes remerciements vont également au professeur **OULMEKKI Abdallah**, Président et membre du jury qui a accepté de présider et de faire partie du jury, veuillez trouver ici mes hautes considérations.

Je tiens à remercier les rapporteurs et membres du jury, **Pr. MAGHAT Hamid**, **Pr. AJANA Mohammed Aziz** et **Pr. SAFFAJ Taoufiq**, qui ont accepté de rapporter et de juger ce travail. Qu'ils trouvent ici l'expression de ma plus haute considération et de ma sincère gratitude.

Je tiens aussi à remercier les examinateurs et membres du jury, **Pr. ZAITAN Hicham, Pr. BOUACHRINE Mohammed et Pr. OUAMMOU Abdelkrim,** pour leur temps, leur expertise et leurs suggestions constructives lors de la revue de ma thèse. Vos commentaires éclairés et vos critiques constructives m'ont permis d'améliorer considérablement la qualité de mon travail.

Je tiens également à exprimer ma reconnaissance à Monsieur **LAKHLIFI Tahar**, professeur émérite de la Faculté des Sciences de Meknès, pour l'intérêt qu'il a manifesté en participant en qualité de membre invité à ce jury.

Mes remerciements vont également à mes collègues de laboratoire, **Said El Rhabori, Lhoucine Naanaai** et **Yassine El Allouche** pour leur soutien tout au long de ce projet. Vos discussions stimulantes et votre collaboration ont été essentielles à mon développement en tant chercheur.

Je suis également reconnaissant envers mes amis et ma famille pour leur soutien inconditionnel et leurs encouragements tout au long de ce parcours académique exigeant. Vos encouragements et votre présence dans ma vie ont été une source constante de motivation.

Je suis honoré d'avoir eu l'opportunité de travailler sur ce projet de thèse et je suis conscient que sans le soutien et la contribution de chacun d'entre vous, cela n'aurait pas été possible. Merci encore pour votre confiance, votre expertise et votre soutien indéfectible.

Résumé

La recherche et dévelopement (R&D) d'un nouveau médicament est un vaste domaine qui applique des techniques scientifiques provenant de divers domaines pour aider à la conception et au développement de nouveaux agents thérapeutiques. Ce processus complexe joue un rôle majeur dans l'avancement de la médecine moderne et dans la lutte contre les maladies. L'objectif de ce travail est de développer de nouveaux agents thérapeutiques pour traiter les maladies d'Alzheimer et de Parkinson en utilisant des techniques de criblage *in silico* et des analyses bio-informatiques. Dans un premier temps, des séries de molécules synthétiques et à base de plantes ont été sélectionnées pour prédire les propriétés pharmaceutiques de nouvelles molécules en utilisant diverses méthodes telles que l'analyse 2D/3D-QSAR, la modélisation par homologique, le docking moléculaire, la simulation par dynamique moléculaire (MD) et l'analyse des propriétés ADME-Tox. Les résultats obtenus dans cette thèse sont présentés en deux grandes parties : La première concerne les résultats relatifs à la maladie de Parkinson tandis que la deuxiéme présente les résultats relatifs à la maladie d'Alzheimer. Tous, les modèles élaborés présentent respectivement des valeurs de validation interne Q^2 et externe R^2_{pred} élevées. De même, les cartes de contour dérivées des analyses CoMFA/CoMSIA indiquent l'importance des champs électrostatiques, l'effet hydrophobe et Accepteurs/Donneurs de liaisons hydrogènes dans l'amélioration de l'activité biologique étudiée. De plus, des simulations par dynamique moléculaire ont confirmé la stabilité des systèmes complexes sélectionnés durant les simulations par MD. Sur la base des méthodes de criblage *in silico* et des méthodes de la modélisation moléculaire, des series de molécules ont été prédites avec des meilleures affinités, une meilleure activité anti-Alzheimer et anti-Parkinson et des propriétés ADME-Tox respectant les règles. Ainsi, ces molécules pourraient être des candidats médicaments potentiels pour les maladies neurodégénératives.

Mots clés : Activité anti-Alzheimer, activité anti-Parkinson, 2D/3D-QSAR, Modélisation par homologie, Docking moléculaire, Simulations par MD, ADME-Tox

Résumé en anglais

New drug research and development (R&D) is a vast field that applies scientific techniques from a variety of fields to help design and develop new therapeutic agents. This complex process plays a major role in the advancement of modern medicine and the fight against disease. This research aims to use *in silico* screening techniques and bioinformatics analysis to identify novel treatment medicines for Alzheimer's and Parkinson's diseases. Initially, a series of synthetic and plant-based compounds were chosen to predict the pharmacological characteristics of novel molecules using a variety of methodologies including 2D/3D-QSAR analysis, homology modeling, molecular docking, molecular dynamics (MD) simulation, and ADME-Tox properties analysis. The results of this study are divided into two sections: the first discusses the findings for Parkinson's disease, and the second discusses the findings for Alzheimer's disease. All of the created models have strong internal validation values Q^2 and external validation values R^2_{pred}. Similarly, the contour maps derived from the CoMFA/CoMSIA analyses indicate the importance of electrostatic fields, the hydrophobic effect and hydrogen bond acceptors/donors in enhancing the studied biological activity. In addition, molecular dynamics simulations confirmed the stability of the complex systems during the MD simulations. Based on the above results, we predicted *in silico* the design of molecules with better affinities, better anti-Alzheimer and anti-Parkinson activity and ADME-Tox properties respecting the rules. These molecules could thus be potential drug candidates for neurodegenerative diseases.

Keywords: Anti-Alzheimer activity, anti-Parkinson activity, 2D/3D-QSAR, Homology modeling, Molecular docking, MD simulations, ADME-Tox,

Table de matières

Dédicace ..... **i**
Remerciements ..... **ii**
Résumé ..... **iv**
Résumé en anglais ..... **v**
Table de matières ..... **vi**
List d'abréviation ..... **ix**
Listes des figures ..... **xi**
Application 2: ..... **xii**
Liste des tableaux ..... **xv**
Introduction générale **1**

Partie I: Bibliographie **3**

Chapitre I: Maladies Neurodégénérative **4**
I. Introduction **5**
II. Maladie d'Alzheimer **7**
II.1. Généralité 7
II.2. Neuropathologies 8
II.2.1. Plaques amyloïdes 8
II.2.2. Dégénérescences neurofibrillaires (DNF) 9
II.2.3. Perte de neurones et de synapses 10
II.3. Rôle de GSK-3 dans la maladie d'Alzheimer 10
II.4. Rôle d'AChE dans la maladie d'Alzheimer 11
II.5. Rôle de BuChE dans la maladie d'Alzheimer 11
II.6. Traitement de la maladie d'Alzheimer 12
III. Maladie de Parkinson (MP) **14**
III.1. Histoire de la maladie de Parkinson 14
III.2. Pathogénèse 15
III.2.1. Hypothèse de l'α-synucléine 15
III2.2. Hypothèse du stress oxydant et de la mitochondrie 16
III.3. Traitement des symptômes moteurs de la MP 16
III.3.1. Lévodopa 17
III.3.2. Inhibiteurs de la catéchol-o-méthyl-transférase (COMT) 17
III.3.3. Inhibiteurs de la monoamine oxydase de type B (MAO-B) 17
III.3.4. Agonistes de la dopamine (DA) 18
Chapitre II : Famille des chalcones **19**
I. Introduction **20**
II. Types de chalcones **20**
II.1. Chalcones naturels 20
II.2. Chalcones synthétiques 21
II.2.1. Synthèse organique 21
II.2.2. Biosynthèse des chalcones 21
III. Activités pharmacologique **22**
III.1. Activité anti-Parkinson 22

III.2. Activité anti-Alzheimer 23
III.3. Autres activités biologiques 24
Chapitre III: Processus de découverte de nouveaux médicaments 26
I. Histoire de la découverte des médicaments 27
II. Processus de recherche et développement (R&D) des médicaments 28
II.1. Identification de la cible protéique 28
II.2. Identification de "hits" 29
II.3. Recherche et optimisation de « leads » 30
II.4. Tests pré-cliniques 30
II.5. Tests cliniques 30
III. Méthodes de criblages virtuels *in silico* 31
III.1. Conception de médicaments basé sur le ligand « Ligand-Based» 33
III.1.1. Recherche par similarité 33
III.1.2. Pharmacophores basé sur le ligand 34
III.1.3. Relation quantitative structure-activité (QSAR) « ligand-based » 36
III.1.3.1. Principe et méthodologie 36
III.1.3.2. Importance de la base de données 38
III.1.3.3. Méthodes d'optimisation 38
III.1.3.4. Étude 2D-QSAR 38
a. Types de descripteurs 38
b. Analyses statistiques et formation du modèle 2D-QSAR 39
b.1. Analyse par composantes principales (ACP) 39
b.2. Régression linéaire multiple (RLM) 40
b.3. Méthode des moindres carrés partiels (PLS) 42
b.4. Régression non linéaire multiple (RNLM) 43
b.5. Réseaux de neurones artificiels (RNA) 43
c.Validation du modèle QSAR 44
c.1. Validation interne 44
 Validation croisée (CV) 44
 Test de randomisation en Y 45
C.2. Validation externe 45
d. Domaine d'applicabilité 46
III.1.3.5. Etudes 3D-QSAR 47
a. Alignement des molécules 47
b. Méthode CoMFA 48
c. Méthode CoMSIA 48
d. Elaboration du modèle 49
e. Validation du modèle 49
f. Interprétation graphique 49
IV. Criblage virtuel basé sur la structure « Structure-Based » 50
IV.1. Modélisation par Homologie 50
IV.2. Docking Moléculaire 51
IV.2.1. Docking de molécules rigides 52
IV.2.2. Docking flexible 53
IV.3. Simulations par dynamique moléculaire 53
IV.3. Méthodes de calcul de l'énergie libre de liaison 55
Chapitre IV: Etudes pharmacocinétiques 57

I. Introduction 58
1. II. Etude ADMET 59
a. Absorption 59
b. Distribution 60
c. Métabolisme 60
d. Excrétion 61
e. Toxicité 61
2. III. Prédiction des spectres d'activité des substances (PASS) 62
3. Références bibliographiques 63

Partie II: Travaux Réalisés (Applications sur la Maladie de Parkinson) 80

Chapitre I : Étude computationnelle de dérivés cétoniques insaturés en tant qu'inhibiteurs de la MAO-B à l'aide de QSAR, ADME/Tox, docking moléculaire et simulations par dynamique moléculaire 83

Chapitre II: Développement de nouveaux inhibiteurs de la monoamine oxydase B (MAO-B) par la combinaison de docking, du 3D-QSAR, de la prédiction ADMET, de la simulation par dynamique moléculaire et de l'énergie libre de liaison MM_GBSA 113

Chapitre III : Identification des agents anti-Parkinsoniens ciblant la protéine MAO-B en utilisant la modélisation par homologie, le docking moléculaire, la simulation par dynamique moléculaire (MD) et les approches ADMET 141

Partie II : Travaux Réalisés (Applications sur la Maladie d'Alzheimer) 162

Chapitre IV : Études de modélisation moléculaire de dérivés de la C-Glycosylfavone en tant qu'inhibiteurs de GSK-3β, basées sur l'étude QSAR et l'analyse de docking 164

Chapitre V: Application des méthodes CADD à la recherche et développement de nouveaux inhibiteurs de la p38a MAP kinase : 2D-QSAR, docking moléculaire, dynamique moléculaire et propriétés ADMET 181

Conclusion générale 208
Annexes 210
Liste des publications 215
Liste des communications orales 218
Liste des communications affichées 221

List d'abréviation

ACh	Acétylcholine
AChE	Acétylcholinestérase
ACP	Analyse en Composantes Principales
ADME	Absorption, Distribution, Métabolisme, Elimination
APP	Protéine Précurseur de l'Amyloïde
Aβ	Amyloïde-β
BHE	Barrière Hémato-Encéphalique
BuChE	Butyrylcholinestérase
CADD	Computer Aided Drug Design
CHS	Chalcone Synthase
CoA	Coenzyme A
CoMFA	Comparative Molecular Field Analysis
CoMSIA	Comparative Molecular Similarity Indices Analysis
COMT :	Catéchol-Ométhyltransférase
COX	Cyclo-Oxygénases
CP	Composante Principale
CV	Cross-Validation
DA	Domaine d'Applicabilité
DM	Dynamique Moléculaire
DNF	Dégénérescence Neurofibrillaire
ERO	Espèces réactives de l'oxygène
FDA	Food et Drug Administration
GSK	Glycogène synthase kinase
HBA	Accepteurs de Liaisons Hydrogène
HBD	Donneurs de Liaisons Hydrogène
HTS	high-throughput screening
LBDD	Ligand based drug design
L-Dopa	Lévodopa
LOO	leave-one-out
LRRK2	Leucine-Rich Repeat Kinase
LSO	Leave-Several-Out
MA	Maladie d'Alzheimer

MAO-B	Monoamine Oxydase de type B
MM-PB/SA	Mécanique Moléculaire-Poisson Boltzmann/Surface
MN	Maladies neurodégénératives
MP	Maladie de Parkinson
MPTP	1-Methyl-4-Phenyl-1,2,3,6-Tetrahydropyridine
MW	Poids Moléculaire
NMDA	N-méthyl-D-aspartate
OMS	Organisation Mondiale de la Santé
PASS	Prediction of Activity Spectra for Substances
PB	Poisson-Boltzmann
PDB	Protein Data Bank
PLS	Régression des Moindres Carrés Partiels
p-Tau	Tau hyperphosphorylé
QSAR	Relation Structure Activity Relationship
R&D	Recherche et Développement
RC	Restriction Calorique
RNA	Réseaux de Neurones Artificiels
REQM	Racine de l'Erreur Quadratique Moyenne
RLM	Régression Linéaire Multiple
RNLM	Régression Non Linéaire Multiple
RSA	Relations Structure-Activité
SBDD	Structure Baed Drug Design
SDEP	Standard Deviation Error of Prediction
SEAL	Steric and Electrostatic ALignment
SIDA	Syndrome d'Immuno-Déficience Acquise
SNc	Substantia Nigra pars compacta
T-DA	Transporteur de Dopamine
α-syn	α-synucléine

Partie Bibliographie

Listes des figures

Figure I.1. Lésions neuropathogiques de la maladie d'Alzheimer 7

Figure I.2. Voies impliquées dans les processus de métabolisation de l'APP 9

Figure I.3. Hydrolyse de l'acétylcholine 11

Figure I.4. Médicaments anti-Alzheimer commercialisés 12

Figure.I.5. Mécanisme de phosphorylation de l'AChE par un organophosphate 14

Figure I.6. Paralysie agitans. Image tirée de W.R. Gowers "Un manuel des maladies du système nerveux " 14

Figure II.1. Représentation structurelle et numérique de la structure de la chalcone 20

Figure II.2. Réaction de condensation de Claisen-Schmidt 21

Figure II.3. Biosynthèse des chalcones 22

Figure II.4. Relation structure activité (RSA) des dérivés de la chalcone 23

Figure II. 5. Activités biologiques présentées par les analogues de la chalcone 24

Figure II.6. Analogues de chalcone utilisés en médecine 25

Figure III.1. Étapes du processus de recherche et développement (R&D) des médicaments 28

Figure III. 2. Classifications des méthodes de criblage in silico 32

Figure III. 3. Etapes principales du modèle pharmacophore 3D (A : accepteur de liaison hydrogène ; D : donneur de liaison hydrogène ; R : aromatique) 34

Figure III. 4. Exemple d'un modèle pharmacophore montrant les différents points pharmacophores d'une molécule dérivée de la foranone 35

Figure III. 5. Méthodologie générale des études QSARs 37

Figure III.6. Perspective dimensionnelle de la technique QSAR 39

Figure III.7. Principe de mise en place de deux composantes principales CP1 et CP2 dans un espace à 3D (x1, x2, x3) lors d'une analyse ACP 40

Figure III.8. Illustration de l'architecture d'un réseau de neurones à trois couches **(A)** et le fonctionnement interne des neurones dans une couche cachée **(B)** 44

Figure III.9. Schéma de construction d'un modèle 3D-QSAR 50

Figure III. 10. Etapes de prédiction d'une structure 3D d'une protéine par homologie 51

Figure III. 11. Schéma simplifié du docking moléculaire 52

Partie expérimentale

Listes des figures

Application 1:

Figure a1. Structure générale des squelettes chimiques possédant des propriétés inhibitrices de la MAO-B 81

Figure a2. Structures chimiques des chalcones (1), des dérivés esters α, β-insaturés (2), et des dérivés amides α, β-insaturés (3) 82

Figure 1. Mechanism of interaction of monoamine oxidase-B (MAO-B) inhibitors. AD = dopamine, L-Dopa = levodopa, AADC = aromatic L-amino acid decarboxylase, 3-OMD= 3-O-methyldopa, COMT= catechol-O-methyltransferase, TH= tyrosine hydroxylase, and DOPAC = 3, 4-dihydroxyphenylacetic acid 88

Figure 2. Chemical structures of chalcones (1), α, β-unsaturated ester derivatives (2), and α, β-unsaturated amide derivatives (3) 88

Figure 3. Core sub-structure of unsaturated ketone derivatives 89

Figure 4. Plot of predicted and observed activity. The blue dots denote internal validation (LOO), whereas the orange dots represent external validation 100

Figure 5. Contribution of descriptors in the generated 2D-QSAR model 100

Figure 6. Applicability domain of the generated 2D-QSAR model. The blue dots represent the training set compounds and the red dots represent the test set compounds 102

Figure 7. Correlation between the binding free energy values and the pIC50 values 104

Figure 8. Docking analysis of compound 10b. (a) 2D view of binding site interactions, (b) 3D view of the binding conformation 105

Figure 9. Docking analysis of compound **10e**. **(a)** 2D view of binding site interactions, **(b)** 3D view of the binding conformation 105

Figure 10. Molecular dynamics simulations of the 10b_2BK3 complex. **(a)** represents the RMSD plot, **(b)** represents the RMSF plot, **(c)** represents the Rg plot, and **(d)** represents the superposition of the final complex structure after MD simulation of 20 ns (green color) and initial complex structures before MD simulation (blue color) 106

Application 2:

Figure a1. Molécules prédites 114

Figure 1. Mechanism of amine oxidation by monoamine oxidases (MAOs) 117

Figure 2. Chemical structure of chalcone (1,3-diphenyl-2-proper-1-one) 117

Figure 3. Superposition of the native ligand (green color) and re-docked ligand (red color) in the binding pocket of MAO-B enzyme 120

Figure 4. Alignment of optimal conformations inside binding pocket of MAO-B enzyme . 121

Figure 5. Correlation between experimental pIC50 values and binding affinity values of compounds. Blue and red colors represent respectively the dataset and predicted compounds 123

Figure 6. Observed and predicted pIC_{50} values. The orange and blue dots, respectively, indicate the training and set compounds 125

Figure 7. Plot of observed versus predicted activities. Blue dots represent the training set compounds and green dots represent the test set compounds calculated by the CoMSIA method 125

Figure 8. Steric contour maps of CoMFA **(a)** and CoMSIA **(b)** based on compound **10b**. The green contours represent regions where bulky groups are favored and yellow contours represent regions where bulky groups are disfavored 127

Figure 9. Electrostatic contour maps of CoMFA **(a)** and CoMSIA **(b)** based on compound 10b. The red color represents regions where electronegative groups are favored and the blue color represents regions where electronegative groups are disfavored 128

Figure 10. H-bond acceptor contour maps of CoMSIA model, based on compound **10b**. The red color represents regions where the addition of H-bond acceptor groups decrease activity and the magenta color represents regions where the addition of H-bond groups increase activity 128

Figure 11. Structure-activity relationship (SAR) information are obtained by 3D-QSAR studies 129

Figure 12. 2D interactions of compound **T1** with the MAO-B enzyme (PDB ID: 2BK3) ... 131

Figure 13. 2D interactions of compound **10b** with the MAO-B enzyme (PDB ID: 2BK3) 131

Figure 14. RMSD plots of 100 ns of MD simulation. (a) for backbone and (b) for ligand. In all figures, the red color represents the 2BK3_T1 complex and the black color represents the 2BK3_b10 complex 135

Figure 15. MD simulation of ligand-protein complex. (a) and (b) represent, respectively, the plots of RMSF (C-alpha) and Rg over the 100 ns of MD silulation at 300 K. In all figures, the red and black colors represent the 2BK3_T1 and 2BK3_b10 complexes 135

Application 3:

Figure a1. Potentiel inhibiteurs de la MAO-B : P34 (Crotolaria rosmosissima), P60 (Tephrosia spinosa) et P81(Artocarpus bracteata) 143

Figure 1. Chemical structures of MAO-B drugs and chalcone : **(1)** selegiline, **(2)** rasagiline, **(3)** safinamide and **(4)** chalcone 145

Figure 2. 3D structure of the modeled hMAO-B protein 148

Figure 3. Homology modeling of hMAO-B protein. (a), superposition of 3D structures between the modeled structure and the 1GOS (Chain A) template, (b), Ramachandran plot of best hMAO-B model, (c), ProSA z-score plot of the modeled structure 149

Figure 4. Binding pocket of the modeled hMAO-B protein 150

Figure 5. 3D-binding site between the target protein and the top-ranked compounds. **(A)** P34 (-10.0 Kcal/mol), **(B)** P60 (-10.77 Kcal/mol), **(C)** P81 (-10.64 Kcal/mol) and **(D)** Safinamid (-7.91 Kcal/mol)..... 156

Figure 6. RMSD plots of 100 ns of MD simulation. **(a)** For backbone and **(b)** for ligand. In all figures, the red black represents the hMAO-B_P60 complex, the red represents hMAO-B_P81 complex, and the green color represents the hMAO-B_Safinamid complex..... 157

Figure 7. MD simulation of ligand-protein complex. **(a)** and **(b)** represent, respectively, the plots of RMSF (C-alpha) and Rg over the 100 ns of MD silulation at 300 K. In all figures, the black, red, and green colors represent the hMAO-B_P60, hMAO-B_P81, and hMAO-B_Safinamid complexes, respectively 158

Application 4:

Fig. 1. Plot of predicted versus observed activities. (a) the generated qsar model, (b) the internal and external validations..... 173

Fig. 2. Williams plot for the developed 2D-QSAR model. Blue values represent training set compounds and red values represent test set compounds 174

Fig. 3. (a) Superposition of the co-crystalized ligand on the GSK-3β pocket (the yellow stick represents the re-docked ligand; the red stick represents the reference ligand). (b) poseview 2D interaction diagram of the reference ligand..... 175

Fig. 4. Docking results of compounds L27 (a) and l6 (b) with GSK-3β protein (PDB code: 1pyx)..... 175

Fig. 5. Molecular docking of compound T1: (a) 2D view of binding site interactions, (b) 3D view of the binding conformation 177

Fig. 6. Molecular docking of compound T3: (a) 2D view of binding site interactions, (b) 3D view of the binding conformation 177

Application 5:

Figure 1. (a) Plot of predicted activities versus experimental ones, (b) standardisd residues of training set compounds..... 195

Figure 2. Williams plot for the developed 2D-QSAR model. Red and blue dots represent the test and training sets compounds, respectively 197

Figure 3. Docking validation. **(a)** Superposition between the re-docked and the original ligands into the binding pocket of the p38α MAPK (PDB ID: 5ML5), **(b)** 2D interactions of the native ligand into the binding p38α MAPK receptor..... 198

Figure 4. Docking analysis of compound L5 into the binding pocket of p38α MAPK receptor. **(a)** 2D ligand-receptor complex interactions, **(b)** 3D ligand-receptor complex interactions. 198

Figure 5. The calculated partial charge of the most active compound in the dataset (Compound L5) 199

Figure 6. Ligand RMSD plot of 50 ns of MD simulation. The green color represents the p38α MAPK _T1 complex, the black color represents the p38α MAPK _T2 complex, and the red color represents the p38α MAPK _L5 complex..... 201

Figure 7. Plot of Rg over the 50 ns of MD simulation at 300 K. The green color represents the p38α MAPK _T1 complex, the black color represents the p38α MAPK _T2 complex, and the red color represents the p38α MAPK _L5 complex..... 202

Figure 8. Graphical representation of hydrogen bonding of p38α MAPK _T1 (green), p38α MAPK _T2 (red) and p38α MAPK _L5 (black) complexes..... 203

Liste des Tableaux

Application 1:

Table 1. Chemical structures of unsaturated ketone derivatives and their corresponding experimental activity 90

Table 2. VIF and p values of the selected descriptors..... 97

Table 3. Correlation matrix of the selected descriptors..... 97

Table 4. Values of selected descriptors, experimental activity and predicted activity..... 98

Table 5. Accepted 2d-qsar model validation tools 99

Application 2:

Table 1. Dataset of unsaturated ketone derivatives and their corresponding experimental activity..... 119

Table 2. Statistical parameters of comfa and comsia models using PLS analysis 124

Table 3. Observed activity, predicted activity, and residual values of the comfa and comsia methods 126

Table 4. Predicted activity and binding affinity score of the newly designed compounds ... 129

Table 5. *in silico* prediction of molecular properties and synthetic accessibility of proposed compounds 132

Table 6. *in silico* ADMET prediction of newly predicted compounds and the most potent inhibitor in the dataset..... 134

Table 7. Binding free energy of 2bk3_b10 and 2bk3_t1 complexes using MM-GBSA method 136

Application 3:

Table 1. Interpretation of some admet parameters 147

Table 2. Binding affinity of the best ranked compounds 151

Table 4. ADMET characteristics of the top ranked compounds 152

Table 5. Binding affinity of the top ranked compounds by admet analysis 153

Table 6. 2D-binding site of P34, P60, P81, and safinamid in the binding pocket of hMAO-B receptor..... 154

Application 4:

Table 1. Chemical structures and their experimental biological activity of c-glycosylfavones 168

Table 2. The correlation matrix of the selected descriptors 172

Table 3. The vif value of the selected descriptors 172

Table 4. Predicted and observed with residual values for QSAR model and leave-one-out (LOO) cross validation 173

Table 5. Chemical structures and their predicted pIC_{50} and binding free energy values 176

Application 5:

Table 1. Structures of chemicals and their corresponding activity values (* represents the test set compounds) 185

Table 2. Matrix of correlation of the selected descriptors 194

Table 4. The value of the observed activity, descriptors, predicted activity, and residual 196

Table 5. Chemical structures and their predicted activity as well as the binding free energy of predicted compounds 200

Table 6. Preadmet server predicted pharmacokinetic properties 204

Table 7. Preadmet server predicted toxicity parameters 204

Table 8. Drug-likeness and synthetic accessibility of most active hits 205

Introduction générale

Les maladies neurodégénératives (MNs), notamment la maladie d'Alzheimer, la maladie de Parkinson, la maladie de Huntington et autres, sont des démences qui entraînent une dégénérescence et/ou une mort progressive des cellules nerveuses du système nerveux central (SNC) [1]. Le taux de démence est en augmentation croissante dans le monde entier. En 2020, plus de 50 millions de personnes dans le monde vivaient avec une démence, dont près de 60 % dans des pays à revenu faible ou intermédiaire [2]. Ce nombre va presque doubler tous les 20 ans, pour atteindre 82 millions en 2030 et 152 millions en 2050 [1]. La nécessité de traiter les MNs est urgente, car l'organisation mondiale de la santé (OMS) prévoit que, dans 20 ans, les MNs qui affectent principalement les fonctions moteurs dépasseront le cancer pour devenir la deuxième cause de décès la plus fréquente, après les maladies cardiovasculaires.

Au cours des dernières années, des efforts importants ont été consacrés au développement de traitements spécifiques pour les MNs, avec un succès limité jusqu'à présent. Les composés polyphénoliques tels que les flavonoïdes et les chalcones présents dans les légumes et les fruits auraient des propriétés anti-âge et réduiraient le risque des MNs [3, 4]. Malgré leur abondance, les recherches sur les avantages de ces composés polyphénoliques pour la santé humaine n'ont commencé que récemment. Des études précliniques et cliniques ont démontré les effets bénéfiques potentiels de flavonoïdes et de chalcones sur les neurones [5, 6].Divers modèles animaux et études de cultures cellulaires ont montré que ces composés étaient très prometteurs pour le développement de cibles thérapeutiques appropriées.

La conception de médicaments assistée par ordinateur (CADD) est devenue un aspect important du processus de découverte de médicaments. En effet, les avancées technologiques dans les domaines de la caractérisation de la structure moléculaire, de la science computationnelle et de la biologie moléculaire ont contribué à la planification de nouveaux médicaments contre les maladies neurodégénératives [7, 8].

L'objectif principal de ce travail consiste à utiliser les outils bio-informatique pour identifier des nouvelles molécules biologiquement actives destinées au traitement des maladies d'Alzheimer et de Parkinson. Pour atteindre cet objectif, le manuscrit de cette thèse est présenté en deux parties. La première partie est consacrée à une synthèse bibliographique tandis que la deuxième partie regroupe les travaux réalisés durant cette thèse.

La partie I intitulé Synthèse Bibliographique est composé de quatre chapitres :

- Le premier chapitre est une recherche bibliographique sur les maladies neurodégénératives, en particulier la maladie d'Alzheimer et la maladie de Parkinson.

Il présente les différentes enzymes responsables de ces démences ainsi que les traitements utilisés pour les soigner.

- Le deuxième chapitre présente l'intérêt des chalcones et leurs dérivés dans la recherche et développement de nouveaux agents thérapeutiques. Dans ce chapitre, nous nous sommes intéressons aux activités Anti-Alzheimer et anti-Parkinson de cette famille de molécule.
- Dans le troisième chapitre de cette partie I, nous présentons un aperçu général sur le processus de la découverte de médicaments.
- Le quatrième chapitre regroupe les différentes méthodes de la modélisation moléculaire utilisées dans la conception de médicaments assistée par ordinateur (CADD).

La partie II intitulé travaux réalisés est composé de cinq chapitres où on expose nos travaux notamment sous forme de 5 publications. Les 3 premières concernent la maladie de Parkinson tandis que la 4ème et 5ème publication traite les résultats relatifs à la maladie d'Alzheimer.

- **Chapitre I :** Étude computationnelle de dérivés cétoniques insaturés en tant qu'inhibiteurs de la MAO-B à l'aide de QSAR, ADME/Tox, docking moléculaire et simulations par dynamique moléculaire.
- **Chapitre II:** Développement de nouveaux inhibiteurs de la monoamine oxydase B (MAO-B) par la combinaison de docking, du 3D-QSAR, de la prédiction ADMET, de la simulation par dynamique moléculaire et de l'énergie libre de liaison MM_GBSA.
- **Chapitre III :** Identification des agents anti-Parkinsoniens ciblant la protéine MAO-B en utilisant la modélisation par homologie, le docking moléculaire, la simulation par dynamique moléculaire (MD) et les approches ADMET.
- **Chapitre IV :** Études de modélisation moléculaire de dérivés de la C-Glycosylfavone en tant qu'inhibiteurs de GSK-3β, basées sur l'étude QSAR et l'analyse de docking.
- **Chapitre V :** Application des méthodes CADD à la recherche et développement de nouveaux inhibiteurs de la p38a MAP kinase : 2D-QSAR, docking moléculaire, dynamique moléculaire et propriétés ADMET.

Enfin, une conclusion générale suivra les résultats de cette thèse et exposera les différentes perspectives.

Partie I: Bibliographie

Chapitre I: Maladies Neurodégénérative

I. Introduction

Avec l'augmentation de la population, en particulier des personnes âgées, le traitement des maladies neurodégénératives (MNs) (dont les symptômes sont la perte de mémoire et les troubles cognitifs dus à la perte de la fonction neuronale) est devenu un nouvel enjeu social et économique [9, 10]. La maladie d'Alzheimer (MA) et la maladie de Parkinson (MP) sont les cas les plus courants des MNs. Elles pourraient être causés par plusieurs facteurs, notamment l'accumulation de protéines amyloïdogènes mal repliées (telles que l'amyloïde-β (Aβ), la protéine *Tau* et l'α-synucléine (α-syn)), le déficit cholinergique, la dyshoméostasie des ions métalliques et le stress oxydatif induit par la surproduction d'espèces réactives de l'oxygène (ERO) [10]. En particulier, les agrégats de protéines amyloïdogènes présents dans les régions intracellulaires et extracellulaires ont fait l'objet de recherches intensives pour comprendre la pathologie de la MA et de la MP. Parmi les agrégats de protéines, les espèces oligomères solubles ont été ciblées comme étant les espèces les plus neurotoxiques [11, 12].

La MA est la forme la plus courante (environ 60 à 80 %) de démence, décrite pour la première fois par le médecin allemand Aloïs Alzheimer en 1907 [13]. Cette démence touche principalement les individus âgés de 65 ans et plus, mais aussi les personnes plus jeunes. Selon l'organisation mondiale de la santé (OMS 2022, environ 55 millions de personnes sont atteintes par la MA dans le monde et ce nombre devrait augmenter de 10 millions chaque année, si des stratégies de réduction des risques ne sont pas mises en place à travers le monde [14].

Les principales caractéristiques de la MA sont (i) le rétrécissement de l'hippocampe et du cortex dans le cerveau, (ii) les plaques séniles principalement composées d'agrégats d'Aβ, (iii) la dégénérescence neurofibrillaire (DNF) contenant les agrégats de la protéine *Tau* hyperphosphorylé (*p-Tau*), (iv) le mauvais compartimentage des ions métalliques et (v) le stress oxydatif [15-17]. Des études antérieures ont montré que, parmi ces caractéristiques, l'Aβ est un facteur de risque majeur dans l'apparition et la progression de la MA. Par conséquent, la production, l'agrégation et la clairance de l'Aβ ont été considérées comme liées à la pathogenèse de la MA. Ainsi, de multiples stratégies ont été testées pour traiter la MA soit en inhibant les β- et γ sécrétases qui sont responsables de la génération de l'Aβ en clivant la protéine précurseur de l'amyloïde (APP) [18, 19], soit en modulant l'agrégation de l'Aβ [10] ou en améliorant l'élimination de l'Aβ [20-22]. Plusieurs études ont révélé que la MA à apparition tardive a été suggérée comme étant associée à la réduction de la clairance de l'Aβ [23]. La diminution liée à l'âge ou à la pathologie de la concentration et/ou de l'activité des

enzymes dégradant l'Aβ (c'est-à-dire la néprilysine (NEP) et l'enzyme dégradant l'insuline (IDE)) pourrait être le déclencheur de l'apparition et de la progression de la MA [24].

En plus de l'Aβ, la protéine *pTau* a été considérée comme une autre cause potentielle de la MA [25]. La protéine *Tau* est fortement exprimée dans les axones neuronaux. L'hyperphosphorylation de cette protéine entraîne la formation des DNFs, qui aboutissent finalement à la formation de plaques dans le cerveau. La protéine *Tau* joue un rôle important dans le trafic neuronal et le maintien de la structure des synapses en stabilisant les microtubules et en soutenant la structure des dendrites. Cependant, lorsque la protéine est hyperphosphorylée et/ou fragmentisée, elle perd ses fonctions d'origine, ce qui entraîne des dommages neuronaux [25-26]. Des études récentes réalisées avec des souris transgéniques ont révélé que la neurotoxicité pourrait provenir des oligomères de *pTau* dans le stade précoce de la MA [27].

La deuxième forme la plus courante des MNs est la MP (plus de 8.5 millions de personnes ont été touchées par la maladie en 2019) qui se caractérise par la perte de neurones dopaminergiques avec les symptômes de tremblement, de rigidité et de lenteur des mouvements [28]. Outre les troubles physiques, des problèmes émotionnels apparaissent également, notamment la dépression et l'anxiété [29]. L'une des principales raisons de l'apparition et de la progression de la MP est la perte de neurones dopaminergiques causée par l'agrégation d'α-Syn, formant des corps de Lewy, dans le cerveau. Ensuite, les neurones endommagés subissent une mort cellulaire entraînant la perte d'astrocytes et une augmentation du nombre de microglies, notamment dans la substantia nigra [29, 30].

L'α-Syn est une protéine soluble de 14 kDa. Elle existe au niveau des terminaux présynaptiques en tant que contrôleur du trafic des vésicules synaptiques [31]. Les voies d'agrégation de l'α-Syn mal repliée sont similaires à celles de l'Aβ. Ils contiennent une phase de latence, une phase d'élongation et un état d'équilibre et forment des agrégats en interagissant à travers la région hydrophobe centrale composée de 11 acides aminés [32-34]. Contrairement à l'α-Syn monomère, ses agrégats sont observés en divers endroits des cellules, ce qui indique que les espèces d'α-Syn peuvent endommager les environnements intracellulaires et extracellulaires. Les espèces oligomériques ont été signalées comme présentant une neurotoxicité en perturbant la membrane lipidique sous forme d'oligomères Aβ annulaires [35, 36].

Pour réussir à traiter les MNs, il est nécessaire de comprendre le processus de génération, d'agrégation et de dégradation des protéines amyloïdogènes (c'est-à-dire Aβ, *Tau* et α-Syn). Afin de réduire la neurotoxicité causée par les agrégats de ces protéines, la dégradation des

protéines amyloïdogènes pourrait être une méthode thérapeutique puissante, tout comme le contrôle des voies de production ou d'agrégation des protéines amyloïdogènes

II. Maladie d'Alzheimer

II.1. Généralité

En 1906, un médecin allemand du nom d'Alois Alzheimer a présenté un cas clinique particulier lors d'un congrès médical à Tübingen, en Allemagne. Il avait examiné le cerveau d'Augusta, une femme de 51 ans décédée après avoir présenté des symptômes tels que des troubles de la mémoire, des changements de comportement et une confusion mentale.

Grâce à des techniques de coloration, l'examen microscopique du cerveau de la patiente a révélé la présence de lésions neuropathologies particulières, notamment des plaques séniles et des dégénérescences neurofibrillaires (Figure I.1). Alois Alzheimer a donc décrit cette nouvelle maladie, qu'il a appelée 'Maladie d'Alzheimer'[37].

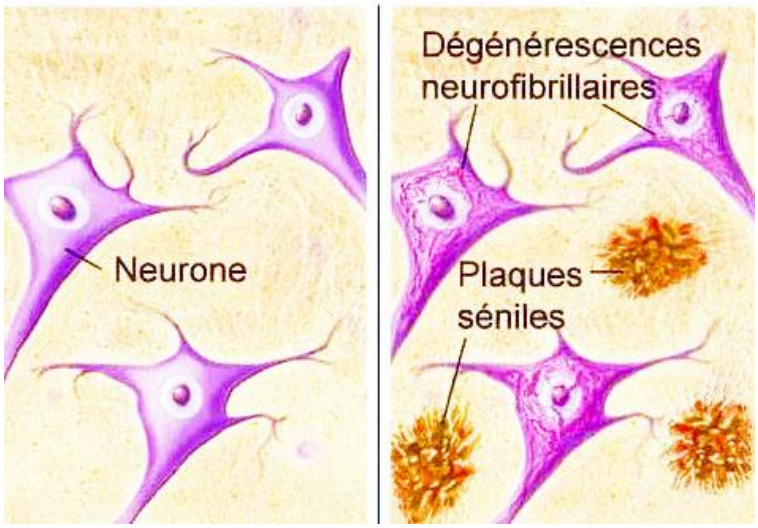


Figure I.1.*Lésions neuropathogiques de la maladie d'Alzheimer*

Cette découverte a marqué un tournant dans la compréhension des troubles cognitifs et a jeté les bases de la recherche future sur la maladie. Cependant, il convient de noter que les travaux d'Alzheimer n'ont pas permis de comprendre pleinement les causes sous-jacentes de la maladie, ni les mécanismes précis qui conduisent aux lésions observées.

Au fil des décennies, la recherche sur la maladie d'Alzheimer s'est intensifiée, notamment pour mieux comprendre les facteurs de risque, les mécanismes cellulaires et moléculaires, ainsi que pour développer des méthodes de diagnostic et des traitements potentiels.

La découverte de la MA a permis de sensibiliser à cette affection et de lancer des efforts considérables pour comprendre et traiter cette maladie neurodégénérative progressive. Aujourd'hui, la MA reste un domaine de recherche actif et un défi majeur pour la médecine et la science.

II.2. Neuropathologies

Les principaux facteurs neuropathologies de la MA sont les plaques amyloïdes extracellulaires, la dégénérescence neurofibrillaire (DNF) intracellulaire, la détérioration synaptique et la mort neuronale. L'hypothèse de la cascade amyloïde suggère que l'agrégation des plaques amyloïdes interfère avec l'activité synaptique et déclenche une série d'effets en aval qui provoquent un dysfonctionnement inter- et intra-neuronal et finalement la mort cellulaire [38].

II.2.1. Plaques amyloïdes

Les plaques amyloïdes proviennent de l'accumulation et des lésions extracellulaires de β amyloïde (Aβ), un produit de la transformation de la protéine précurseur amyloïde (ou APP) et le principal constituant des plaques séniles ou neuritiques. L'APP est une protéine transmembranaire de type 1, traitée par clivage séquentiel via β- ou α-sécrétase pour produire les fragments C-terminaux (C99 et C83), sAPPβ (voie amyloïdogène) ou sAPPα (voie non amyloïdogène), respectivement **(**Figure I. 2). Le fragment C99 est ensuite clivé par γ-sécrétase pour donner AICD (ou APP Intra Cellular Domain) et les peptides Aβ de longueur variable, qui ne serent pas dégradé par les lysosomes, ce qui produit une accumulation d'Aβ ($A\beta_{1-40}$ et $A\beta_{1-42}$). L'élévation des niveaux de peptides $A\beta_{1-42}$ conduit à une agrégation de l'amyloïde (plaques amyloïde) qui provoque une toxicité neuronale [39]. Ces dépôts extracellulaires empêchent le fonctionnement des neurones et entraîne une mauvaise transmission des messages nerveux et provoquent un effet toxique pour le cerveau dû à une entrée massive de calcium, ce qui va augmenter la production de radicaux libres responsables du stress oxydant dans le cortex cérébral [40].

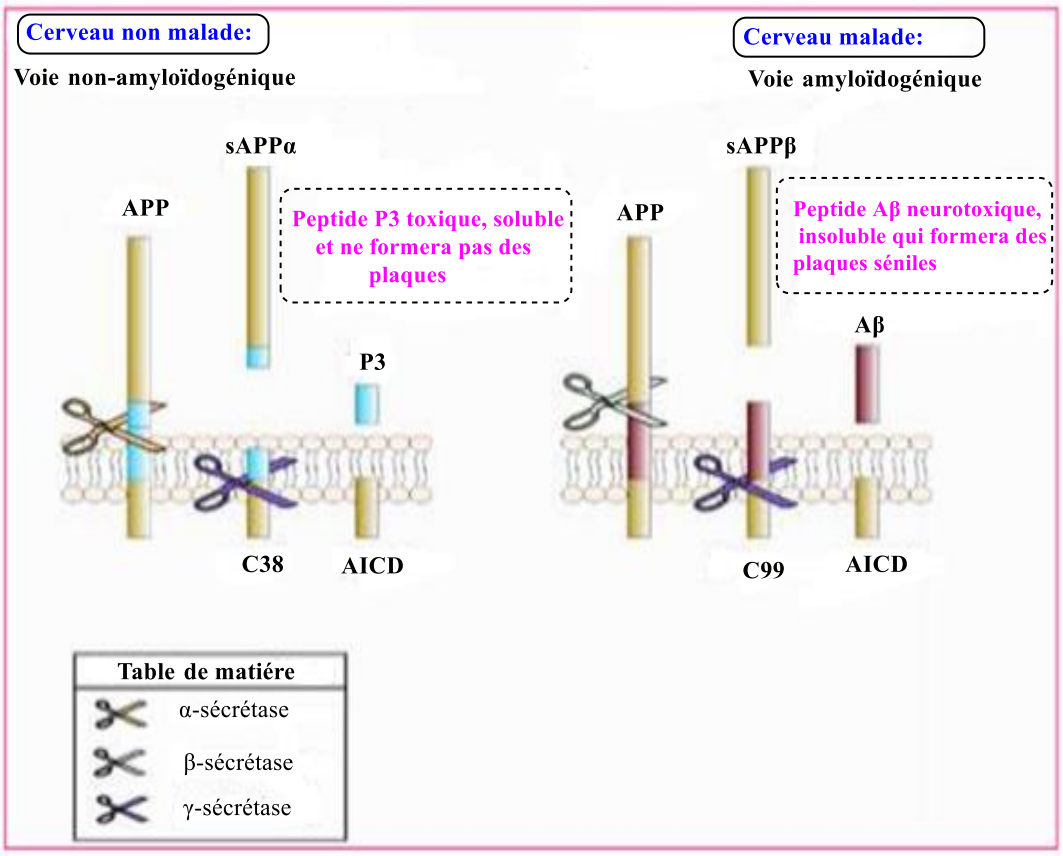


Figure I.2. *Voies impliquées dans les processus de métabolisation de l'APP*

II.2.2. Dégénérescences neurofibrillaires (DNF)

La protéine *Tau* est l'une des protéines impliquées dans l'assemblage des microtubules. Il s'agit d'un composant essentiel pour une croissance axonale et un développement neuronal. La DNF correspond à une accumulation anormale de filaments hélicoïdaux intra-neuronaux de protéine *Tau* hyper-phosphorylée [17, 41], qui se déposent préférentiellement dans les neurones du lobe temporal mésial (en particulier l'hippocampe), de la région pariétotemporale latérale et des cortex d'association frontaux. Dans le cerveau normal, l'équilibre entre phosphorylation et déphosphorylation de la protéine *Tau* module la stabilité du cytosquelette et, par conséquent, la morphologie axonale. L'hyper-phosphorylation de la protéine *Tau* provoque une diminution de l'affinité de la protéine pour tubuline, donc une déstabilisation du cytosquelette et une accumulation de protéines *Tau* sous la forme de filaments appariés en hélice [42]. Par ailleurs, il a été prouvé que l'hyper-phosphorylation est induite par l'exposition de peptides Aβ provoquait la dysrégulation et l'activation de certaines protéines kinases (GSK-3B, CDK5, ERK2) qui entrainent une hyper-phosphorylation de *Tau* [43, 44]. Il est donc plausible que *Tau*

et Aβ agissent dans des voies parallèles, causant la MA et renforçant leurs effets toxiques respectifs [45].

II.2.3. Perte de neurones et de synapses

Dans la MA, les plaques Aβ provoquent la mort des neurones du noyau basal de Meynert, ce qui entraîne une diminution de la synthèse et de la libération d'acétylcholine (Ach), ainsi qu'une augmentation de l'activité de l'acétylcholinestérase, une altération des voies de signalisation de l'acétylcholine muscarinique et une diminution de la signalisation et de la fonction cholinergiques. Ces changements perturbent indirectement l'activité des récepteurs de N-méthyl-D-aspartate (NMDA), ce qui entraîne une neurotoxicité du glutamate. Comme la plaque est concentrée dans les ganglions de la base, des parties du lobe temporal et du néocortex, la mémoire et les fonctions exécutives sont affectées. La perte des neurones du cortex médian et du locus coeruleus dans le tronc cérébral entraîne des déficits en sérotonine et en norépinéphrine, respectivement. Dans la MA, la dysphorie et l'insomnie sont causées par une activité cérébrale de l'activité sérotonergique et adrénergique cérébrale [45].

II.3. Rôle de GSK-3 dans la maladie d'Alzheimer

La glycogène synthase kinase -3 (GSK-3) est une sérine/thréonine kinase initialement découverte pour son rôle majeur dans la signalisation intracellulaire induite par l'insuline. Chez les mammifères, deux gènes de GSK-3 ont été identifiés [46], la GSK-3α et la GSK-3β. Les deux gènes sont presque identiques dans leur domaine catalytique (98% d'identité) mais divergent au niveau des régions N- et C-terminales [46]. La GSK-3β très exprimée dans les tissus neuronaux, a la particularité d'être active ; elle est négativement régulée par phosphorylation sur son résidu sérine 9. Son activité est inhibée par plusieurs stimulus extérieurs, dont l'insuline, les ligands Wnt et la voie AKt. La GSK-3β et également phosphorylée sur la tyrosine 216, probablement par autophosphorylation, ce qui augmente son activité. Récemment, la GSK-3β est l'un des principaux acteurs impliqués dans l'hyper-phosphorylation de la protéine *Tau*, qui est une protéine liée aux microtubules, impliquée dans la stabilisation du cytosquelette neuronal [47]. Lorsque la GSK-3β phosphoryle *Tau* de manière anormalement élevée, elle rend *Tau* incapable de se lier aux microtubules, ce qui rend les microtubules instables. En conséquence, la protéine *Tau* dissociée s'accumule et s'agrège dans les neurones, ce qui entraîne la formation de DNF et la mort des neurones [48]. Des études utilisant des systèmes automatisés d'imagerie par fluorescence à haut débit ont démontré qu'une augmentation anormale de l'activité de GSK-3β peut entraîner des perturbations dans

la localisation/le trafic de la β-sécrétase et une sécrétion importante d'Aβ [49]. Dans un modèle de souris où la protéine précurseur de l'amyloïde est surexprimée, la GSK-3β peut réduire l'Aβ et sa toxicité, améliorer les performances comportementales induites par l'Aβ et réduire la perte neuronale. Par conséquent, cette kinase est impliquée dans presque toutes les voies conduisant à la MA. Malgré ces résultats préliminaires, les inhibiteurs très puissants et sélectifs de la GSK-3β qui ont progressé dans les phases cliniques n'ont pas réussi à démontrer l'efficacité attendue/souhaitée [50, 51].

II.4. Rôle d'AChE dans la maladie d'Alzheimer

L'acétylcholinestérase (AChE) est une enzyme cholinergique que l'on trouve principalement dans le système nerveux et sanguin des eucaryotes supérieurs. Elle décompose ou hydrolyse immédiatement l'acétylcholine (ACh), un neurotransmetteur responsable de la transmission de l'influx nerveux d'un neurone à l'autre en passant par les synapses, en acide acétique et en choline (Figure I.3) [52]. Le rôle principal de l'AChE est de mettre fin à la transmission neuronale et à la signalisation entre les synapses pour empêcher la dispersion de l'ACh et l'activation des récepteurs voisins. L'interaction de l'AChE avec le substrat ACh entraîne la dégradation, l'hydrolyse et l'inactivation de l'ACh, puis le contrôle de la concentration d'ACh au niveau des synapses.

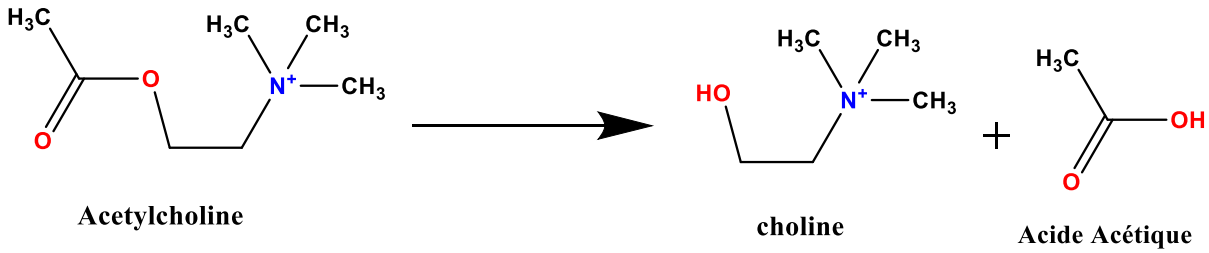


Figure I.3.*Hydrolyse de l'acétylcholine*

II.5. Rôle de BuChE dans la maladie d'Alzheimer

La butyrylcholinestérase (BuChE), également appelée pseudocholinestérase ou cholinestérase non spécifique, est une sérine hydrolase qui catalyse l'hydrolyse des esters de choline [53]. Les premières preuves de l'existence de la BuChE remontent à 1932, lorsque l'isolement d'une activité enzymatique catalysant l'hydrolyse des esters de choline a été réalisé pour la première fois [54]. La BuChE est largement distribuée dans le cerveau humain et est exprimée à des niveaux élevés dans la substance blanche [55], dans la glie [56] et dans certaines populations de neurones. L'activité enzymatique du BuChE comprend l'hydrolyse de l'acétylcholine, des polypeptides acylés à longue chaîne [57] et d'une grande variété d'autres esters [58], des amides

d'amines aromatiques [59] et la détoxification des organophosphates synthétiques hautement toxiques. Bien que la BuChE ait été associée à de nombreuses autres fonctions dans le cerveau normal et dans des maladies telles que la MA [60-63], sa capacité à hydrolyser l'acétylcholine lie cette enzyme à deux processus importants, la neurotransmission cholinergique et la modulation immunitaire [64, 65].

II.6. Traitement de la maladie d'Alzheimer

Plusieurs recherches ont mis en évidence que le traitement des émotions reste relativement préservé au cours du vieillissement. Selon l'US FDA (Food and Drug Administration) américaine, deux aspects fondamentaux sont utilisés dans le traitement de la MA : Le traitement des plaques séniles ou amyloïdes et le traitement de la transmission cholinergique.

Les anti-cholinestérases sont les premiers médicaments ayant une action démontrée dans la MA. Ils augmentent la quantité d'ACh dans la synapse, en bloquant l'AChE qui dégrade l'ACh dans la fente synaptique. Les inhibiteurs de l'acétylcholinestérase sont classés comme irréversibles ou réversibles en fonction de leur puissance et de leur durée d'action. Les inhibiteurs réversibles les plus couramment utilisés, aux différents stades de la maladie sont le Donepezil ou l'Huperzine A (Aricept®), la Tacrine (Cognex®), la Galantamine (Reminyl®) et le Rivastigmine (Exelon®) (Figure I.4).

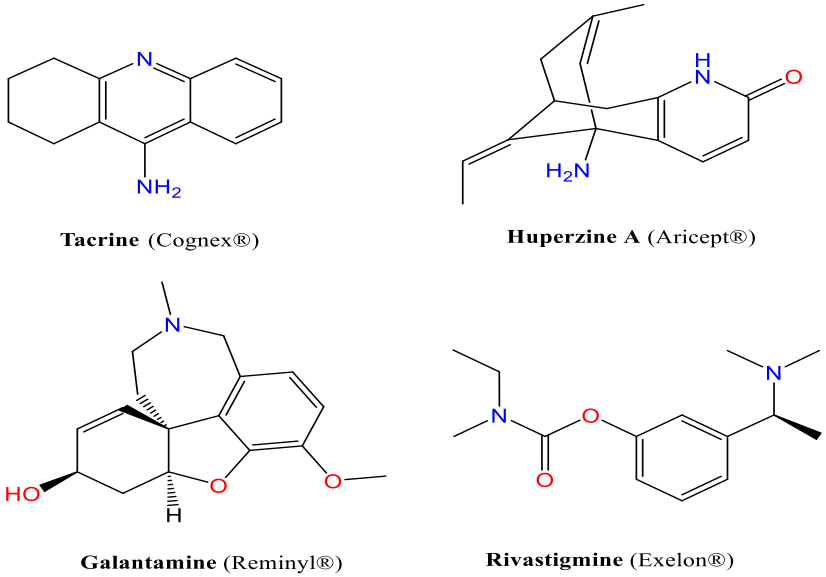


Figure I.4.*Médicaments anti-Alzheimer commercialisés*

La Tacrine est le premier médicament suggéré pour le traitement de la démence de type Alzheimer, agit principalement en inhibant l'activité enzymatique de l'AChE, mais aussi celle de la BuChE [66]. Malheureusement, comme toute prescription médicale, ce médicament,

commercialisé sous le nom Cognex®, est souvent associé à des effets indésirables. L'effet indésirable majeur de la Tacrine est son hépatotoxicité [67], qui se traduit par une élévation réversible des transaminases sanguines hépatiques, en particulier de l'alanine amino-transférase (ALAT), une faible sélectivité des effets cholinergiques périphériques [68] et une inhibition des canaux sodiques et potassiques. Les propriétés pharmacologiques de la Tacrine entraînent d'autres effets indésirables et très gênants tels que des nausées, des vomissements, des troubles gastro-intestinaux et neurologiques, etc.), ce qui fait que ce médicament a été récemment interdit et même retiré du marché aux Etats-Unis, au profit d'inhibiteurs de deuxième génération [69]. Le Donépézil (Aricept®) et la Galantamine (Reminyl®) sont des médicaments réversibles de l'AChE. Ils ciblent l'AChE et l'inhibent de façon compétitive. Le Donépézil interagit avec les cholinestérases et désactive l'activité de cette enzyme, empêchant ainsi la dégradation de l'acétylcholine [70]. Au niveau des synapses, cela conduit finalement à une augmentation du niveau d'acétylcholine [71]. Cependant, son mécanisme d'action exact n'a pas encore été exploré. Les effets indésirables les plus fréquents du Donépézil sont des troubles digestifs. La Galantamine obstrue l'AChE, qui est responsable de la dégradation du neurotransmetteur ACh. Cette suppression de l'AChE entraîne l'augmentation du neurotransmetteur ACh qui joue un rôle important dans la transmission de l'influx nerveux. Enfin, la rivastigmine (Exelon®) [72], de la famille des carbamates, est un inhibiteur de type "slow-binding" des deux formes acétylcholine et butylcholine estérases. Cette capacité à inhiber deux enzymes impliquées dans la maladie est probablement responsable de sa meilleure efficacité pour augmenter les taux d'acétylcholine, comparée au donepezil et à la Galanthamine (qui n'en inhibent qu'une seule).

De plus, les organophosphates sont des inhibiteurs irréversibles de l'AChE. Ces composés sont constitués d'un atome de phosphate sur lequel sont attachés différents groupes chimiques, qui déterminent les propriétés exactes du composé. Ils créent une lésion biochimique en phosphorylant le site estérasique de l'enzyme, s'opposant à l'hydrolyse de l'acétylcholine. La phosphorylation de l'AChE est réalisée par la formation d'une liaison covalente entre le groupe phosphate et le résidu Ser 200 du site actif. Dans un deuxième temps, une déalkylation va rendre irréversible cette phosphorylation entrainant le vieillissement de l'enzyme ce qui correspond à une inactivation quasi-définitive. (Figure I.5).

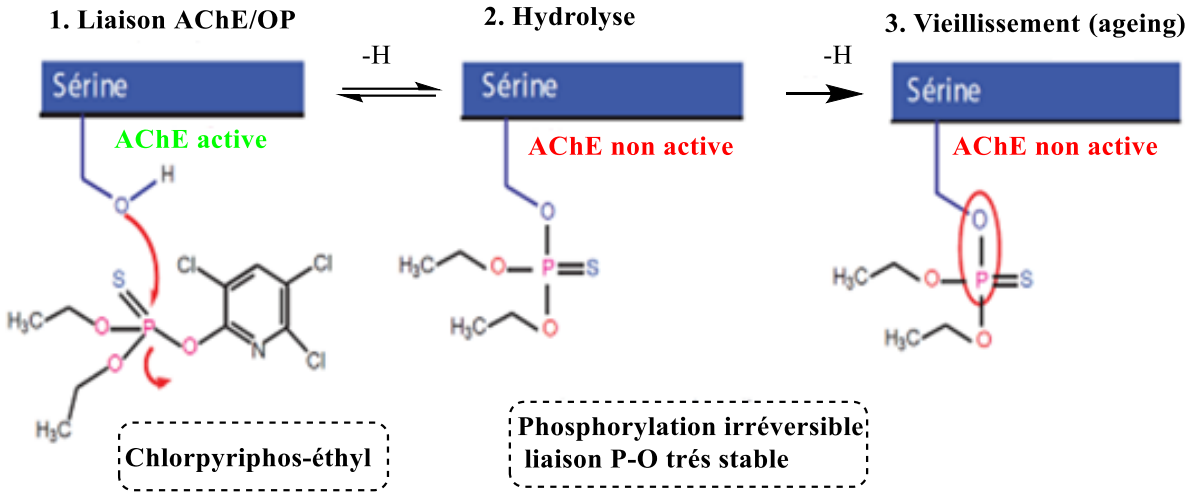


Figure.I.5. *Mécanisme de phosphorylation de l'AChE par un organophosphate*

III. Maladie de Parkinson (MP)

III.1. Histoire de la maladie de Parkinson

En 1817, le médecin londonien James Parkinson a été le premier à caractériser la maladie de Parkinson (MP) dans sa publication intitulée "An Essay on the Shaking Palsy". Parkinson a décrit plusieurs patients présentant des tremblements involontaires des membres au repos, une démarche traînante, une posture voûtée des problèmes de sommeil et de constipation. Il a noté la nature progressive de la maladie et le grand handicap qu'elle entraînait et l'a appelée *paralysis agitans* (Figure I.6). Au fil des siècles, cette pathologie ainsi que son tableau clinique seront précisés, notamment par le neurologue français Jean-Martin Charcot (1825-1893) qui lui donne le nom de « Maladie de Parkinson ».



Figure I.6. *Paralysie agitans. Image tirée de W.R. Gowers "Un manuel des maladies du système nerveux "*

La prise en charge de la maladie de Parkinson a constamment évolué au fil du temps. Dans les années 1950, Carlsson a découvert que la lévodopa inversait l'akinésie induite par la réserpine, ouvrant ainsi la voie à son utilisation comme traitement de la maladie de Parkinson. Cotzias et ses collègues ont découvert que la lévodopa améliorait les symptômes de la maladie de Parkinson en 1967, lorsqu'il a administré de la DL dihydroxyphénylalanine (DOPA) à des patients atteints de parkinsonisme et obtenu des résultats favorables. La lévodopa, associée à la carbidopa, est devenue le pilier du traitement médical et le demeure à ce jour.

Des avancées significatives ont été réalisées à la fin du XXe siècle et au XXIe siècle concernant les mécanismes physiologiques et les traitements supplémentaires de la maladie de Parkinson. Alors que la lévodopa reste la pierre angulaire de la prise en charge médicale, un lien a été établi entre la lévodopa à forte dose, la progression avancée de la maladie et l'apparition de fluctuations motrices. Les agonistes de la dopamine, les inhibiteurs de la monoamine oxydase de type B (MAO-B) et les inhibiteurs de la catéchol-Ométhyltransférase (COMT) ont trouvé leur place dans le traitement de la maladie, ainsi que de nouvelles formes de lévodopa.

III.2. Pathogénèse

La pathogénèse de la MP implique plusieurs mécanismes complexes conduisant à la disparition des neurones dopaminergiques. Il existe actuellement deux théories principales sur la pathogenèse de la MP. La première théorie suggère que le mauvais repliement et l'accumulation toxique des formes insolubles d'α-synucléine sont la clé du processus neurodégénératif. La seconde théorie suggère que le dysfonctionnement mitochondrial et le stress oxydant qui en résulte, en sont à l'origine.

III.2.1. Hypothèse de l'α-synucléine

Dans la maladie de Parkinson, les neurones dopaminergiques de la substantia nigra pars compacta (SNc) dégénèrent, ce qui entraîne une perte de dopamine dans les ganglions de la base, une zone du cerveau responsable de la coordination de la motricité fine. Ceci conduit finalement à l'apparition des symptômes cliniques de la maladie de Parkinson tels que la bradykinésie, la rigidité musculaire, les tremblements au repos et l'instabilité posturale [73]. α-synucléine s'accumule dans la MP sporadique dans les corps cellulaires neuronaux et se transforme pour former des corps de Lewy et des neurites de Lewy dans le cerveau, la moelle épinière et le système nerveux périphérique [73]. Cependant, la pathologie de Lewy est également observée dans un sous-ensemble de patients neurologiquement sains (dans une

condition appelée maladie accidentelle à corps de Lewy), ce qui indique que des facteurs supplémentaires autres que la seule pathologie de Lewy pourraient être nécessaires à la toxicité de l'α-synucléine chez les patients [74]. La duplication du locus de l'α-synucléine entraîne des formes autosomiques dominantes de la MP avec un début tardif (~60 ans), typique des patients atteints de la MP sporadique. En revanche, la triplication du locus α-synucléine entraîne une apparition précoce de la MP (<40 ans) [75], ce qui démontre que l'augmentation des taux d'α-synucléine accélère la pathogenèse de la MP. Les gènes associés à la MP, comme la leucine-rich repeat kinase (LRRK2), peuvent également développer la pathologie des corps de Lewy [76-78]. On a constaté initialement que la surexpression de LRRK2 accélérait la progression de la neuropathologie chez les souris α-synucléine A53T (tetO-A53T), alors que la déplétion de LRRK2 retardait cette progression [79]. Des études ultérieures ont démontré que le knock-out ou l'inhibition de LRRK2 empêchait la neurodégénérescence induite par l'α-synucléine dans des modèles animaux [80, 81], et que LRRK2 muté augmente les niveaux d'α-synucléine et son recrutement dans les inclusions neuronales [82]. Ceci suggère en outre des interactions fonctionnelles potentielles entre la toxicité de l'α-synucléine et d'autres gènes associés à la MP.

III2.2. Hypothèse du stress oxydant et de la mitochondrie

Un élément de l'étiologie de la MP est à l'origine de l'hypothèse selon laquelle le dysfonctionnement mitochondrial et le stress oxydant qui en résulte seraient impliqués dans la pathogénèse de la MP. Depuis l'implication de la 1-méthyl-4-phényl-1,2,3,6-tétrahydropyridine (MPTP) dans l'étiologie de la MP, il a en effet été démontré que sa neurotoxicité, caractérisée par la mort des neurones dopaminergiques, résultait d'une inhibition du complexe I (NADH-déshydrogénase) de la chaine respiratoire mitochondriale (CRM) [83]. Différents modèles ont ainsi démontré que des caractéristiques de la MP pouvaient être reproduites par inhibition du complexe I de la CRM et donc que la voie nigrostriée (la voie dopaminergique) est intrinsèque et sélectivement sensible à un dysfonctionnement du complexe I. Par la suite, une réduction de l'activité du complexe I et une diminution de la respiration mitochondriale ont été constatées dans les neurones de la substance noire (SN) de patients décédés, atteints par la MP [84].

III.3. Traitement des symptômes moteurs de la MP

Les traitements pharmacologiques des symptômes moteurs de la MP sont principalement basés sur la dopamine. Les préparations de lévodopa, les agonistes de la dopamine et les inhibiteurs de la monoamine-oxydase-B (MAO-B) sont des traitements initiaux utiles. Chez les jeunes

individus présentant des tremblements importants, les agents anticholinergiques (par exemple, le trihexyphénidyl) sont utiles, mais la prudence est de mise en raison du risque d'effets indésirables, notamment sur le plan cognitif.

III.3.1. Lévodopa

La lévodopa est le premier médicament utilisé pour traiter la MP et reste le plus puissant [85, 86]. Elle est le précurseur immédiat de la dopamine, qui peut traverser la barrière hémato-encéphalique (BHE) [87]. Elle permet aux neurones dopaminergiques de produire davantage de dopamine et de soulager les symptômes. Elle est généralement associée à la carbidopa, qui bloque le métabolisme de la lévodopa dans la périphérie, augmentant ainsi la biodisponibilité dans le système nerveux central et diminuant les effets secondaires périphériques, notamment les nausées.

III.3.2. Inhibiteurs de la catéchol-o-méthyl-transférase (COMT)

Les inhibiteurs de la COMT réduisent la dégradation de la lévodopa en 3-O-méthyldopa et augmentent la demi-vie plasmatique de la lévodopa [88]. Les inhibiteurs de la COMT sont utilisés conjointement avec la lévodopa pour améliorer le temps d'élimination en fin de dose, mais ils peuvent augmenter la dyskinésie. Les inhibiteurs de la COMT actuellement disponibles comprennent l'Entacapone et la Tolcapone, un inhibiteur de la COMT qui fait l'objet d'un avertissement de type boîte noire en raison de sa toxicité hépatique potentielle [89, 90].

III.3.3. Inhibiteurs de la monoamine oxydase de type B (MAO-B)

L'oxydation par la monoamine oxydase de type B (MAO-B) dans les cellules gliales est un mécanisme majeur d'élimination de la dopamine libérée au niveau synaptique, après la recapture présynaptique par le transporteur de dopamine (T-DA) [90]. L'inhibition de cette enzyme prolonge et augmente les concentrations synaptiques de dopamine et l'efficacité symptomatique de l'inhibition de la MAO-B à l'aide de l'inhibiteur sélectif sélégiline en tant qu'adjuvant de la Levodopa a déjà été démontrée dans les années 1970. Des études plus récentes ont établi l'efficacité antiparkinsonienne de la sélégiline en monothérapie et de la rasagiline, un nouvel inhibiteur de la MAO-B. Ce dernier est également révélé efficace lorsqu'il a été ajouté à la lévodopa chez les patients présentant des fluctuations motrices [89]. Alors que la sélégiline et la rasagiline sont toutes deux des inhibiteurs irréversibles de l'enzyme MAO-B. La safinamide, agit comme un inhibiteur réversible de la MAO-B [91].

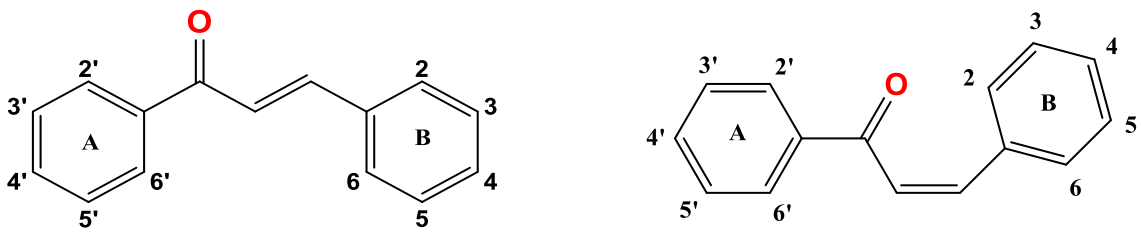
III.3.4. Agonistes de la dopamine (DA)

Les agonistes de la dopamine activent directement les récepteurs de la dopamine, ce qui atténue les symptômes de la MP. Les agonistes dopaminergiques non ergotés sont utilisés à la fois en monothérapie et en traitement d'appoint dans le traitement de la MP [89, 92]. Ils ont une demi-vie plus longue (plus de 6 heures) que la lévodopa, mais ont également une incidence plus élevée d'effets secondaires psychiatriques, notamment des hallucinations et des troubles du contrôle des impulsions, ainsi que des "crises de sommeil" potentielles (c'est-à-dire des épisodes d'endormissement soudain). Les agonistes dopaminergiques comprennent le pramipexole [92, 93], le ropinirole (disponible en formulations à libération immédiate et à libération prolongée) [94], la rotigotine (formulation transdermique) [94] et l'apomorphine à utiliser en sous-cutané (injection) comme médicament de secours pour les périodes d'arrêt aigu [95].

Chapitre II : Famille des chalcones

I. Introduction

La chalcone représentée en figure II.1 est un terme générique donné aux composés portant la structure 1,3-diphénylprop-2-én-1-one [96, 97]. Ce sont des précurseurs à chaine ouverte pour la biosynthèse des flavonoïdes et des isoflavonoïdes et se présentent principalement comme des composés polyphénoliques dont la couleur passe du jaune à l'orange. Les chalcones possèdent deux cycles aromatiques A et B reliés par un système carbonyle à trois carbones α, β-insaturés [98]. Ces composés peuvent être naturels ou synthétiques. En fonction de la disposition des substituants autour de la double liaison centrale, les chalcones existent sous la forme d'isomères trans (E, **1**) ou cis (Z, **2**). Dans la plupart des cas, l'isomère E est plus stable du point de vue thermodynamique, ce qui en fait la configuration prédominante parmi les chalcones. La configuration de l'isomère Z est instable en raison des effets stériques importants entre le groupe carbonyle et le cycle B [99].



(1) Trans (E) **(2)** Cis (Z)

Figure II.1. *Représentation structurelle et numérique de la structure de la chalcone*

II. Types de chalcones

II.1. Chalcones natureles

De nombreux composés phytochimiques sont obtenus à partir des sources végétales et sont utilisés comme agents médicinaux efficaces. Les chalcones constituent un des métabolites secondaires important obtenus à partir de nombreuses plantes comestibles comme les légumes, le thé, les épices et les produits alimentaires naturels [100].

Le tableau II.1 (voir Annexes) regroupe différents types de chalcones isolées à partir de différentes plantes. La chalcone est notamment isolée à partir des feuilles, des grains ou des racines de plante ou du fruit.

II.2. Chalcones synthétiques

Bien que les chalcones soient présentes naturellement, elles pourraient être disponibles en plus grandes quantités grâce à des méthodes de synthèses efficaces. Les chalcones peuvent en effet être préparées par deux voix de synthèse, à savoir la synthèse organique et la biosynthèse.

II.2.1. Synthèse organique

La condensation de Claisen-Schmidt est la méthode la plus utilisée pour la préparation des chalcones. Il s'agit d'une condensation aldolique, entre un aldéhyde aromatique et une cétone conduisant à la formation d'une liaison carbone-carbone donnant une énone conjuguée comme produit final. L'exemple le plus étudié de la condensation de Claisen-Schmidt est la réaction d'arylméthylcétone (acétophénone) avec l'arylaldéhyde (benzaldéhyde) en présence d'un alcali alcoolique (hydroxyde de sodium) comme catalyseur (Figure II.2).

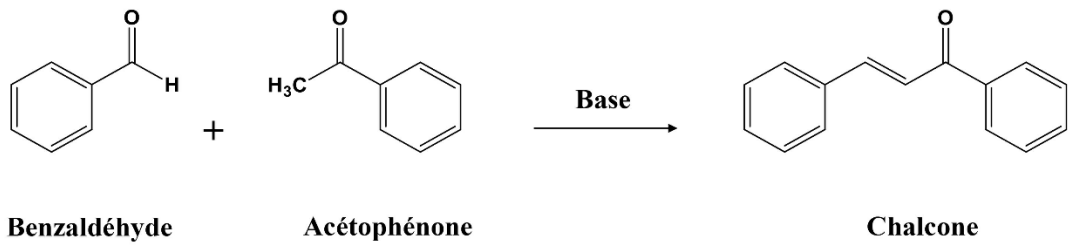


Figure II.2. *Réaction de condensation de Claisen-Schmidt*

Cette réaction entre le benzaldéhyde et l'acétophénone a servi de modèle dans le cadre de la chimie verte d'une part, et d'autre part elle s'est révélée évoluer également sans solvant, c'est-à-dire une réaction en phase solide [101,102]. Les chalcones ont également été synthétisées dans ce cadre, à partir des mêmes réactifs de départ dans l'eau à haute température.

II.2.2. Biosynthèse des chalcones

Plusieurs rapports ont été publiés sur la biosynthèse des chalcones et l'approche la plus fiable a été décrite par Andersen et Markham (2006) [103] (Figure II.3). Dans l'approche biosynthétique, la formation des chalcones se fait principalement par condensation de trois molécules d'acétate dérivées du malonyl-CoA et d'une molécule de p-coumaroyl-CoA (CoA = Coenzyme A), conduisant à un précurseur tétrakétide. Cette réaction de condensation est catalysée par l'enzyme chalcone synthase (CHS). Le tétrakétide peut être soumis à une cyclisation par deux voies différentes. La voie primaire où le tétrakétide subit une cyclisation

en présence de CHS (qui agit comme catalyseur) et donne une chalcone tri-oxygénée dans le cycle A, appelée chalconaringénine ou 2′,4′, 6′,4-tétrahydroxychalcone. Dans la seconde voie, le tétrakétide subit une réduction en présence de NADPH, puis une cyclisation en présence de CHS donne l'isoliquiritigénine (6′-déoxychalcone).

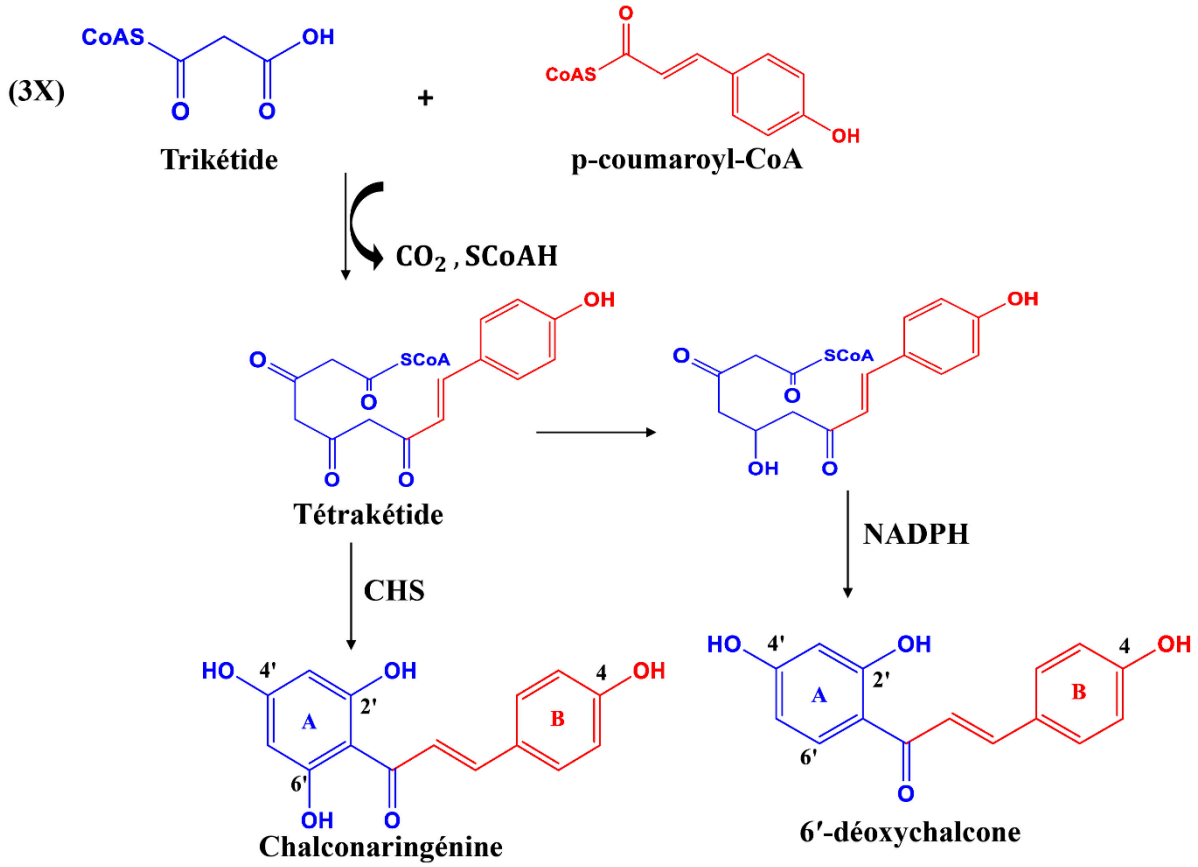


Figure II.3. *Biosynthèse des chalcones*

III. Activités pharmacologique

La chalcone et ses dérivées présentent plusieurs activités pharmacologiques favorables contre la maladie de Parkinson et contre la maladie d'Alzheimer. En effet, les chalcones ont la capacité d'inhiber plusieurs enzymes ou récepteurs ciblant ces maladies.

III.1. Activité anti-Parkinson

Dans la recherche de nouveaux inhibiteurs spécifique et sélectifs de la MAO-B, les chalcones sont apparues comme des molécules prometteuses [104] (Figure II. 4). De nombreuses études ont montré que ces molécules ont une activité inhibitrice sélective et réversible contre l'enzyme MAO-B dans la gamme nano-molaire (nM). L'introduction de divers groupements fonctionnels

sur le cycle A et le cycle B des chalcones permet d'obtenir un profil inhibiteur significatif contre la MAO-B. Par exemple, l'introduction d'hétérocycles tels que le furane, le thiophène, l'indole, l'imidazole ou des entités hétérocycliques de morpholine, ainsi que les substitutions hydroxyle, méthoxyle et chloro sur le cycle A des chalcones, permettent d'obtenir des inhibiteurs de la MAO-B puissants [105-109]. En outre, des études antérieures, ont montré que la présence de groupes donneurs d'électrons lipophiles tels que diméthylamino et éthyle en position para du cycle B des chalcones favorise l'inhibition de la MAO-B [110, 111]. De plus, la présence d'un atome d'halogène sur le cycle B permet d'obtenir des composés inhibiteurs de la MAO-B hautement sélectif [112].

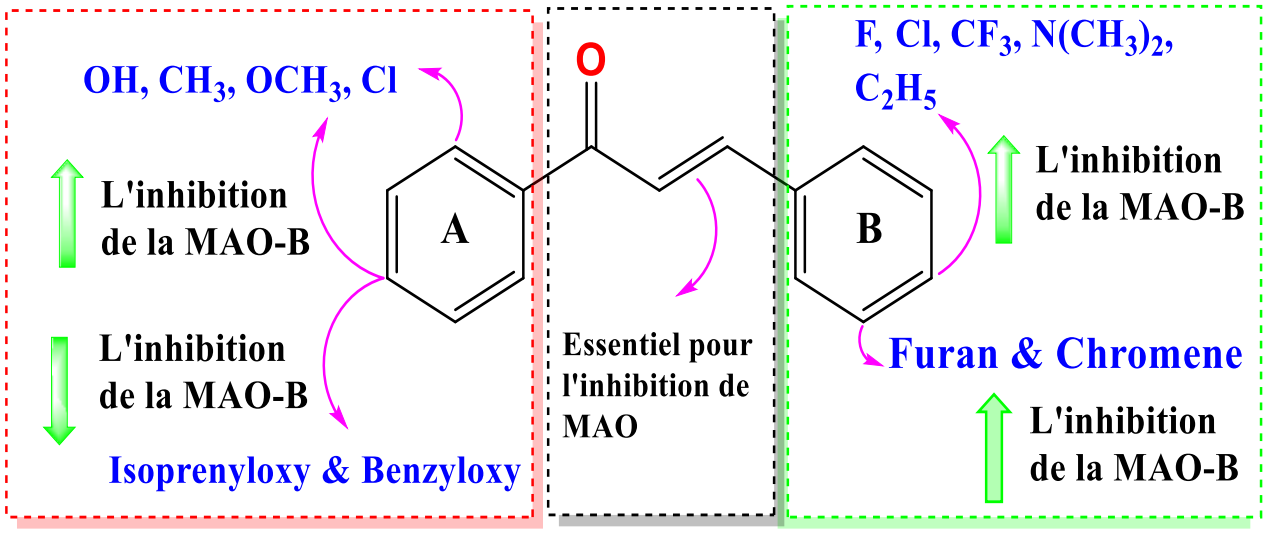


Figure II.4. Relation structure activité (RSA) des dérivés de la chalcone

III.2. Activité anti-Alzheimer

Les chalcones et leurs dérivés présentent des activités biologiques favorables contre la maladie d'Alzheimer (MA) [113-115]. Elles ont la capacité d'inhiber plusieurs enzymes ou récepteurs ciblant la MA [116, 117]. L'introduction de divers groupements fonctionnels sur le cycle A et le cycle B des chalcones confère un profil inhibiteur significatif contre les enzymes responsables de la MA (BACE-1, BuChE, AChE, etc) [118-123]. Par exemple, la présence des groupes aminoalky sur le cycle A des chalcones favorisent l'activité inhibitrice d'AChE [124,125]. L'introductions de groupes donneurs d'électrons tels que le groupement méthyle, l'hydroxyle et méthoxy sur la position para de cycle B a entraîné une inhibition modérée de l'AChE [126]. La présence des groupes lipophiles (F, CF_3, OCH_3, etc) augmente le pouvoir inhibiteur de la BACE-1 [126]. De plus, plusieurs études antérieures montrent que les dérivés

de la chalcone ont une activité puissante par rapport aux médicaments de référence (Rivastigmine, Donepizel, Galanthamine) [127, 117, 128].

III.3. Autres activités biologiques

Les nombreuses activités biologiques des chalcones (Figure II. 5), seraient dues à la présence de la fonction carbonyle α, β-éthylénique. En effet, les chalcones ont été étudiées aussi pour leurs propriétés anti-inflammatoires [97], anti-fongiques [129, 130], anti-oxydantes [131], anti-cancéreuses [132, 133], …etc. Des études bibliographiques montrent que le xanthoangelol induit l'apoptose et inhibe la prolifération tumorale et les métastases dans plusieurs lignées cellulaires cancéreuses [134, 135]. La broussochalcone A [136], le diméthylaminochalcone [137] et la cardamonine [138] présentent une activité anti-inflammatoire [139]. La licochalcone A, une substance présente dans les racines de la réglisse chinoise, a montré une activité antipaludique [140]. La 4′,4-dichlorochalcone présente une activité antilipidémique [141]. Il existe d'autres chalcones ayant une activité biologique très intéressante, comme la butein [142] et l'isoliquiritigenin [143]. Ces chalcones, comme le trans-resvératrol, activent remarquablement l'activité enzymatique des sirtuines [144], imitent les effets bénéfiques de la restriction calorique (RC), retardent le processus de vieillissement et augmentent la longévité chez *la levure bourgeonnante Saccharomyces cerevisiae*. La figure II.6 regroupe quelques analogues de la chalcone utilisés en médecine.

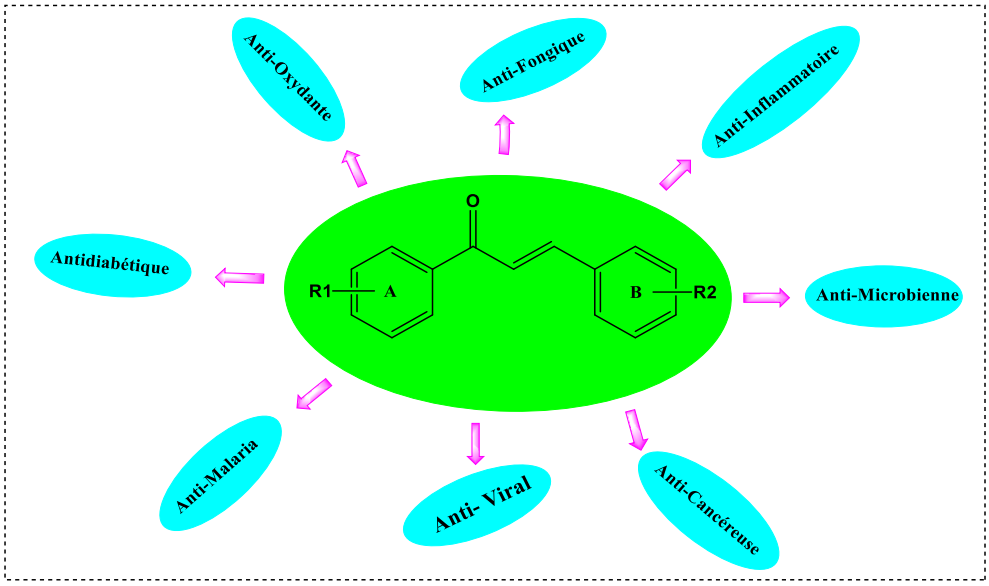


Figure II. 5. *Activités biologiques présentées par les analogues de la chalcone*

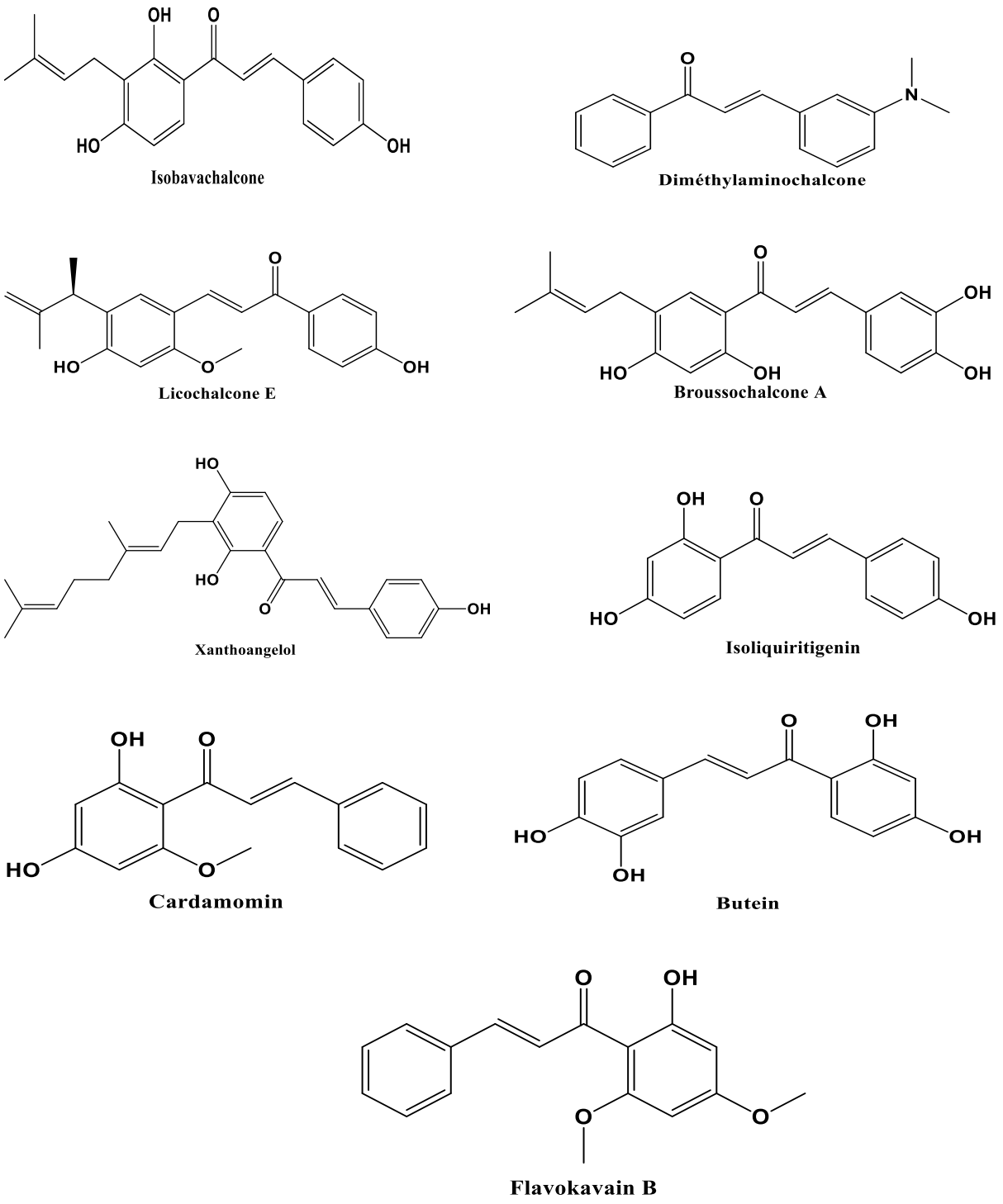


Figure II.6. *Analogues de chalcone utilisés en médecine*

Chapitre III: Processus de découverte de nouveaux médicaments

I. Histoire de la découverte des médicaments

Un médicament est, par essence, toute substance ayant des propriétés curatives ou préventives à l'égard des maladies humaines et animales. Pendant des siècles, les médicaments étaient des produits naturels, découverts par hasard, sans mécanisme d'action connu, et pas toujours synthétisables industriellement. La découverte des médicaments était essentiellement due à la sérendipité. En d'autres termes, la découverte de médicaments était un "heureux hasard".

Historiquement, l'observation des micro-organismes par van Leeuwenhoek avec le premier microscope en 1670 et leur lien avec les maladies ont permis à Pasteur d'évoquer l'idée d'un médicament en tant que molécule dotée d'un mécanisme d'action biologique. En effet, au début du vingtième siècle, Ehrlich a déclaré dans "Principe de la chimiothérapie" qu'une dose appropriée d'un produit chimique peut interférer avec la croissance des micro-organismes à l'origine d'une maladie.

Ce "principe" est devenu une réalité dans les années 1930, lorsque Fleming a découvert la pénicilline. Après cette découverte, d'importants efforts pour synthétiser ou isoler des composés similaires à la pénicilline ont conduit au développement de familles d'antibiotiques. À partir des années 1950, la compréhension du génome humain exprimé, aidée par les progrès décisifs de la biologie cellulaire et moléculaire, a permis la conception rationnelle de médicaments. Dans ce nouveau concept, les biologistes formulent des hypothèses rationnelles sur un mécanisme approprié expliquant les maladies. Ainsi, la découverte de médicaments est devenue moins fortuite et plus systématique puisque nous nous concentrons désormais sur l'association des maladies à des protéines spécifiques des voies biologiques afin de les cibler avec des composés organiques. Trouver des cibles associées aux maladies et leurs ligands est l'objectif de la découverte rationnelle de médicaments.

Enfin, à la suite de l'épidémie de SIDA dans les années 1980, un effort de recherche considérable a abouti au développement de la chimie combinatoire, permettant la synthèse des composés bioactifs présentant diverses activités pharmacologiques. Aujourd'hui, même si le hasard joue encore un rôle important dans le développement des médicaments, les progrès récents des méthodes de modélisation moléculaires dites *in silico* permettent de rationaliser le processus de découverte de médicaments.

II. Processus de recherche et développement (R&D) des médicaments

La conception rationalisée de médicaments consiste à associer une maladie à une ou plusieurs cibles moléculaires, généralement des protéines appelées cibles thérapeutiques, impliquées dans les voies de signalisation responsable dans le développement de la maladie. Ces cibles sont donc identifiées par une compréhension fondamentale de la maladie après une étude à long terme et sont généralement validées par des expériences in-vitro/in-vivo. L'objectif du processus de découverte de médicaments est alors d'identifier un petit composé (têtes de sériés) qui se lie à la protéine cible de manière à inhiber le développement de la maladie. Cette petite molécule doit également être capable d'atteindre la protéine cible dans l'organisme sans déclencher d'effets secondaires nocifs indésirables.

Même si les progrès dans les études expérimentales, le processus de découverte des médicaments reste long et coûteux. Aujourd'hui, on estimait que le processus global de développement d'un médicament nécessitait 1.8 milliard de dollars US et une période moyenne de 7.3 ans [145]. En outre, environ 53% des médicaments récemment développés n'atteignent pas la phase préclinique et principalement en raison d'effets secondaires intolérables, d'effets toxicologiques inacceptables et d'interactions médicamenteuses inattendues.

La figure III.1 présente les principales étapes du processus de découverte de médicaments.

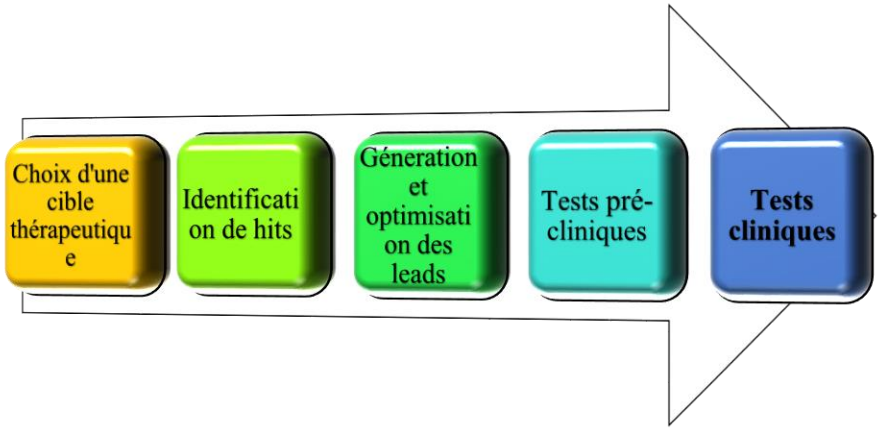


Figure III.1. *Étapes du processus de recherche et développement (R&D) des médicaments*

Les cinq principales étapes du processus de découverte d'un médicament sont présentées ci-dessous

II.1. Identification de la cible protéique

L'identification de la cible thérapeutique est une étape fondamentale dans le domaine de la recherche et du développement des médicaments. Cette étape consiste à identifier des nouvelles

cibles thérapeutiques (gènes, protéines ou processus biologiques impliqués dans une maladie) qui jouent un rôle crucial dans le développement d'une maladie, et qui pourraient être ciblées par des médicaments pour traiter cette maladie.

Généralement, toute protéine pourrait en principe être ciblée. Sur la base des médicaments actuellement disponibles, le génome médicamenteux se limite principalement à quelques familles de protéines. En effet, la moitié des cibles "humaines" connues sont des récepteurs couplés aux protéines G, une famille de protéines de la membrane cellulaire impliquées dans la transduction du signal. Les canaux ioniques, c'est-à-dire les protéines membranaires qui régulent le flux d'ions à travers les membranes, constituent le deuxième groupe le plus important de cibles médicamenteuses connues.

II.2. Identification de "hits"

Une deuxième étape consiste à identifier un composé moléculaire qui inhibe (ou active) la protéine cible, et modifie sa fonction. En général, une molécule qui se lie à une protéine et module sa fonction est appelée un ligand. Dans le contexte de l'identification de médicaments, les composés initialement identifiés comme actifs contre la cible sont appelés "hits". De plus, on parle souvent de l'interaction entre un ligand et sa cible. En termes d'activité, il s'agit d'une valeur continue quantifiant l'altération de l'activité de la cible déclenchée par le ligand, telle qu'évaluée par un test biologique. Par exemple, l'activité d'un ligand contre une cible enzymatique est mesurée comme la concentration à laquelle il diminue l'action catalytique de la cible de 50%, également appelée IC_{50}.

En pratique, les succès sont souvent identifiés par des expériences de criblage à haut débit (high-throughput screening, HTS). Il s'agit de grands essais conçus avec une lecture simple, pour permettre un criblage massif et parallèle de centaines de milliers de médicaments en une journée. Les expériences de criblage à haut débit ont été rendues possibles par des percées technologiques dans les domaines de la miniaturisation et de la robotique, qui permettent de réaliser les essais biologiques en parallèle, ainsi que par la chimie combinatoire, qui permet de générer de grandes bibliothèques moléculaires à partir de blocs élémentaires.

Dans l'industrie pharmaceutique, une campagne HTS "classique" consiste généralement en un criblage de 1 à 2 millions de composés. Elle est généralement suivie d'un essai dose-réponse pour évaluer le niveau de bioactivité des médicaments candidats en évaluant leur modulation de l'activité cible. Les "hits" identifiés sont également criblés dans un test dit orthogonal, où

une autre technologie de criblage mécaniquement différente est utilisée pour éliminer les composés qui interfèrent avec la technologie de criblage.

II.3. Recherche et optimisation de « leads »

Un bon niveau de bio-activité ne suffit pas, et les molécules candidates à l'essai clinique doivent répondre à une série de critères différents. Tout médicament potentiel doit présenter un excellent profil biologique (en particulier une faible toxicité) et d'excellentes propriétés pharmacocinétiques. Au lieu de pharmacocinétique, les experts utilisent principalement les termes Absorption, Distribution, Métabolisme, Elimination (ADME) qui caractérisent la capacité d'un hit à atteindre sa cible protéique puis à être éliminé sans danger par l'organisme. En effet, l'accumulation de substances étrangères est susceptible d'être toxique et d'entraîner des effets secondaires néfastes.

Les structures de ces molécules sont ensuite optimisées afin de répondre aux propriétés de toxicité requises sans perdre l'activité. Globalement, la transformation d'une molécule "hit" en un médicament candidat est un processus d'optimisation multidimensionnel complexe. À la fin de cette optimisation, la molécule candidate est appelée "lead".

II.4. Tests pré-cliniques

Cette phase est destinée à préparer les essais des composés chez l'homme. Elle comprend la synthèse à l'échelle, la formulation, la toxicologie et la conception des essais pré-cliniques [146]. La difficulté et le coût de la synthèse des composés sont pris en compte tout au long du processus de découverte, mais deviennent particulièrement importants à ce stade.

Cette étape a pour but de préparer la voie aux essais cliniques où le médicament candidat est administré à des humains [147]. Elle est couverte par un régime réglementaire rigoureux et de nombreuses étapes de la phase préclinique sont couvertes par des réglementations et la nécessité de travailler selon certaines directives légales.

II.5. Tests cliniques

L'étape suivante consiste à valider le médicament en trois étapes d'essais cliniques, dont les résultats sont jugés par les agences réglementaires (Food and Drug Administration -FDA- au États-Unis et l'Agence européenne pour l'évaluation des produits médicaux -EMEA- en Europe).

En phase I, le médicament est testé sur des volontaires sains (ou parfois sur des patients qui n'ont pas d'alternative thérapeutique), afin d'évaluer sa toxicité chez l'homme. La phase II vise à tester différentes doses du médicament afin d'estimer la dose thérapeutique. La phase III est un essai en double aveugle, qui vise à comparer l'efficacité du nouveau médicament à celle d'un traitement de référence. Chacune de ces trois phases nécessite généralement plusieurs années.

Au cours du processus de développement et d'élaboration d'un médicament, plusieurs approches bioinformatiques peuvent être utilisées. Ceci afin de sélectionner des composés et particulièrement lors des phases précoces en effectuant un criblage virtuel *in silico*.

III. Méthodes de criblages virtuels *in silico*

L'objectif de criblage est l'exploration de bases de composés dans le but de chercher des molécules d'intérêt thérapeutique [148]. Il regroupe un ensemble de techniques computationnelles. Le criblage virtuel est très efficace dans les processus de conception de médicaments, notamment avec les progrès en informatique et en bio-informatique [148]. Les méthodes de criblages in silico permettent un gain considérable en coût et en temps lors de la sélection des composés les plus prometteurs. Ils sont notamment très efficaces lors des phases d'identification des *hits* ou d'optimisation des *leads* [149]. Les méthodes de criblage sont très importantes car elles consistent à prédire l'affinité d'un très grand nombre de molécules collectées en chimiothèques (banque de donnée de molécules) vis-à-vis d'un site actif ciblé afin d'aboutir à un nombre réduit de molécules présentant l'activité souhaitée. Les chimiothèques criblées expérimentalement ou virtuellement doivent présenter une diversité moléculaire afin de mieux couvrir l'espace chimique et augmenter les chances d'identifier de nouvelles touches « hits ».

Enfin, les molécules doivent également avoir les propriétés physico-chimiques caractérisant un médicament comme la biodisponibilité et les propriétés ADME (Adsorption, Distribution, Métabolisme et Excrétion). Il existe des règles élémentaires comme celles de Lipinski [150], de Viber [151] et de triangle d'or [152] qui permettent de filtrer rapidement des chimiothèques et éliminer toutes les molécules qui ne répondent pas à ces règles.

Les méthodes de criblage *in silico* ou bien les approches de conception de médicaments assistée par l'ordinateur (Computer Aided Drug Design ou CADD) sont classées en deux grandes familles : La conception de médicaments basée sur la structure (Structure Based Drug Design

ou SBDD) et la conception de médicaments basée sur le ligand (Ligand Based Drug Design ou LBDD). Lorsque la structure 3D de la cible biologique est résolue expérimentalement, les méthodes basées sur la structure sont utilisées pour évaluer la capacité des ligands à établir des interactions avec la protéine cible [153]. Par contre, lorsque les valeurs d'activité pour la cible étudiée d'un ensemble de ligands sont disponibles, on utilise les méthodes basées sur le ligand [154].

Les relations structure-activité (RSA) de ligands sont donc analysées pour prédire de nouveaux composés susceptibles d'être actifs [155]. Les deux méthodes « SBDD » et « LBDD » peuvent être utilisées l'une à la suite de l'autre [156]. Lorsque ces deux types de données sont disponibles simultanément. Enfin, différentes méthodes ont été développées au sein de ces deux grandes familles (Figure III. 2).

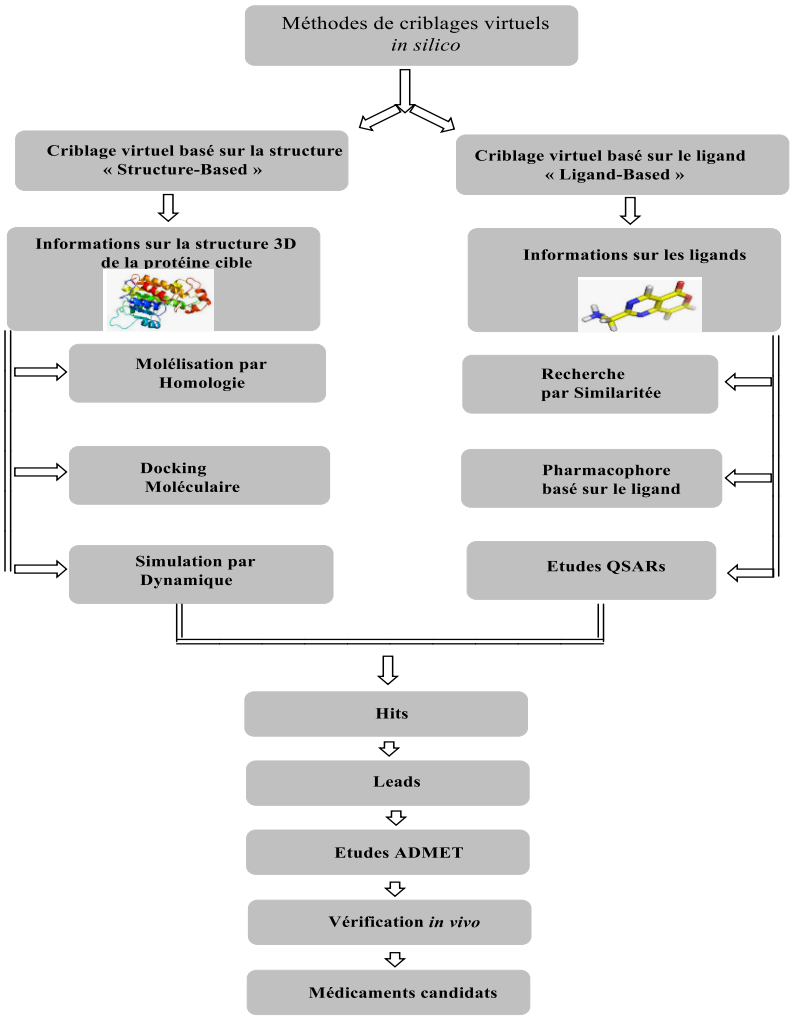


Figure III. 2. *Classifications des méthodes de criblage in silico*

III.1. Conception de médicaments basé sur le ligand « Ligand-Based»

La conception de médicaments basée sur les ligands est une approche largement utilisée dans la recherche de nouveaux médicaments candidats. Elle est employée lorsque la structure de la protéine cible est inconnue. Ces approches reposent surtout sur la connaissance préalable de ligands possédant une activité vis-à-vis de la cible thérapeutique connue. Ensuite, il sera possible d'utiliser ces ligands comme première base de « hits » dans le but d'identifier d'autres composés similaires et qui présentent des caractéristiques communes d'activité [154].

L'hypothèse que les molécules similaires vont avoir tendance à présenter des profils d'activités très proches à similaires est le principe de base de toutes les méthodes « ligand-based ». Différents types de descripteurs moléculaires pourront être calculés pour quantifier la similitude entre les composés.

Ces méthodes montrent leurs intérêts dans le processus de découverte de médicament, même en l'absence de structure expérimentale 3D du récepteur cible [157]. Parmi ces méthodes, nous pouvons citer : la recherche par similarité, le criblage de pharmacophores et les approches QSAR (Quantitative Structure Activity Relationship) [158].

III.1.1. Recherche par similarité

Le principe majeur de la chimie médicinale est l'hypothèse selon laquelle des composés structurellement similaires présentent des activités biologiques similaires. Sur la base de ce principe, des techniques de recherche par similarité ont été appliquées pour représenter les structures chimiques afin de permettre une comparaison structurelle rapide dans le but d'identifier les composés structurellement similaires ou de regrouper des collections sur la base de la similarité structurelle. Sur le plan méthodologique, le ligand de référence et la base de données des composés peuvent être représentés et comparés à l'aide d'empreintes moléculaires, une chaîne composée de chiffres binaires qui représentent la présence (1) ou l'absence (0) de fragments ou d'atomes représentatifs dans la structure chimique. Ces empreintes digitales varient considérablement en longueur et en complexité, allant de simples représentations topologiques à des arrangements pharmacophores 3D multipoints complexes [159, 160]. Ensuite, la similarité des vecteurs binaires doit être exprimée d'une manière qui peut être quantifiée. La métrique la plus couramment utilisée pour explorer la similarité est le coefficient de Tanimoto, qui est équivalent au nombre de bits communs mis à 1 dans les deux empreintes

digitales divisé par le nombre total de bits mis à 1 entre les deux empreintes. Il suppose que deux structures peuvent être considérées comme similaires si ce coefficient est supérieur à 0,85.

III.1.2. Pharmacophores basés sur le ligand

Un pharmacophore est défini comme étant un ensemble de caractéristiques stériques et électrostatiques nécessaires pour assurer d'une part les interactions avec une cible biologique spécifique. D'autres part, pour déclencher (ou bloquer) sa réponse biologique. Au cours des années, un pharmacophore a été généralement décrit comme l'orientation spatiale relative (en termes de distances, d'angles, de plans et de centroïdes) des éléments structurels et stéréoélectroniques dérivés d'un groupe de molécules qui sont responsables de la liaison à une cible commune (Figure III. 3). Les caractéristiques chimiques les plus souvent prises en compte sont les accepteurs et les donneurs de liaisons hydrogène, les groupes chargés positivement et négativement, les cycles aromatiques et les fragments aliphatiques ou aromatiques. Parfois, on inclut également des sphères de volume d'exclusion [161]. Il est dit basé sur le ligand (ou « ligand-based ») lorsqu'il est déterminé à partir de la structure de composés actifs de référence, sans connaître ou sans prendre en compte la structure de la cible biologique.

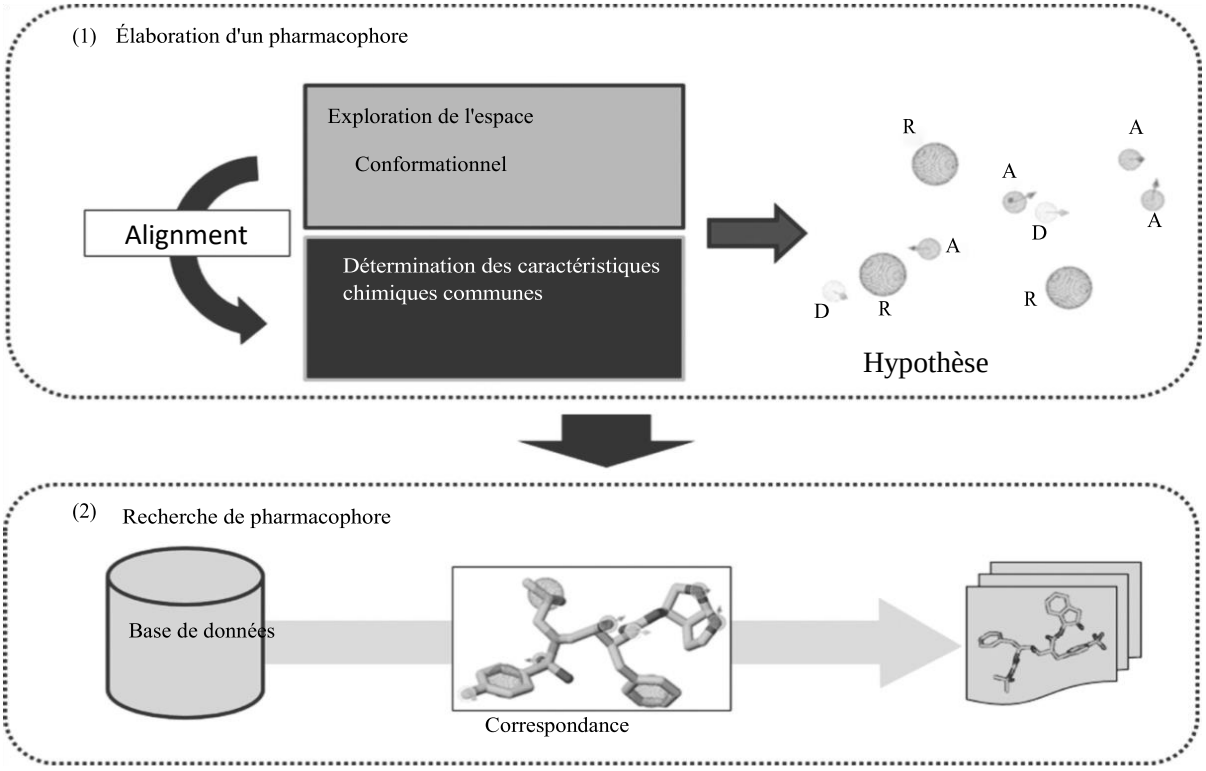


Figure III. 3. *Etapes principales du modèle pharmacophore 3D (A : accepteur de liaison hydrogène ; D : donneur de liaison hydrogène ; R : aromatique)*

La modélisation pharmacophore basée sur le ligand nécessite deux étapes principales : l'exploration de l'espace conformationnel des ligands et la détermination des caractéristiques chimiques essentielles communes qui sont responsables des interactions. Au cours de ce processus, les caractéristiques moléculaires qui ne sont pas systématiquement observées dans les composés actifs doivent être rendues facultatives ou supprimées du modèle. De plus, des contraintes spatiales peuvent être employées dans les moitiés occupées par des composés inactifs et affinées pour éviter de rendre le modèle trop restrictif. Après le raffinement du modèle, des études de validation utilisant des mesures statistiques doivent être réalisées pour déterminer la capacité du modèle à discriminer les composés actifs et inactifs connus. La figure III.4 représente un exemple de modèle pharmacophore montrant les différents points pharmacophores d'une molécule dérivée de la foranone.

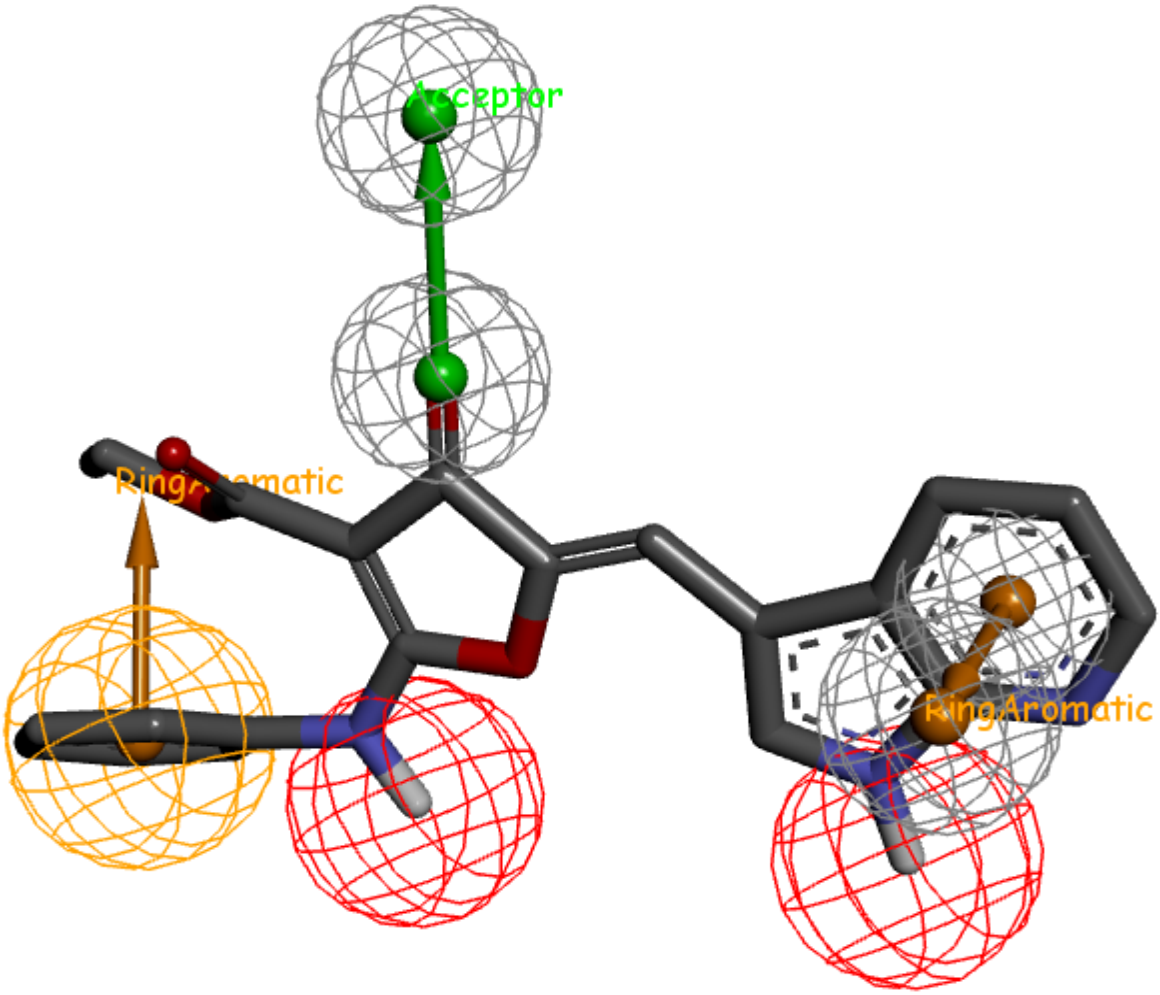


Figure III. 4. *Exemple d'un modèle pharmacophore montrant les différents points pharmacophores d'une molécule dérivée de la foranone*

Catalyst [163], MOE [164], Phase [165] et LigandScout [166] sont des outils de modélisation pharmacophore populaires appliqués à la découverte de nouvelles molécules. Bien que, le principe commun à tous ces outils de criblage est l'utilisation d'un modèle de pharmacophores comme requête pour le criblage virtuel de bases de données constituées de structures 3D de petites molécules [166].

III.1.3. Relation quantitative structure-activité (QSAR) « ligand-based »

III.1.3.1. Principe et méthodologie

Les méthodes de relations quantitatives structure-activité (QSAR) permettent d'établir une relation quantitative entre les caractéristiques structurelles (descripteurs moléculaires) des composés et leurs propriétés biologiques. Les études QSAR reposent sur le principe qui est dit que les structures similaires présentent des propriétés (activités) similaires. Pour établir un modèle QSAR, un groupe de ligands présentant l'activité biologique souhaitée est déterminé, puis une relation quantitative est établie entre les descripteurs moléculaires et l'activité biologique, selon l'équation suivante :

$$\text{Activité Biologique} = f(\text{Descripteurs Moléculaires}) \quad \textbf{Eq. III. 1}$$

Les étapes principales de la méthodologie des études QSAR sont les suivantes : Dans un premier temps, les composés dont l'activité biologique (par exemple les valeurs d'IC_{50}) sont identifiés ; ils constituent l'ensemble d'apprentissage. Ensuite, les descripteurs moléculaires (constitutionnel lD, topologiques 2D, tridimensionnelle 3D) les plus appropriés auxquels le profil biologique peut être attribué sont déterminés. Puis, les méthodes de régression qui mettent en corrélation les descripteurs moléculaires avec l'activité biologique (RLM, PLS, etc.) sont développées afin de décrire les variations de l'activité des molécules de l'ensemble d'apprentissage. Enfin, la capacité prédictive externe du modèle QSAR est déterminée par l'ensemble de test. Le détail de méthodologie d'études QSARs est présenté dans la figure III. 5 suivante :

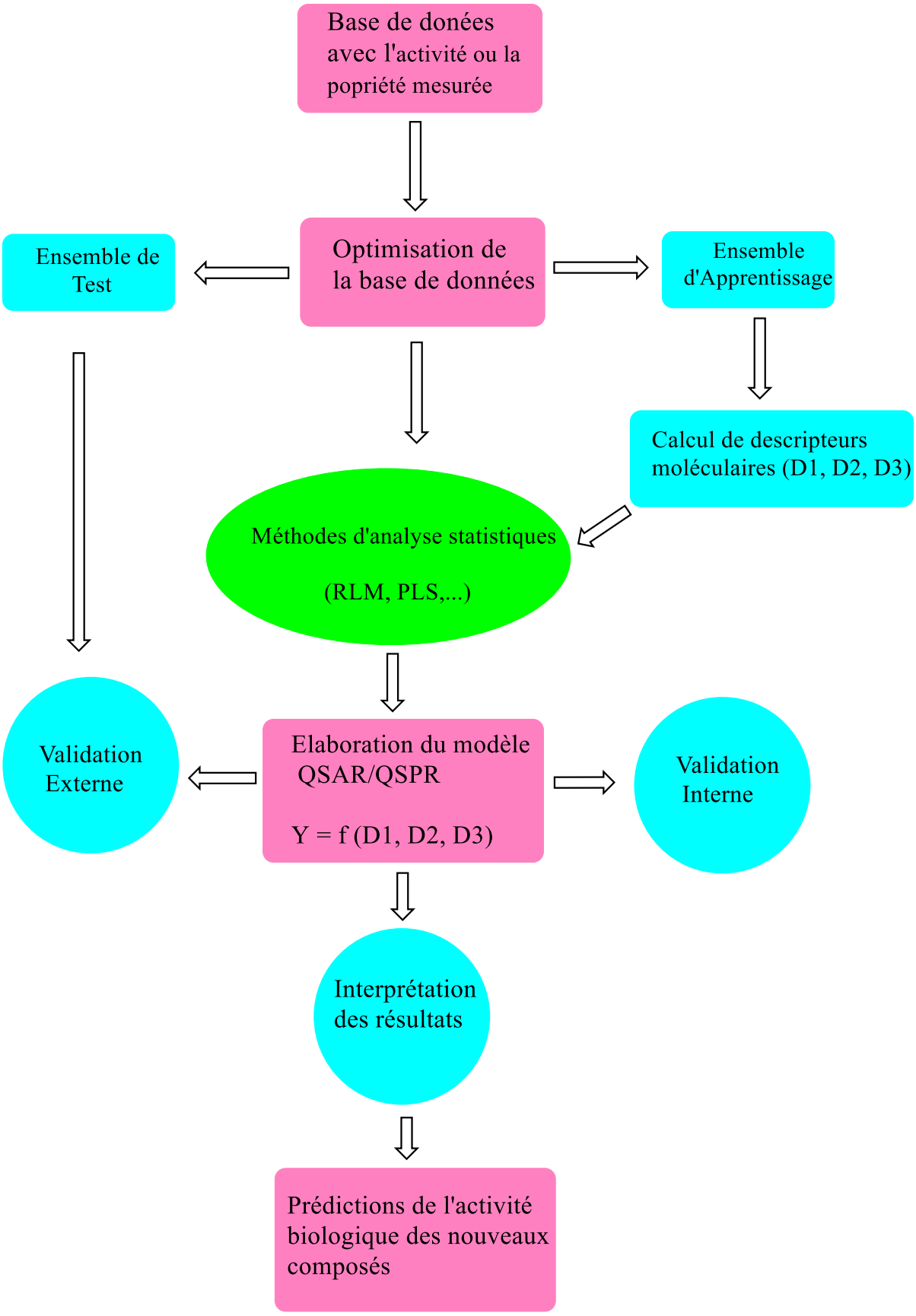


Figure III. 5. *Méthodologie générale des études QSARs*

III.1.3.2. Importance de la base de données

Le choix de la base de données est primordial dans le développement des modèles QSARs. De ce fait, le model QSAR dépend étroitement des données expérimentales de référence. Il est conseillé de considérer un nombre de molécules d'au moins 20, dans la même unité de mesure de puissance biologique (homogénéité des données) et avec le même mécanisme d'action. La gamme d'activité biologique doit être aussi large (au moins 2 unités logarithmiques) et aussi continue que possible [167]. D'autre part, les données doivent être obtenues suivant un protocole expérimental unique étant donné que les conditions expérimentales ont une forte influence sur les valeurs obtenues. La définition de la propriété en termes de conditions expérimentales constitue un point important de la démarche. Finalement, l'ensemble de données doit être divisé en un ensemble d'apprentissage et un ensemble de test.

III.1.3.3. Méthodes d'optimisation

L'étude QSAR ou bien les méthodes de la modélisation moléculaires impliquent l'utilisation des méthodes de calcul théoriques (mécanique quantique, *ab-initio* ou semi-empirique, mécanique moléculaire, dynamique moléculaire…). Ils permettent d'une part de déterminer la représentation graphique de la géométrie ou de la configuration des atomes d'une molécule. D'autre part, ils permettent d'évaluer les propriétés physico-chimiques de la molécule étudiée.

III.1.3.4. Étude 2D-QSAR

Les relations quantitatives structure-activité 2D (2D-QSAR) sont des modèles qui relient l'activité ou propriété biologique aux caractéristiques structurelles (descripteurs moléculaires) calculées dans l'espace 2D pour un groupe de molécule. Les données utilisées pour établir l'équation 2D-QSAR sont représentées par un ensemble de valeurs quantitatives dans lesquelles chaque ligne représente un composé et chaque colonne représente un descripteur moléculaire calculé.

a. Types de descripteurs

Les descripteurs peuvent être classés de plusieurs façons. En général, il existe plusieurs types de descripteurs comme les descripteurs topologiques, les descripteurs électroniques et les descripteurs quantiques. Il est intéressant de souligner que la majorité des chercheurs QSAR préfèrent classer les types de descripteurs en fonction de leurs dimensions (Figure III. 6).

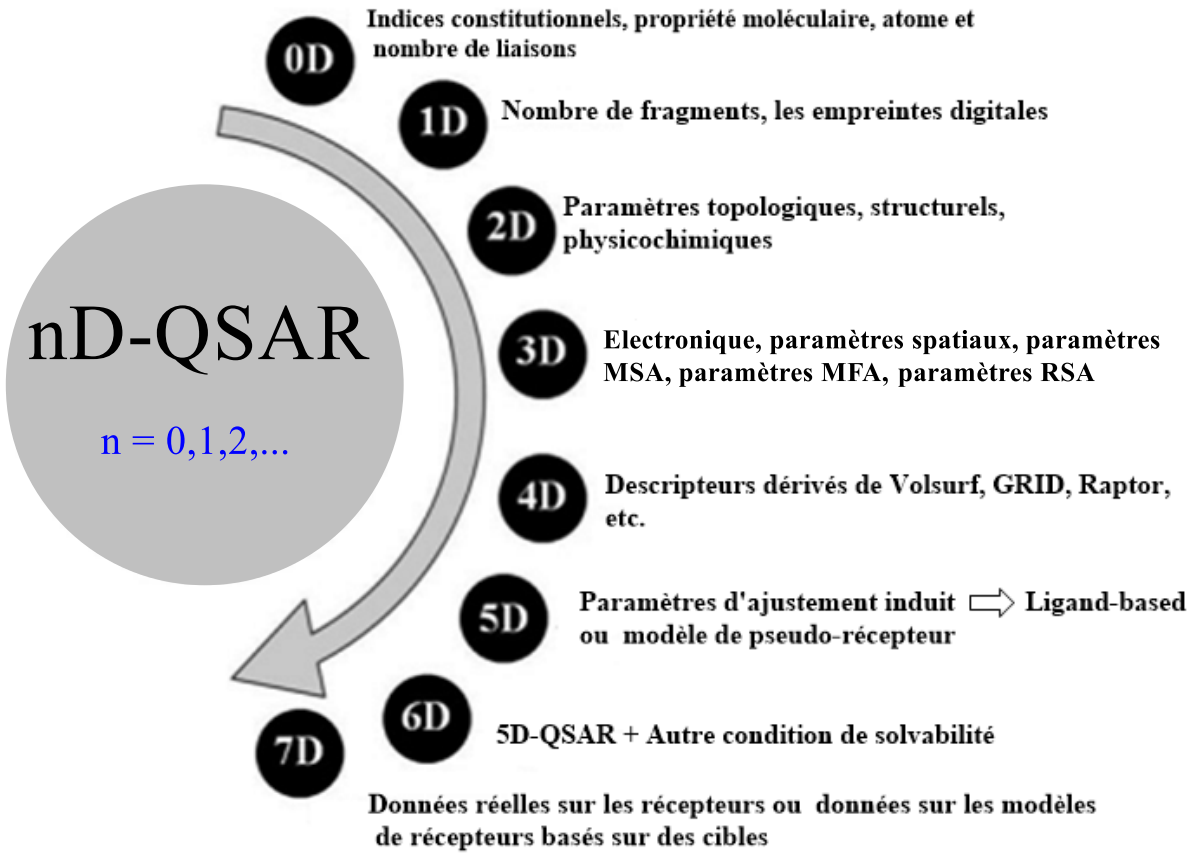


Figure III.6. Perspective dimensionnelle de la technique QSAR

b. Analyses statistiques et formation du modèle 2D-QSAR

b.1. Analyse en composantes principales (ACP)

L'Analyse en composantes principales (ACP) est une technique d'analyse de données quantitatives utilisée pour résumer l'ensemble des informations encodées dans les structures des composés [168]. Cette méthode consiste à transformer un ensemble de variables corrélées en un nouvel ensemble de variables, appelées composantes principales, qui sont moins nombreuses mais indépendantes. La figure III. 7, représente graphiquement le principe de la méthodologie pour 3 variables X1, X2 et X3 et deux composantes principales CP1 et CP2.

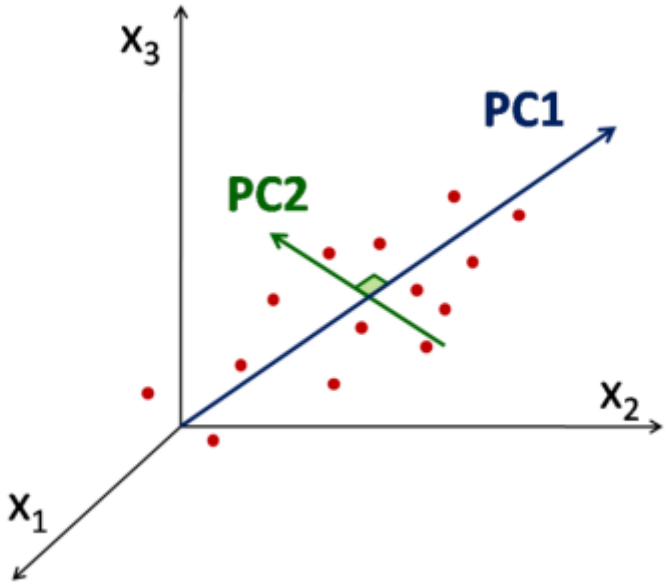


Figure III.7. *Principe de mise en place de deux composantes principales CP1 et CP2 dans un espace à 3D (x1, x2, x3) lors d'une analyse ACP*

Lors d'une étude QSAR ; la méthode ACP est utilisée pour réduire le nombre de descripteurs sans perdre d'information. Les **n** individus appelés variables dépendantes sont dans un premier temps représentés dans un espace à **m** dimensions (variables indépendantes). Le principe de la méthode consiste à obtenir une représentation approchée du nuage des **n** individus dans un sous-espace de plus faible dimensionnalité. Ces nouvelles dimensions (PCs) sont une combinaison linéaire des dimensions initiales (descripteurs moléculaires). Elles sont créées de telle façon que le maximum d'information soit présent sur la première dimension (1er axe factoriel PC1). A signaler que les axes factoriels suivants sont choisis de la même façon avec comme contrainte d'être orthogonaux aux autres axes.

b.2. Régression linéaire multiple (RLM)

La régression linéaire multiple (RLM) est l'une des méthodes de modélisation la plus fondamentale et la plus courante pour la conception d'un modèle QSAR. La méthode RLM est privilégiée pour sa simplicité et sa facilité d'interprétation car le modèle suppose une relation linéaire entre une variable dépendante, Y (l'activité/propriété biologique) et une série de n variables indépendantes, Xi (descripteurs moléculaires). Ainsi, avec la notion de Xi, l'activité d'un composé inconnu peut être prédite par le modèle ajusté. L'équation suivante représente une expression générale d'un modèle RLM :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots .. + \beta_k X_k \qquad \textbf{Eq. III. 2}$$

Où β_0 est la constante du modèle, X_1, ..., X_k sont des descripteurs moléculaires avec leurs coefficients correspondants $\beta 1$,..., β_k (pour les descripteurs moléculaires 1 à k). Ces coefficients peuvent être obtenus par l'utilisation d'estimateurs comme la méthode des moindres carrés qui minimise la somme des résidus au carré.

Les coefficients dans le modèle RLM établi représentent le degré d'influence des descripteurs sur l'activité biologique. Un coefficient positif suggère que le descripteur moléculaire correspondant contribue positivement à l'activité biologique. Par contre, une contribution négative est reflétée par un coefficient négatif. Ces interprétations peuvent cependant ne pas être exactes du fait que les descripteurs colinéaires ont le potentiel d'influencer les coefficients de telle sorte que des valeurs erronées peuvent être attribuées. Il est donc primordial que les descripteurs moléculaires du modèle soient indépendants les uns des autres et le nombre d'observations pour la construction du modèle doit être au moins cinq fois supérieur au nombre de descripteurs utilisés [169].

La qualité d'un modèle RLM est déterminée à partir d'un certain nombre de paramètres comme décrit ci-dessous :

- **Coefficient de détermination (R^2) :**

Le coefficient de détermination représente le pourcentage de l'information retenu par le modèle développé. Ce coefficient est calculé par la formule :

$$R^2 = 1 - \frac{\sum(Y_{i\,exp} - Y_{i\,pred})^2}{\sum(Y_{i\,exp} - \bar{Y}_i\,exp)^2} \qquad \textbf{Eq. III. 3}$$

Où $Y_{i\,exp}$ et $Y_{i\,pred}$ sont, respectivement, les valeurs expérimentales et prédites de la variable dépendante ; $\bar{Y}_i\,exp$ est la valeur moyenne des valeurs expérimentales. Un bon coefficient est caractérisé un score proche de 1.

- **Coefficient de détermination ajusté (R_a^2)**

Une valeur élevée de R^2 n'est pas nécessairement l'indication d'un bon modèle statistique qui s'adapte bien aux données disponibles. En effet, l'augmentation du nombre de descripteurs dans le modèle entraîne une augmentation de la valeur de R^2, mais ceci conduit à une diminution du degré de liberté et à une faible fiabilité statistique. Pour mieux refléter la variance expliquée (la fraction de la variance des données expliquée par le modèle), on utilise le $R^2_{\text{ajusté}}$ qui a été défini de la manière suivante :

$$R_a^2 = \frac{(n-1) \times R^2 - p}{n-1-p} \qquad \textbf{Eq. III. 4}$$

Dans l'expression ci-dessus, n est le nombre total de composés dans l'ensemble d'apprentissage et p est le nombre de descripteurs dans le modèle généré.

- **Racine de l'erreur quadratique moyenne (REQM))**

Pour un bon modèle, la racine de l'erreur quadratique moyenne (REQM) doit être faible et celle-ci est définie comme suit :

$$REQM = \sqrt{\frac{\sum_i (Y_i \text{ pred} - Y_i \text{ exp})^2}{n}} \qquad \textbf{Eq. III. 5}$$

Où $y_{i\ exp}$ et $y_{i\ pred}$ sont les valeurs expérimentales et calculées pour l'échantillon de molécule i, respectivement.

- **Rapport de variance (F)**

Pour évaluer la signification globale des modèles QSAR élaborés, le rapport de variance (le rapport entre le carré moyen de la régression et le carré moyen des déviations) peut être défini comme suit :

$$F = \frac{\frac{\sum\left(Y_{i\,pred} - \bar{Y}_i\ mean\right)^2}{p}}{\frac{\sum\left(Y_{exp} - Y_{pred}\right)^2}{n-1-p}} \qquad \textbf{Eq. III. 6}$$

La valeur F a deux degrés de liberté : p et N - p - 1. La valeur F calculée d'un modèle doit être significative à $p < 0,05$. Pour la signification globale des coefficients de régression, la valeur F doit être élevée.

b.3. Méthode des moindres carrés partiels (PLS)

La régression des moindres carrés partiels (ou PLS) [40, 41] est une généralisation de la méthode RLM [172]. Cette méthode a été introduite dans les années soixante par le statisticien Herman Wold. Elle est utilisée lorsque le nombre de descripteurs dépasse largement le nombre de molécules et prend en considération la corrélation entre l'activité biologique est les descripteurs moléculaires. Cette méthode utilise à la fois des principes de l'ACP et de la RLM. Elle consiste à remplacer l'espace initial des variables par un espace de plus faible dimension, sous-entendu par un petit nombre de variables synthétiques appelées « facteurs latentes», construites de façon itérative. Les variables retenues sont orthogonales (non corrélées), et sont

des combinaisons linéaires des variables initiales. L'équation et la qualité du modèle PLS peut être exprimée sous la même forme que dans la méthode RLM.

b.4. Régression non linéaire multiple (RNLM)

La régression non linéaire multiple (RNLM) est une méthode de régression non linéaire (exponentielle, logarithmique, polynomiale, ...) qui consiste à ajuster un modèle mathématique, reliant l'activité ou propriété biologique aux descripteurs moléculaires [173]. Dans l'étude QSAR, la fonction polynomiale est largement utilisée en s'appuyant sur les descripteurs proposés par le modèle linéaire qui seront élevés à la puissance 2 selon l'équation suivante :

$$y = a_0 + \sum_{i=1}^{n} a_i X_i + b_i X_i^2 \qquad \textbf{Eq. III. 7}$$

Où y est la variable à modéliser (l'activité ou propriété biologique), xi sont les descripteurs moléculaires (les variables explicatives), n est le nombre de composés dans l'ensemble d'apprentissage (les observations) ; a_0 est la constante de l'équation du modèle, a_i et b_i sont les coefficients de descripteurs dans l'équation du modèle.

b.5. Réseaux de neurones artificiels (RNA)

Les réseaux de neurones artificiels (ou RNAs) s'inspirent du système nerveux humains [174]. Des neurones interconnectés constituent un système permettant de traiter les informations données au réseau. Dans l'étude QSAR les RNAs associent des descripteurs à l'activité ou propriété par le biais d'une série de connexions pondérées, ou synapses, entre les neurones (Figure III.8). Les synapses transportent et manipulent l'information via une fonction de transfert souvent non linéaire [175]. Les réseaux de neurones de base comportent trois couches : la couche d'entrée, la couche cachée et une couche de sortie. Les réseaux de neurones sont formés de manière itérative, chaque période de formation étant appelée une époque. Leur formation implique la minimisation de l'erreur entre l'activité connue et l'activité prédite par le réseau de neurones. Comme les réseaux de neurones sont hautement non linéaires et ont la capacité de minimiser les erreurs au fil du temps, ils constituent des modèles incroyablement puissants, mais il est très facile de les former [176]. Des méthodologies telles que l'arrêt précoce sont utilisées pour éviter le surentraînement [177]. Dans ce cas, l'ensemble de données QSAR original est divisé en un ensemble d'apprentissage et un ensemble de validation ; le réseau de neurones est formé sur l'ensemble de d'apprentissage, la performance est évaluée en fonction de l'erreur quadratique moyenne, et lorsque la performance ne s'améliore plus, la formation du réseau de neurones est arrêtée. Outre la susceptibilité au surentraînement, la nature complexe

des réseaux de neurones rend l'interprétation très difficile. Il est possible de réajuster les poids des neurones de telle sorte à minimiser l'erreur [178].

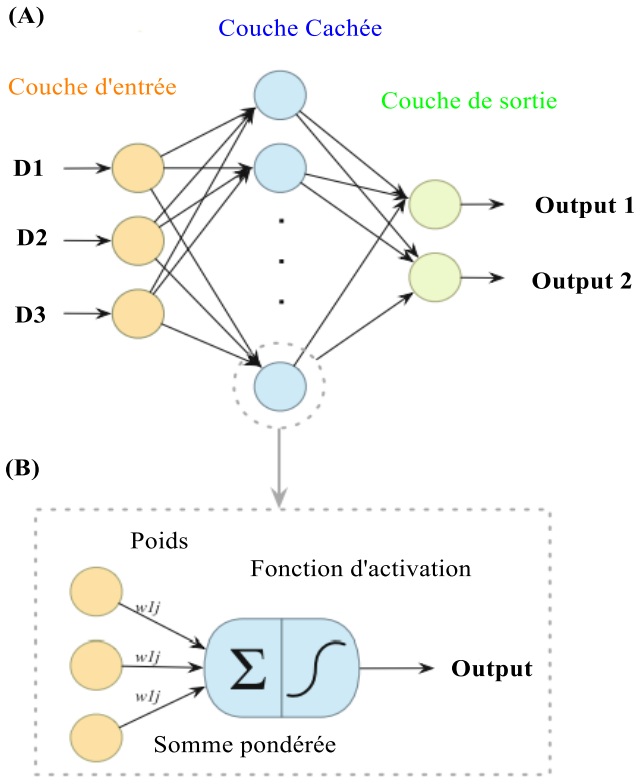


Figure III.8. *Illustration de l'architecture d'un réseau de neurones à trois couches* ***(A)*** *et le fonctionnement interne des neurones dans une couche cachée* ***(B)***

c.Validation du modèle QSAR

c.1. Validation interne

- ✓ **Validation croisée (CV)**

La validation interne d'un modèle QSAR est effectuée sur la base des molécules utilisées dans le modèle développé [179]. Cette méthode consiste à exclure un nombre donné de composés de l'ensemble d'apprentissage et à reconstruire le modèle, qui est ensuite utilisé pour prédire les composés exclus. Si un seul composé est exclu, la technique est appelée LOO (leave-one-out), sinon LSO (leave-several-out). Un coefficient de détermination LOO ou LSO, généralement indiqué par q^2, est calculé en évaluant la précision de la prédiction de ces composés. La formule q^2 est similaire au coefficient de détermination (R^2), mais les valeurs prédites proviennent de modèles construits sans le composé prédit.

$$q^2 = 1 - \left[\frac{\sum_i (Y_i\ \text{obs} - Y_i \text{pred})^2}{\sum_i (Y_i\ \text{obs} - \overline{Y}_i\ mean)^2}\right] \qquad \textbf{Eq. III. 8}$$

Le nombre d'itérations dépend des nombres de molécules dans l'ensemble d'apprentissage et le modèle correspondant devraient être raisonnablement si l'on obtient un coefficient moyen élevé après validation croisée, c'est-à-dire au moins supérieur à 0.5.

De nombreuses discussions ont impliqué la communauté scientifique sur la fiabilité de q^2 pour évaluer la robustesse du modèle et fournir une évaluation préliminaire de la prévisibilité du modèle. Il est souvent considéré que la procédure LOO peut être trop optimiste dans l'évaluation des performances du modèle, de sorte que la procédure LSO devrait également être employée, car elle augmente la perturbation infligée à l'ensemble de données. En particulier avec de grands ensembles de données, la méthode LOO pourrait donner une perturbation trop faible aux données, de sorte que le coefficient q^2_{LOO} pourrait se rapprocher du R^2 dans l'ajustement [179]. De plus, d'autres auteurs [180] ont affirmé que les paramètres validés par recoupement ne sont pas suffisants pour donner une idée sur prédictivité du modèle sur les composés externes.

- ✓ **Test de randomisation en Y**

La randomisation en Y est effectuée afin d'assurer la robustesse du modèle QSAR développé [181]. Dans le test de randomisation en Y, la validation est effectuée en permutant les valeurs de réponse (Y) par rapport à la matrice X qui est restée inchangée. Après avoir répété ce test plusieurs fois, la valeur moyenne au carré du modèle randomisé (R_r^2) devrait être faible, ce qui confirme que le modèle QSAR généré est robuste et n'est pas dû à une corrélation fortuite. Si un score élevé est obtenu, le modèle original n'est pas acceptable, car il est dû à une corrélation fortuite ou à une redondance structurelle dans l'ensemble d'apprentissage. L'écart entre les valeurs du coefficient de corrélation moyen au carré du modèle randomisé (R_r^2) et le coefficient de corrélation au carré du modèle non aléatoire (R^2) se reflète dans la valeur du paramètre ${}^cR_p^2$, qui est calculé à partir de l'équation suivante :

$$ {}^cR_p^2 = R * \sqrt{R^2 - (Average\ R_r)^2} \qquad \textbf{Eq. III. 9} $$

Où, R_r est la valeur moyenne de 'R' des modèles aléatoires. La valeur ${}^cRp^2$ doit être supérieure à 0.5 pour passer ce test.

c.2. Validation externe

La validation externe d'un modèle QSAR est évaluée par un ensemble de composés qui ne sont pas employés dans le processus de développement du modèle [182]. Elle assure la prévisibilité

et l'applicabilité du modèle QSAR développé pour la prédiction de molécules non testées [183]. Le pouvoir de prédiction peut être mieux caractérisé en évaluant (1) le coefficient de détermination de la validation externe entre les valeurs expérimentales et prédites ; (2) l'ordonnée à l'origine des lignes de régression, c'est-à-dire k et k′, la première associée aux valeurs prédites par rapport aux valeurs observées et la seconde au cas inverse ; (3) les coefficients de détermination R^2_0 et R'^2_0. Ces derniers coefficients caractérisent la régression linéaire avec l'ordonnée à l'origine Y fixée à 0, le premier étant associé aux valeurs prédites par rapport aux valeurs observées, le second étant lié aux valeurs observées par rapport aux valeurs prédites et (4) les pentes k et k′ des lignes de régression, reliant respectivement les valeurs prédites par rapport aux valeurs observées et les valeurs observées par rapport aux valeurs prédites.

Les conditions suivantes, selon le travail proposé par Golbraikh et al. (2003) [180], ont été adoptées pour conclure qu'un modèle QSAR a un pouvoir de prédiction acceptable si :

1. $q^2 > 0.5$
2. $R^2_{Test} > 0.6$
3. $\frac{R^2 - R_0^2}{R^2} < 0.1$ et $0.85 < k < 1.15$ ou $\frac{R^2 - {R_0'}^2}{R^2} < 0.1$ et $0.85 < k' < 1.15$
4. $\left|R^2 - {R_0'}^2\right| < 0.3$

La raison d'utiliser R^2_0 et d'exiger des valeurs de k proches de 1 est que lorsque les activités réelles par rapport aux activités prédites sont comparées, un ajustement exact est requis, et pas seulement une corrélation. Il est préférable que la troisième condition soit satisfaite en utilisant à la fois les coefficients R^2_0 et R'^2_0, mais cette exigence est extrêmement stricte comme le confirme l'ensemble des travaux de DEMETRA, et la condition 4, moins stricte, a été adoptée. Il faut souligner que tous les paramètres de validation pertinents, R'^2_0, R^2_0, k et q^2 ont été vérifiés pour les ensembles d'apprentissage et de test.

d. Domaine d'applicabilité

L'utilité d'un modèle QSAR réside dans sa capacité à prédire avec précision de nouveaux produits chimiques. Ainsi, une fois le modèle QSAR construit, son domaine d'applicabilité (DA) doit être défini. Un modèle n'est considéré comme valide que s'il est capable de faire des prédictions dans son domaine d'apprentissage et seules les prédictions pour de nouveaux composés relevant de son domaine d'applicabilité peuvent être considérées comme fiables s'ils sont à l'intérieur de domaine d'applicabilité. La méthode la plus courante pour définir le DA

est basée sur la détermination de la valeur de levier de chaque composé [25]. Le diagramme de Williams représenté par les résidus standardisés en fonction des valeurs de levier (hi)) est utilisé dans la présente étude pour visualiser le DA du modèle QSAR.

$$h_i = x_i^T (X^T X)^{-1} x_i \qquad (i = 1, \ldots, n), \qquad \textbf{Eq. III. 10}$$

Où xi est le vecteur descripteur du composé considéré, X est la matrice des descripteurs dérivée des valeurs des descripteurs de l'ensemble d'apprentissage, le seuil est défini comme suit :

$$h^* = \frac{3(d+1)}{n} \qquad \textbf{Eq. III. 11}$$

Où n est le nombre de composés dans l'ensemble d'apprentissage, d est le nombre de descripteurs dans le modèle proposé, un effet de levier (h) supérieur au seuil (h*) indique que la réponse prédite est une extrapolation du modèle et, par conséquent, qu'elle peut être peu fiable.

III.1.3.5. Etudes 3D-QSAR

Les relations quantitatives structure-activité tridimensionnelles (3D-QSAR) sont des modèles qui relient l'activité ou propriété biologique aux caractéristiques structurelles (descripteurs moléculaires) calculées dans l'espace tridimensionnel pour un groupe de molécules. Les méthodes CoMFA (Comparative Molecular Field Analysis) et CoMSIA (Comparative Molecular Similarity Indices Analysis) sont les méthodes les plus utilisées dans l'approche 3D-QSAR pour étudier la relation quantitative de structure-activité au niveau tridimensionnel [184, 185].

a. Alignement des molécules

L'alignement des molécules et la sélection de la conformation "active" sont les clés du succès d'un modèle 3D-QSAR par les méthodes CoMFA et CoMSIA. Plusieurs approches sont disponibles pour superposer les structures moléculaires [186] (atome par atome, basé sur le pharmacophore, basé sur le récepteur, etc). Cela permet de comparer les molécules sur la base de la différence du groupement structurel des composés étudiés. Les règles d'alignement doivent être aussi reproductibles que possible, et il faut donc éviter tout ajustement manuel ou arbitraire [187]. Cette étape comprend également la génération de conformations tridimensionnelles et éventuellement leur sélection à partir d'une recherche conformationnelle [188].

b. Méthode CoMFA

Pendant plusieurs années, l'approche 3D-QSAR a été utilisée comme synonyme de CoMFA [189], qui a été la première méthode à mettre en œuvre le concept dans une approche 3D-QSAR, c'est-à-dire que l'activité biologique d'un ligand peut être prédite à partir de sa structure 3D. Jusqu'à présent, la méthode CoMFA est probablement la méthode 3D-QSAR la plus couramment appliquée [189, 190]. L'idée sous-jacente de la méthode CoMFA est que les différences dans une propriété cible, par exemple les activités biologiques, sont souvent étroitement liées à des changements équivalents dans les formes et les forces des champs d'interaction non covalents entourant les molécules. Autrement dit, les champs stérique et électrostatique fournissent toutes les informations nécessaires à la compréhension des propriétés biologiques d'un ensemble de composés. Ainsi, les molécules sont placées dans une grille cubique et les énergies d'interaction entre la molécule et une sonde définie sont calculées pour chaque point de grille. Normalement, seuls deux potentiels, à savoir un potentiel stérique sous la forme d'une fonction de Lennard-Jones et un potentiel électrostatique sous la forme d'une simple fonction de Coulomb, sont utilisés dans une étude CoMFA. Il est évident que la description de la similarité moléculaire n'est pas une tâche triviale, tout comme la description du processus d'interaction des ligands avec les cibles biologiques correspondantes. Dans l'application standard de CoMFA, seules les contributions enthalpiques de l'énergie libre de liaison sont fournies par les potentiels utilisés. Cependant, de nombreux effets de liaison sont régis par des contributions hydrophobes et entropiques. Par conséquent, il faut caractériser à l'avance les principales contributions attendues des forces et déterminer si, dans ces conditions, CoMFA sera effectivement en mesure de trouver des résultats réalistes. Les étapes principales pour construire un modèle CoMFA sont présentées dans la figure III. 9.

c. Méthode CoMSIA

En raison des problèmes associés à la forme fonctionnelle du potentiel de Lennard-Jones utilisé dans la plupart des méthodes CoMFA, Klebe et al. [191] ont développé une méthode basée sur les indices de similarité, appelée CoMSIA (comparative molecular similarity indices analysis). Au lieu de champs basés sur les champs de force, CoMSIA est basé sur des indices de similarité qui sont obtenus en utilisant une forme fonctionnelle adaptée de l'algorithme SEAL (Steric and Electrostatic ALignment). Trois indices différents liés aux potentiels stérique, électrostatique et hydrophobe ont été utilisés dans leur étude de l'ensemble de données de référence des stéroïdes classiques. Des modèles de qualité statistique comparable en ce qui concerne la

validation croisée de l'ensemble d'apprentissage, ainsi que la capacité prédictive d'un ensemble de test, ont été dérivés en utilisant CoMSIA. L'avantage de cette méthode réside dans les fonctions utilisées pour décrire les molécules étudiées, ainsi que dans les cartes de contour résultantes. Les cartes de contour obtenues à partir de la méthode CoMSIA sont généralement plus informatives, par rapport à celles obtenues par la méthode CoMFA. La méthode CoMSIA évite également les valeurs de coupure utilisées dans la méthode CoMFA pour restreindre les fonctions potentielles en prenant des valeurs inacceptables. Les étapes principales pour construire un modèle CoMSIA sont présentées dans la figure III. 9.

d. Elaboration du modèle

Par l'utilisation de la méthode de régression des moindres carrés partiels (PLS), les champs d'interaction moléculaire (ou MIF) calculés de type CoMFA et CoMSIA sont corrélées avec la "variable dépendante" (activité biologique, espace Y), afin de trouver toute relation possible entre elles. Comme les variables indépendantes (espace X) sont numériquement beaucoup plus importantes que le nombre de composés testés, la méthode PLS est utilisée pour extraire les composantes principales (PC) afin de réduire la dimensionnalité de l'espace X et d'établir une corrélation statistique valable avec l'espace Y [169]. La qualité des modèles est évaluée par le coefficient de détermination (R^2), le nombre de composant (N), la valeur de Fisher (F) et la valeur de RMSE comme indiqué précédemment.

e. Validation du modèle

Le modèle élaboré doit être vérifié en termes de robustesse, de corrélation aléatoire et de capacité de prédiction [169]. Pour cela, les coefficients q^2 et SDEP (Standard Deviation Error of Prediction) sont utilisés à chaque étape de validation. La robustesse statistique du modèle est évaluée par validation croisée (CV) [192, 193], tandis que l'absence de corrélation fortuite peut être évaluée par la procédure de brouillage en Y (Y-randomisation) [169, 192]. Bien que, selon Clark et al. [192], la validation croisée soit souvent utilisée à tort pour indiquer la capacité prédictive du modèle, l'utilisation d'ensembles de tests externes devrait être considérée comme la seule méthode possible pour évaluer la capacité prédictive d'un modèle [167].

f. Interprétation graphique

Contrairement aux QSAR classiques (2D-QSAR), le modèle 3D-QSAR peut être visualisé par un certain nombre de polyèdres obtenus en connectant des points de grille similaires. La

visualisation est normalement obtenue en sélectionnant des nœuds isovaleurs qui peuvent représenter différentes caractéristiques chimiques déterminées par la sonde utilisée. À chaque point de grille peuvent être associées différentes valeurs : le champ réel (MIF) ; les coefficients PLS à un nombre donné de composantes principales pour obtenir les graphiques de coefficients PLS ; les cartes de contour standard associant à chaque point de grille ; le produit du coefficient PLS par l'écart type MIF correspondant calculé sur toutes les molécules de l'ensemble d'apprentissage. Au moyen d'une analyse approfondie et fastidieuse des résultats graphiques, il est possible de décrire comment les caractéristiques structurelles moléculaires de l'ensemble d'apprentissage peuvent influer favorablement ou défavorablement sur l'activité et ainsi prédire *in silico* de nouvelles molécules ayant des activités biologiques potentielles.

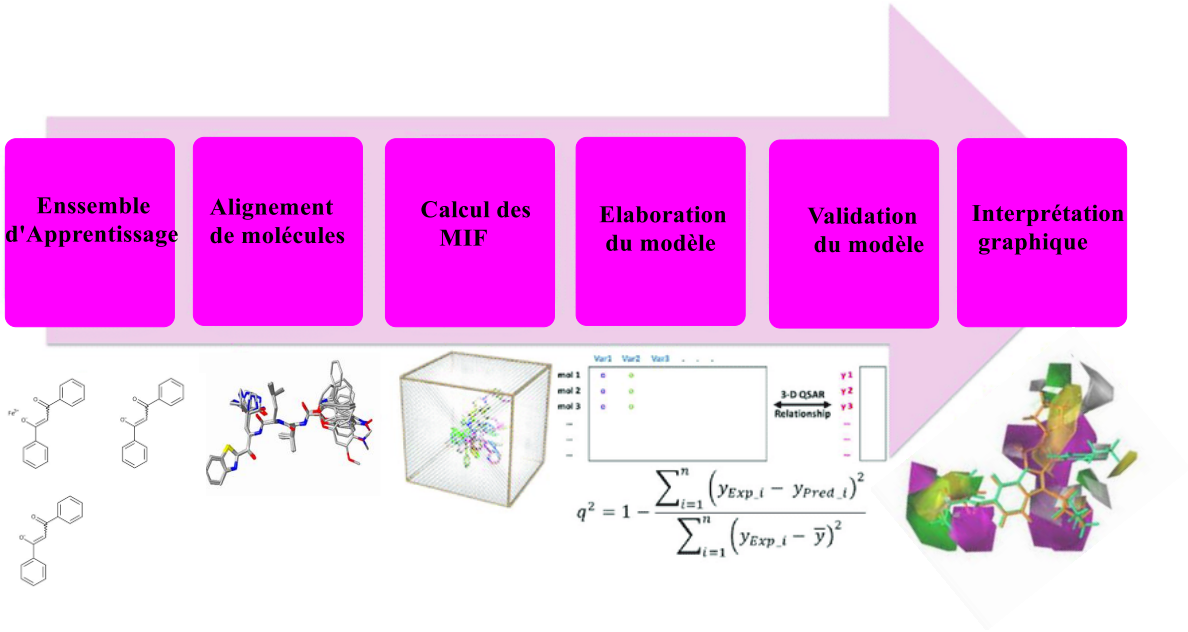


Figure III.9. Schéma de construction d'un modèle 3D-QSAR

IV. Criblage virtuel basé sur la structure « Structure-Based »

IV.1. Modélisation par Homologie

La détermination de la structure tridimensionnelle (3D) des cibles thérapeutiques par des méthodes expérimentales continuera à alimenter et à entretenir la base de données des protéines. En Mars 2022, plus de 190404 structures protéiques expérimentales sont disponibles dans la base de donné PDB (Protein Data Bank, http://www.rcsb.org/pdb). Les méthodes expérimentales telles que la méthode de la cristallographie aux rayons X et la méthode de la spectroscopie RMN [194, 195] sont des méthodes longues et plus coûteuses. De plus, elles échouent pour toutes les protéines, en particulier les protéines membranaires [196]. Pour pallier à ce problème,

différentes approches ont été utilisées pour prédire la structure 3D d'une cible thérapeutique à partir de sa séquence premières (2D) d'acides aminés. La modélisation par homologie (ou homology modeling) est l'approche la plus précise qui est basée sur l'hypothèse que dans une même famille les protéines avec des séquences similaires ont des structures 3D similaires. Le développement d'un modèle homologue est un processus comportant plusieurs étapes, que l'on peut résumer dans la figure III. 10, suivante :

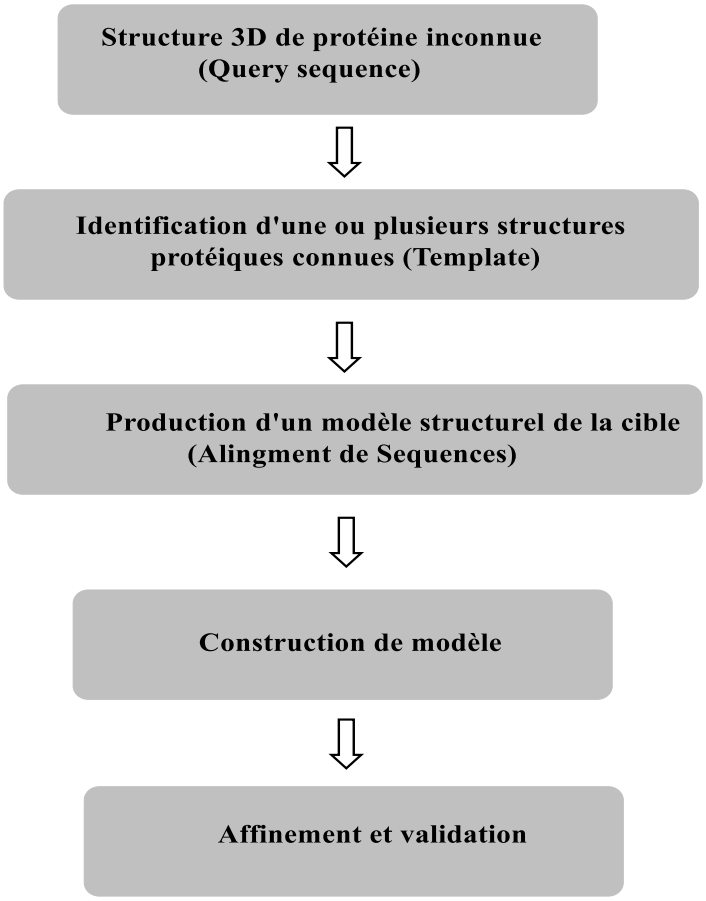


Figure III. 10. *Etapes de prédiction d'une structure 3D d'une protéine par homologie*

Il existe plusieurs programmes informatiques et serveurs Web qui automatisent le processus de modélisation par homologie, notamment le programme Modeller et le serveur Web SWISS Model.

IV.2. Docking Moléculaire

Le docking moléculaire est l'un des outils les plus utilisés pour la conception de médicaments assistée par ordinateur. Il vise principalement à trouver la position d'un ligand dans le site de liaison d'une protéine cible [197, 198] (Figure III. 11). Ce type d'analyse computationnelle des interactions protéine-ligand joue un rôle essentiel dans la recherche et développement de

nouveaux agents thérapeutiques ainsi que dans la compréhension des processus biochimiques fondamentaux [199, 200]. La prédiction de la pose d'un ligand dans le site de liaison d'une protéine, peut être réalisée à l'aide de différentes approches, parmi lesquelles on trouve, le docking rigide, le docking flexible ou le docking semi-flexible.

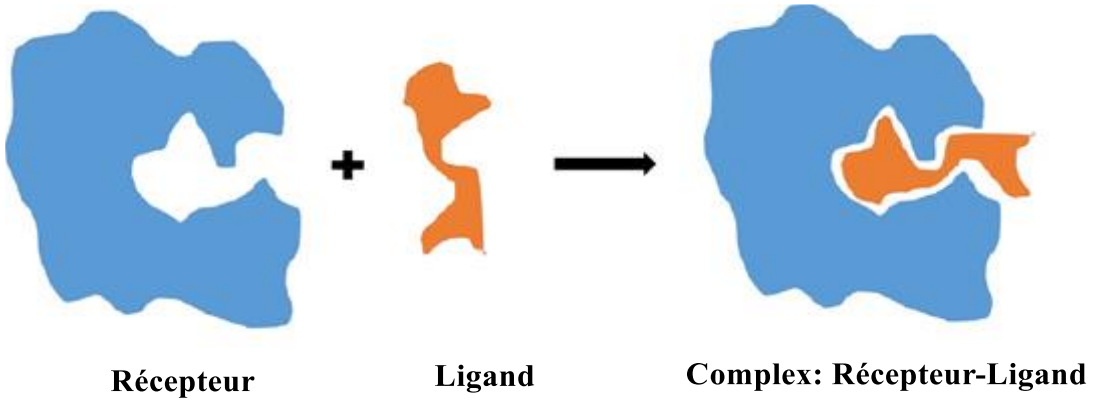


Figure III. 11. *Schéma simplifié du docking moléculaire*

Différents programmes de docking moléculaire ont été proposés au fil des ans, notamment DOCK [201], AutoDock4 [202], GLIDE142, GOLD [203], PLANTS [204], ICM2, FlexX [203] et d'autres encore [205, 206].

IV.2.1. Docking de molécules rigides

Jusqu'au milieu du XXe siècle, le mécanisme de liaison d'un ligand à un récepteur était compris comme un processus statique impliquant deux formes stériques complémentaires (concept « clé-serrure ») [207], selon lequel le ligand qui représente la clé, est complémentaire au niveau géométrique du site actif du récepteur, qui représente la serrure. Considérant donc le ligand et le récepteur comme deux entités rigides, la recherche de la pose optimale se limite au positionnement en énumérant toutes les translations et les rotations possibles pour le ligand à l'intérieur du site actif rigide. Ensuite toutes les poses incompatibles avec le site actif sont éliminées. Enfin, les poses sélectionnées se voient attribuées un score et les meilleures sont optimisées [208]. L'intérêt de cette approche de docking rigide dans sa rapidité d'exécution qui permet de réaliser un premier filtrage de très larges chimiothèques en éliminant les molécules aberrantes (trop grandes, mauvaise complémentarité avec le site de liaison,…) [209] avant de procéder, par exemple, à la mise en œuvre de méthodes de docking flexible.

IV.2.2. Docking flexible

Bien que l'approche de docking rigide permet l'obtention rapide de résultats préliminaires, elle repose toutefois sur des outils heuristiques non-représentatifs du mécanisme biologique d'interaction entre un ligand et son récepteur, puisque la liaison du ligand à un site de liaison se fait en réalité dans la majorité des cas par un processus de sélection de conformation qui se termine par une étape d'adaptation du ligand et du récepteur l'un à l'autre [210]. Cette approche simpliste du problème a donc été améliorée par l'introduction de la flexibilité sur le ligand avec l'essor des moyens de calcul, permettant l'exploration des conformations adoptées par celui-ci lors de docking dans le site de liaison [208]. En effet cette approche, permettant d'explorer les conformations adoptées par le ligand lors de sa fixation au site de liaison, est maintenant classiquement prise en compte dans la plupart des études de docking, et représente actuellement l'approche la plus utilisée dans la plupart des logiciels car elle permet une évaluation rapide de bases de milliers, voire de millions de molécules [211]. A l'inverse, la prise en compte de la flexibilité du récepteur, proposée par de plus en plus de logiciels de docking [242] est encore trop onéreuse en temps de calcul pour être systématiquement envisagée dans les études de docking.

Les algorithmes de docking permettant de simuler une flexibilité moléculaire peuvent être classés en trois catégories : les algorithmes de recherche systématique, stochastique ou déterministe [244].

IV.3. Simulations par dynamique moléculaire

La dynamique moléculaire (DM) calcule le comportement d'un système moléculaire en fonction du temps. Des informations détaillées sur les fluctuations et les changements de conformation des protéines et des acides nucléiques sont obtenues en intégrant les lois du mouvement de Newton. La force agissant sur une particule accélérée de masse m_i est donnée par :

$$F_i = m_i\, a_i \qquad \textbf{Eq. III. 12}$$

La force est dérivée de la fonction d'énergie potentielle par rapport aux coordonnées internes :

$$F = -\frac{\partial U\,(r)}{\partial ri} \qquad \textbf{Eq. III. 13}$$

En combinant Eq. III. 12 et Eq. III. 13, on obtient la deuxième loi de Newton :

$$-\frac{\partial U(r)}{\partial ri} = mi\ \frac{\partial^2 r}{\partial t^2} \qquad \textbf{Eq. III. 14}$$

Comme les mouvements de toutes les particules sont couplés, la solution analytique d'**Eq. III. 14** n'est pas possible. Il existe de nombreuses méthodes permettant d'effectuer une intégration numérique pas à pas des équations différentielles ordinaires couplées, telles que l'algorithme de Verlet, l'algorithme LeapFrog et l'algorithme de Verlet pour la vitesse [217].

Dans la simulation par DM, les systèmes infiniment grands sont représentés par un ensemble de conditions limites utilisant une petite partie appelée cellule unitaire. Cette cellule unitaire est répliquée dans tout l'espace dans les trois directions cartésiennes pour former un réseau infini. Toutes les copies périodiques des particules, appelées images, se déplacent solidairement dans la même direction pendant la simulation, de sorte que le nombre de particules dans la boîte de simulation est toujours conservé. Cette procédure est appelée conditions périodiques aux limites (CPL, en anglais periodic boundary conditions - PBC) [218].

Les propriétés thermodynamiques des systèmes sont obtenues en utilisant des ensembles statistiques. L'ensemble est la distribution de probabilité des états d'un système et il est défini par un petit ensemble de paramètres tels que la pression (P), la température (T), l'énergie (E), le volume (V), le potentiel chimique (μ) et le nombre de particules (N). Les ensembles les plus courants en dynamique moléculaire sont l'ensemble microcanonique, l'ensemble canonique, l'ensemble isobarique-isotherme et le grand ensemble canonique. Dans l'ensemble microcanonique (NVE), le système est isolé avec un nombre de particules, un volume et une énergie constants. Dans l'ensemble canonique (NVT), le nombre de particules dans le système, le volume et la température sont fixes. L'ensemble isobarique isotherme (NPT) est caractérisé par un nombre d'atomes, une pression et une température constants. L'état thermodynamique du grand ensemble canonique (μVT) est décrit par un potentiel chimique, un volume et une température fixes.

Pour les simulations de systèmes biologiques, l'ensemble canonique est préféré en raison de son efficacité de calcul. Différents algorithmes de thermostat sont utilisés afin d'ajuster la température constante pendant les simulations. Ces algorithmes utilisent la relation entre l'énergie cinétique (E_K) et la température instantanée (T).

$$E_K = \frac{N_{df}}{2}\ K_\beta T = \frac{1}{2}\sum_i^{N_{df}} \frac{Pi^2}{mi}$$

Où N_{df} représente le nombre de degrés de liberté, k_β est la constante de Boltzmann, p_i et m_i expriment respectivement la quantité de mouvement (momentum) et la masse de la particule i.

Dans le thermostat d'Andersen [219], le système est couplé thermiquement avec un bain de chaleur fictif qui impose la température désirée. La méthode est stochastique et modifie aléatoirement la vitesse des particules thermostatées en choisissant parmi une distribution de Maxwell-Boltzmann. Le thermostat de Berendsen [220] utilise le changement d'échelle de la vitesse de façon périodique pour maintenir la température. Dans cette approche, le système est supposé être faiblement couplé à un bain de chaleur et les vitesses des particules sont redimensionnées à chaque pas de temps en fonction de l'énergie cinétique moyenne ciblée. Dans le thermostat de Langevin [221], à chaque pas de temps, toutes les particules sont exposées à une force aléatoire et leur vitesse est réduite par une friction constante. La quantité de mouvement des particules suit une distribution Maxwell-Boltzmann à la température souhaitée.

IV.3. Méthodes de calcul de l'énergie libre de liaison

Les interactions entre les biomolécules sont à la base de tous les processus biologiques. L'énergie libre de liaison mesure la stabilité à l'équilibre thermique et indique la quantité d'énergie libérée lors de la formation d'un complexe. L'énergie libre de liaison est le changement d'énergie libre lorsque deux entités (A et B) se combinent pour former un complexe et peut être calculée comme la différence entre les potentiels chimiques du complexe et ceux des entités individuelles A et B. Pendant la formation d'un tel complexe, l'entropie de conformation est compensée en fonction de la faisabilité de la complexation.

La perturbation de l'énergie libre [222] et l'intégration thermodynamique [223] sont les méthodes les plus précises et les plus rigoureuses parmi celles qui calculent directement la différence d'énergie libre et ces calculs sont basés sur les changements de la structure moléculaire pendant la dynamique moléculaire ou les simulations de Monte-Carlo. Toutefois, ces méthodes sont très coûteuses en termes de calcul. Les méthodes de Mécanique Moléculaire-Poisson Boltzmann/Surface (MM-PB/SA) et de Mécanique Moléculaire-Généralisée Born/Surface (MM-GB/SA) [224, 225] sont plus efficaces en termes de coût et de simplicité. Ces méthodes nécessitent la simulation directe des points finaux des états liés et non liés en utilisant la solvatation implicite. La procédure peut être brièvement décrite en trois étapes. Dans la première étape, l'échantillonnage des coordonnées est préparé pour les simulations MD du complexe protéine-ligand afin d'échantillonner les configurations pour l'analyse énergétique.

Dans la deuxième étape, les énergies potentielles de la phase gazeuse et les énergies libres de solvatation sont calculées à partir d'une moyenne d'ensemble. Dans la dernière étape, la variation estimée de l'entropie est calculée. Ensuite, l'énergie libre de liaison finale est calculée par :

$$\Delta G^0_{bind,solv} = \Delta G^0_{bind,vacuum} + \Delta G^0_{solv,complex} - \Delta G^0_{solv,ligand} - \Delta G^0_{solv,receptor}$$

$\Delta G^0_{bind,vacuum}$, est obtenu par :

$$\Delta G^0_{bind,vacuum} = \Delta E^0_{Molecular\ Mechanics} - T\Delta S^0$$

Le terme d'énergie de solvatation G_{solv} est généralement obtenu en résolvant l'équation de Poisson-Boltzmann (PB) ou en utilisant le modèle de Born généralisé (GB).

Chapitre IV: Etudes pharmacocinétiques

I. Introduction

La recherche et le développement de médicaments est un processus extrêmement complexe et coûteux, qui comprend la sélection de la maladie, l'identification et la validation de la cible, la découverte et l'optimisation de la piste, ainsi que les essais précliniques et cliniques [226]. Avec le développement des méthodes *in silico* ces dernières années, le nombre de nouvelles entités moléculaires (NEM) approuvées par FDA (Food and Drug Administration) a augmenté de manière remarquable [227]. Le manque d'efficacité et de sécurité sont les deux principales causes de l'échec des médicaments [228], ce qui signifie que les propriétés d'absorption, de distribution, de métabolisme, d'excrétion et de toxicité (ADMET) des produits chimiques jouent un rôle essentiel à chaque étape de la découverte et du développement des médicaments. Il est donc nécessaire de trouver des molécules efficaces ayant de meilleures propriétés ADMET.

Au cours de la phase initiale de la découverte d'un médicament, le concept de « drug-like » constitue une ligne directrice utile [229, 230]. Lipinski et al, ont proposé la « règle de cinq » en 1997, qui a été le filtre original et le plus connu de la pharmacocinétique basé sur des règles distinguant si une molécule est bien absorbée par la voie orale ou non : le poids moléculaire (MW) ≤ 500, le coefficient de partage octanol/eau (log P) ≤ 5, le nombre de donneurs de liaisons hydrogène (HBD) ≤ 5, le nombre d'accepteurs de liaisons hydrogène (HBA) ≤ 10 et le nombre de liaison relatives ≤ 10 [231]. Selon la règle de cinq, une molécule ne serait pas active par voie orale si elle viole deux ou plusieurs paramètres. Cependant, ces règles ne sont pas adaptées aux produits naturels complexes car elles ont été déduites de petites molécules relativement simples. Par la suite, plusieurs règles/filtres de « drug like » similaires à la « règle de cinq » ont été proposées telles que la règle de Veber, la règle de Ghose, la règle de Muegge, etc. Par exemple, Ghose et al, ont montré que plus de 80 % des composés (6304 molécules de la base de données CMC (Comprehensive Medicinal Chemistry)) respectent les limites suivantes : -0,4 ≤ log P ≤ 5,6, 160 ≤MW ≤ 480, 40 ≤ MR (réfractivité molaire) ≤ 130, et le nombre total d'atomes entre 20 et 70 [232]. Les filtres d'apparentement des médicaments basés sur les propriétés physicochimiques accélèrent le processus de développement des médicaments [233]. Cependant, un certain nombre d'études montrent que ces règles d'apparentement des médicaments basés sur les propriétés physicochimiques ont leurs limites [232]. Pour étudier le « drug-like », un certain nombre de modèles de prédiction ont été élaborés à l'aide de méthodes d'apprentissage automatique, telles que la machine à vecteurs de support (ou SVM) [234], les

réseaux de neurones artificiels (ANN) [235], le partitionnement récursif (ou RP) [236], et la méthode bayésienne naïve (ou NB) [236].

1. II. Etude ADMET

L'étude ADMET (Absorption, Distribution, Métabolisme, Élimination et Toxicité) décrit le devenir des médicaments dans l'organisme, en particulaire dans le corps humain. La principale cause de rejet d'un médicament candidat est ses mauvaises propriétés pharmacocinétiques. Le développement d'un médicament pendant la phase clinique entraine des pertes importantes en termes de temps et de coûts. La mise en place de filtres ADMET des chimiothèques avant tout processus de criblage a donné d'excellents résultats. En effet, même si des tentatives d'amélioration de la prédiction des propriétés ADMET sont encore à l'étude, le taux d'échec attribuable à une mauvaise pharmacocinétique ou biodisponibilité a chuté de 40 % à moins de 10 % en 10 ans. Actuellement, le manque d'efficacité et la toxicologie sont les principales causes d'arrêt de développement de médicaments candidat [226].

a. Absorption

L'absorption est un processus complexe régi par une grande variété de facteurs, y compris non seulement les propriétés intrinsèques de la substance (taille moléculaire, solubilité (Log Saq), constante d'ionisation pKa et coefficient de partage octanol/eau (Log P)), mais aussi les conditions physiologiques à l'intérieur de l'organisme (pH local, surface d'absorption) et les activités des enzymes, des transporteurs et des porteurs le long du tractus gastro-intestinal (GI). Dans le domaine de la pharmacocinétique, l'absorption désigne le processus par lequel les médicaments traversent les organes de l'organisme pour atteindre la circulation systémique. Les médicaments ont différentes voies d'administration : orale, intraveineuse, intramusculaire, intrapéritonéale, oculaire, nasale, rectale, intratrachéale et sous-cutanée. Dans ce travail, nous nous concentrons principalement sur la voie orale car c'est la voie d'administration la plus courante.

L'absorption intestinale humaine est généralement mesurée comme le pourcentage de la dose qui atteint la veine porte après avoir traversé la paroi intestinale et constitue la base de la plupart des modèles d'absorption *in silico*. Le pourcentage de la dose qui reste après l'absorption et métabolisme hépatique de premier passage est défini comme la biodisponibilité orale (F) du composé. En d'autres termes, la biodisponibilité décrit le passage d'une substance du site

d'absorption dans la circulation systémique et n'est généralement pas équivalente à la quantité d'une substance absorbée.

b. Distribution

Apres l'absorption d'un composé par la membrane intestinale, il doit être distribué dans le corps pour atteindre sa cible. La distribution est définie comme le transport du médicament d'un tissu/organe à un autre tissu/organe. Parmi tous les tissus, le sang transporte principalement le médicament vers d'autres tissus en fonction de l'affinité. En général, le principe actif du médicament présent sous deux formes dans le plasma : une forme active libre, qui permet au principe actif de diffuser librement et d'atteindre l'organe cible où il exercera son action ; une forme liée aux protéines plasmatiques, sous laquelle le principe actif ne peut diffuser, et donc atteindre son site d'action. Les médicaments acides se lient principalement à l'albumine, et les médicaments basiques à l'acide alpha glycoprotéique.

La barrière hémato-encéphalique (BHE) est un critère important de la régulation de l'environnement interne du cerveau. Les médicaments agissant dans le système nerveux central doivent traverser le BHE. Les paramètres moléculaires « règle de cinq » ont été conçus pour faciliter les prédictions de la pénétration à la BHE. Ils participent à la détermination de la capacité des molécules à traverser la BHE [236]. Par exemple les molécules de la masse moléculaire inférieure à 450 Dalton ou de PSA (polar surface area) inférieure à 100 $Å^2$ sont considérées comme plus susceptibles de pénétrer le BHE.

c. Métabolisme

Le métabolisme est la biotransformation ou la modification chimique que subit une substance médicamenteuse dans l'organisme. Le résultat de ces biotransformations est généralement la réduction de la substance en une forme inerte, appelée métabolite inactif, plus facilement éliminable. Les substances médicamenteuses ne sont cependant pas toutes inactivées par le métabolisme ; certains médicaments dits précurseurs ne deviennent actifs qu'après leur biotransformation, d'autres sont biotransformés en un ou plusieurs métabolites actifs. La plupart des réactions métaboliques sont enzymatiques et se produisent dans le foie, organe très vascularisé et riche en enzymes. D'autres organes comme les poumons, le tube digestif et les reins y contribuent également, mais de façon beaucoup moins importante.

Le métabolisme peut être classé en deux phases principales, la phase I et la phase II. Le métabolisme de phase I est principalement régi par les enzymes du cytochrome P450 (CYP) et

comprend des réactions d'oxydation, de réduction et d'hydrolyse. Le métabolisme de phase II consiste principalement en des réactions de conjugaison. Des fragments augmentant l'hydrophilie sont ajoutés aux composés exogènes ou à leurs métabolites. Les métabolites obtenus à la suite des réactions de phase I et/ou de phase II dans le foie peuvent subir d'autres réactions de biotransformation en interagissant avec d'autres enzymes [237]. La réaction de phase I génère généralement les sites de réaction de conjugaison pour la phase II ; cependant, la phase I ne précède pas nécessairement la phase II.

d. Excrétion

Les médicaments sont éliminés de l'organisme par métabolisme et excrétion. Les voies d'excrétion les plus courantes sont l'excrétion rénale et/ou biliaire. Le rein est le principal contributeur à l'excrétion du médicament et de certains métabolites. Les grosses moléculaires et de nombreux médicaments conjugués à l'acide glucuronique sont généralement excrétés dans la bile. La filtration à travers les membranes glomérulaires est le principal mécanisme d'excrétion dans les urines. Cependant, certains médicaments sont excrétés par des mécanismes de transport actifs servant de substrats aux transporteurs épithéliaux tubulaires ou biliaires.

Les systèmes de transport actifs de médicaments responsables de leurs éliminations [238] sont les transporteurs de peptides, les transporteurs d'anions organiques et de cations peuvent également transporter des composés et/ou leurs métabolites sur la surface basolatérale des membranes rénale ou hépatique.

D'autres familles de transporteurs, tels que : P-gp, BCRP et MDR, assurent la médiation de l'efflux sur les surfaces apicales, entraînant la sécrétion et l'élimination. L'excrétion passive peut théoriquement être prédite à l'aide de certaines des approches permettant de prédire la distribution tissulaire, car elle est déterminée par des paramètres physicochimiques similaires, et des propriétés physiologiques (débit sanguin, liaison aux protéines, lipophilie, pKa), éventuellement avec des limites différentes. La modélisation *in silico* des substrats de la P-gp et de l'inhibiteur peut être utilisée pour prédire l'excrétion car elle agit comme une pompe à effluent de l'intérieur de la cellule rénale dans le conduit collecteur [238].

e. Toxicité

Complémentairement aux propriétés ADME, il est également possible de filtrer les médicaments candidats selon la toxicité. En effet, un médicament candidat présentant des effets secondaires graves lors des phases cliniques est immédiatement abandonné, et ce généralement

de manière définitive. Il est donc très important de tenter de prédire la toxicité d'un composé à partir de sa structure lors des phases précoces de la recherche et développement. La toxicité peut être classifiée selon l'organe affecté (hépatotoxicité, néphrotoxicité, cardiotoxicité, …) ou selon le mécanisme de toxicité (génotoxicité, mutagénicité, carcinogénicité, ...). De nombreuses méthodes *in silico* ont donc été mises en place pour tenter de prédire la toxicité, les effets secondaires et le métabolisme des composés [238], c'est ce qu'on appelle la toxicologie *in silico* ou toxicologie computationelle [239].

1. III. Prédiction des spectres d'activité des substances (PASS)

La prédiction des spectres d'activité des substances (Prediction of Activity Spectra for Substances ou PASS) est un outil d'évaluation des réponses biologiques des composés chimiques. PASS prédit environ 2500 types d'activités biologiques sur la base de la formule structurelle du composé chimique. Les types biologiques comprennent les effets pharmacologiques principaux et secondaires (par exemple, antihypertenseur, hépatoprotecteur, sédatif, etc.), les mécanismes d'action (agoniste de la 5-hydroxytryptamine, inhibiteur de l'acétylcholinestérase, etc.), les toxicités spécifiques (mutagénicité, carcinogénicité, tératogénicité, etc.) et les noms des enzymes de métabolisation (substrat du CYP1A, substrat du CYP3A4, etc.) [240]. L'approche PASS est basée sur l'analyse des relations structure-activité (SAR) pour l'ensemble d'apprentissage comprenant actuellement environ 60 000 médicaments, médicaments candidats, leads et composés toxiques dont l'activité biologique a été déterminée expérimentalement [241]. Les résultats de la prédiction sont une liste de types d'activité, avec la probabilité de présence (Pa) et d'absence (Pi) pour chaque activité particulière. Par définition, les probabilités Pa et Pi peuvent également être interprétées comme les mesures d'appartenance à des sous-ensembles de composés "actifs" et "inactifs" [240].

Références bibliographiques

1. Salman MM, Al-obaidi Z, Kitchen P, et al (2021) Advances in Applying Computer-Aided Drug Design for Neurodegenerative Diseases. Int J Mol Sci 9:1–22. https://doi.org/10.1021/jm990129n

2. Livingston G, Sommerlad A, Orgeta V, et al (2017) Dementia prevention, intervention, and care. Lancet 390:2673–2734. https://doi.org/10.1016/S0140-6736(17)31363-6

3. Królicka E, Kieć-Kononowicz K, Łażewska D (2022) Chalcones as Potential Ligands for the Treatment of Parkinson's Disease. Pharmaceuticals 15:. https://doi.org/10.3390/ph15070847

4. Rehuman NA, Oh JM, Abdelgawad MA, et al (2022) Development of Halogenated-Chalcones Bearing with Dimethoxy Phenyl Head as Monoamine Oxidase-B Inhibitors. Pharmaceuticals 15:. https://doi.org/10.3390/ph15091152

5. Ullah A, Munir S, Badshah SL, et al (2020) Important flavonoids and their role as a therapeutic agent. Molecules 25:1–39. https://doi.org/10.3390/molecules25225243

6. Guglielmi P, Mathew B, Secci D, Carradori S (2020) Chalcones: Unearthing their therapeutic possibility as monoamine oxidase B inhibitors. Eur J Med Chem 205:112650. https://doi.org/10.1016/j.ejmech.2020.112650

7. Llorach-Pares L, Nonell-Canals A, Avila C, Sanchez-Martinez M (2022) Computer-Aided Drug Design (CADD) to De-Orphanize Marine Molecules: Finding Potential Therapeutic Agents for Neurodegenerative and Cardiovascular Diseases. Mar Drugs 20:. https://doi.org/10.3390/md20010053

8. Scotti L, Scotti M (2015) Computer Aided Drug Design Studies in the Discovery of Secondary Metabolites Targeted Against Age-Related Neurodegenerative Diseases. Curr Top Med Chem 15:2239–2252. https://doi.org/10.2174/1568026615666150610143510

9. Derrick JS, Lim MH (2015) Tools of the Trade: Investigations into Design Strategies of Small Molecules to Target Components in Alzheimer's Disease. ChemBioChem 16:887–898. https://doi.org/10.1002/cbic.201402718

10. Chao A, Sieminski PJ, Owens CP, Goulding CW (2019) Iron Acquisition in Mycobacterium tuberculosis. Chem Rev 119:1193–1220. https://doi.org/10.1021/acs.chemrev.8b00285

11. Kayed R, Head E, Thompson JL, et al (2003) Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. Science (80-) 300:486–489. https://doi.org/10.1126/science.1079469

12. Haass C, Selkoe DJ (2007) Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid β-peptide. Nat Rev Mol Cell Biol 8:101–112. https://doi.org/10.1038/nrm2101

13. Stelzmann RA, Norman Schnitzlein H, Reed Murtagh F (1995) An english translation of alzheimer's 1907 paper, "über eine eigenartige erkankung der hirnrinde." Clin Anat 8:429–431. https://doi.org/10.1002/ca.980080612

14. World Health Organization (WHO) 2022 (2022) https://www.who.int/health-topics/dementia#tabtab_1

15. Frisoni GB, Laakso MP, Beltramello A, et al (1999) Hippocampal and entorhinal cortex atrophy in frontotemporal dementia and Alzheimer's disease. Neurology 52:91–100. https://doi.org/10.1212/wnl.52.1.91

16. Hardy J, Selkoe DJ (2002) The amyloid hypothesis of Alzheimer's disease: Progress and problems on the road to therapeutics. Science (80-) 297:353–356. https://doi.org/10.1126/science.1072994

17. Grundke-Iqbal I, Iqbal K, Tung YC, et al (1986) Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. Proc Natl Acad Sci U S A 83:4913–4917. https://doi.org/10.1073/pnas.83.13.4913

18. De Strooper B, Annaert W, Cupers P, et al (1999) A presenilin-1-dependent γ-secretase-like protease mediates release of notch intracellular domain. Nature 398:518–522. https://doi.org/10.1038/19083

19. Lieberman DE (2007) This material is the copyright of the original publisher . Unauthorised copying and

distribution is prohibited . Rev Lit Arts Am 24:573–580. https://doi.org/10.2165/11536850-000000000-00000

20. Leissring MA, Lu A, Condron MM, et al (2003) Kinetics of amyloid β-protein degradation determined by novel fluorescence- and fluorescence polarization-based assays. J Biol Chem 278:37314–37320. https://doi.org/10.1074/jbc.M305627200

21. Marr RA, Guan H, Rockenstein E, et al (2004) Neprilysin regulates amyloid β peptide levels. J Mol Neurosci 22:5–11. https://doi.org/10.1385/JMN:22:1-2:5

22. Nalivaeva NN, Turner AJ (2019) Targeting amyloid clearance in Alzheimer's disease as a therapeutic strategy. Br J Pharmacol 176:3447–3463. https://doi.org/10.1111/bph.14593

23. Wang M, Beckmann ND, Roussos P, et al (2018) The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. Sci Data 5:1–16. https://doi.org/10.1038/sdata.2018.185

24. Kurochkin I V., Guarnera E, Berezovsky IN (2018) Insulin-Degrading Enzyme in the Fight against Alzheimer's Disease. Trends Pharmacol Sci 39:49–58. https://doi.org/10.1016/j.tips.2017.10.008

25. Goedert M, Spillantini MG, Jakes R, et al (1989) Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. Neuron 3:519–526. https://doi.org/10.1016/0896-6273(89)90210-9

26. Ballatore C, Lee VMY, Trojanowski JQ (2007) Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. Nat Rev Neurosci 8:663–672. https://doi.org/10.1038/nrn2194

27. Ward SM, Himmelstein DS, Lancia JK, Binder LI (2012) Tau oligomers and tau toxicity in neurodegenerative disease. Biochem Soc Trans 40:667–671. https://doi.org/10.1042/BST20120134

28. Zhang Z, Kang SS, Liu X, et al (2017) Asparagine endopeptidase cleaves α-synuclein and mediates pathologic activities in Parkinson's disease. Nat Struct Mol Biol 24:632–642. https://doi.org/10.1038/nsmb.3433

29. Ahn EH, Kang SS, Liu X, et al (2020) Initiation of Parkinson's disease from gut to brain by δ-secretase. Cell Res 30:70–87. https://doi.org/10.1038/s41422-019-0241-9

30. Maragakis NJ, Rothstein JD (2006) Mechanisms of Disease: Astrocytes in neurodegenerative disease. Nat Clin Pract Neurol 2:679–689. https://doi.org/10.1038/ncpneuro0355

31. Bendor JT, Logan TP, Edwards RH (2013) The function of α-synuclein. Neuron 79:1044–1066. https://doi.org/10.1016/j.neuron.2013.09.004

32. Meuvis J, Gerard M, Desender L, et al (2010) The conformation and the aggregation kinetics of α-synuclein depend on the proline residues in its C-terminal region. Biochemistry 49:9345–9352. https://doi.org/10.1021/bi1010927

33. Giasson BI, Murray IVJ, Trojanowski JQ, Lee VMY (2001) A Hydrophobic Stretch of 12 Amino Acid Residues in the Middle of α-Synuclein Is Essential for Filament Assembly. J Biol Chem 276:2380–2386. https://doi.org/10.1074/jbc.M008919200

34. Rodriguez JA, Ivanova MI, Sawaya MR, et al (2015) Structure of the toxic core of α-synuclein from invisible crystals. Nature 525:486–490. https://doi.org/10.1038/nature15368

35. Wong YC, Krainc D (2017) α-synuclein toxicity in neurodegeneration: Mechanism and therapeutic strategies. Nat Med 23:1–13. https://doi.org/10.1038/nm.4269

36. Kayed R, Pensalfini A, Margol L, et al (2009) Annular protofibrils area structurally and functionally distinct type of amyloid oligomer. J Biol Chem 284:4230–4237. https://doi.org/10.1074/jbc.M808591200

37. Möller H-J, Graeber MB (1998) The case described by Alois Alzheimer in 1911. Eur Arch Psychiatry Clin Neurosci 248:111–122. https://doi.org/10.1007/s004060050027

38. Ulep MG, Saraon SK, McLea S (2018) Alzheimer Disease. J Nurse Pract 14:129–135. https://doi.org/10.1016/j.nurpra.2017.10.014

39. Murphy MP, Levine H (2010) 10.3233@Jad-2010-1221. J Alzheimers Dis 19:311–323. https://doi.org/10.3233/JAD-2009-1221

40. Selkoe DJ, Schenk D (2003) Alzheimer's Disease: Molecular Understanding Predicts Amyloid-Based Therapeutics. Annu Rev Pharmacol Toxicol 43:545–584. https://doi.org/10.1146/annurev.pharmtox.43.100901.140248

41. Noble W, Hanger DP, Miller CCJ, Lovestone S (2013) The importance of tau phosphorylation for neurodegenerative diseases. Front Neurol 4 JUL:1–11. https://doi.org/10.3389/fneur.2013.00083

42. Mandelkow EM, Mandelkow E (1998) Tau in Alzheimer's disease. Trends Cell Biol 8:425–427. https://doi.org/10.1016/S0962-8924(98)01368-3

43. Cedazo-Mínguez A, Popescu BO, Blanco-Millán JM, et al (2003) Apolipoprotein E and β-amyloid (1-42) regulation of glycogen synthase kinase-3β. J Neurochem 87:1152–1164. https://doi.org/10.1046/j.1471-4159.2003.02088.x

44. Takashima A, Noguchi K, Sato K, et al (1993) tau Protein kinase I is essential for amyloid β-protein-induced neurotoxicity. Proc Natl Acad Sci U S A 90:7789–7793. https://doi.org/10.1073/pnas.90.16.7789

45. Scheltens P, Blennow K, Breteler MMB, et al (2016) Alzheimer's disease. Lancet 388:505–517. https://doi.org/10.1016/S0140-6736(15)01124-1

46. Woodgett JR (1990) Molecular cloning and expression of glycogen synthase kinase-3/factor A. EMBO J 9:2431–2438. https://doi.org/10.1002/j.1460-2075.1990.tb07419.x

47. Clavaguera F, Goedert M, Tolnay M (2010) Induction et propagation de la pathologie par la protéine tau chez un modèle murin de la maladie d'Alzheimer. Médecine/Sciences 26:121–124. https://doi.org/10.1051/medsci/2010262121

48. Aston NM, Bamborough P, Buckton JB, et al (2009) p38α mitogen-activated protein kinase inhibitors: Optimization of a series of biphenylamides to give a molecule suitable for clinical progression. J Med Chem 52:6257–6269. https://doi.org/10.1021/jm9004779

49. Rampa A, Gobbi S, Concetta Di Martino RM, et al (2017) Dual BACE-1/GSK-3β Inhibitors to Combat Alzheimer's Disease: A Focused Review. Curr Top Med Chem 17:3361–3369. https://doi.org/10.2174/1568026618666180112161406

50. Del Ser T, Steinwachs KC, Gertz HJ, et al (2013) Treatment of Alzheimer's disease with the GSK-3 inhibitor tideglusib: A pilot study. J Alzheimer's Dis 33:205–215. https://doi.org/10.3233/JAD-2012-120805

51. Matsunaga S, Fujishiro H, Takechi H (2019) Efficacy and safety of glycogen synthase kinase 3 inhibitors for Alzheimer's disease: A systematic review and meta-analysis. J Alzheimer's Dis 69:1031–1039. https://doi.org/10.3233/JAD-190256

52. McHardy SF, Wang HYL, McCowen S V., Valdez MC (2017) Recent advances in acetylcholinesterase Inhibitors and Reactivators: an update on the patent literature (2012-2015). Expert Opin Ther Pat 27:455–476. https://doi.org/10.1080/13543776.2017.1272571

53. Bowen DM, Smith CB, White P, Davison AN (1976) Neurotransmitter-related enzymes and indices of hypoxia in senile dementia and other abiotrophies. Brain 99:459–496. https://doi.org/10.1093/brain/99.3.459

54. Stedman E, Stedman E, Easson LH (1932) Choline-esterase. An enzyme present in the blood-serum of the horse. Biochem J 26:2056–2066. https://doi.org/10.1042/bj0262056

55. ORD MG, THOMPSON RH (1952) Pseudocholinesterase activity in the central nervous system. Biochem J 51:245–251. https://doi.org/10.1042/bj0510245

56. Koelle GB (1954) The histochemical localization of cholinesterases in the central nervous system of the rat. J Comp Neurol 100:211–235. https://doi.org/10.1002/cne.901000108

57. De Vriese C, Gregoire F, Lema-Kisoka R, et al (2004) Ghrelin degradation by serum and tissue homogenates: Identification of the cleavage sites. Endocrinology 145:4997–5005.

https://doi.org/10.1210/en.2004-0569

58. Gatley SJ (1991) Activities of the enantiomers of cocaine and some related compounds as substrates and inhibitors of plasma butyrylcholinesterase. Biochem Pharmacol 41:1249–1254. https://doi.org/10.1016/0006-2952(91)90665-R

59. Darvesh S, McDonald RS, Darvesh K V., et al (2006) On the active site for hydrolysis of aryl amides and choline esters by human cholinesterases. Bioorganic Med Chem 14:4586–4599. https://doi.org/10.1016/j.bmc.2006.02.021

60. Darvesh S, Hopkins DA, Geula C (2003) Neurobiology of butyrylcholinesterase. Nat Rev Neurosci 4:131–138. https://doi.org/10.1038/nrn1035

61. Ballard C, Greig N, Guillozet-Bongaarts A, et al (2005) Cholinesterases: Roles in the Brain During Health and Disease. Curr Alzheimer Res 2:307–318. https://doi.org/10.2174/1567205054367838

62. Geula C, Darvesh S (2004) Butyrylcholinesterase, cholinergic neurotransmission and the pathology of Alzheimer's disease. Drugs of Today 40:711–721. https://doi.org/10.1358/dot.2004.40.8.850473

63. Podoly E, Shalev DE, Shenhar-Tsarfaty S, et al (2009) The butyrylcholinesterase K variant confers structurally derived risks for Alzheimer pathology. J Biol Chem 284:17170–17179. https://doi.org/10.1074/jbc.M109.004952

64. Li B, Stribley JA, Ticu A, et al (2000) Abundant tissue butyrylcholinesterase and its possible function in the acetylcholinesterase knockout mouse. J Neurochem 75:1320–1331. https://doi.org/10.1046/j.1471-4159.2000.751320.x

65. Mesulam MM, Guillozet A, Shaw P, et al (2002) Acetylcholinesterase knockouts establish central cholinergic pathways and can use butyrylcholinesterase to hydrolyze acetylcholine. Neuroscience 110:627–639. https://doi.org/10.1016/S0306-4522(01)00613-3

66. Uyar T, Malterud KE, Anthonsen T (1978) Two new dihydrochalcones from Myrica gale. Phytochemistry 17:2011–2013. https://doi.org/10.1016/S0031-9422(00)88753-1

67. Niederhofer H (2007) Treating autism pharmacologically: Also tacrine might improve symptomatology in some cases [1]. J Child Neurol 22:1054. https://doi.org/10.1177/0883073807305857

68. Knapp MJ, Knopman DS, Solomon PR, et al (1994) A 30-Week Randomized Controlled Trial of High-Dose Tacrine in Patients With Alzheimer's Disease. JAMA J Am Med Assoc 271:985–991. https://doi.org/10.1001/jama.1994.03510370037029

69. O'Brien PJ, Irwin W, Diaz D, et al (2006) High concordance of drug-induced human hepatotoxicity with in vitro cytotoxicity measured in a novel cell-based model using high content screening. Arch Toxicol 80:580–604. https://doi.org/10.1007/s00204-006-0091-3

70. Shamsi A, Al Shahwan M, Ahamad S, et al (2020) Spectroscopic, calorimetric and molecular docking insight into the interaction of Alzheimer's drug donepezil with human transferrin: implications of Alzheimer's drug. J Biomol Struct Dyn 38:1094–1102. https://doi.org/10.1080/07391102.2019.1595728

71. Tatiparti K, Iyer A (2021) Autism Spectrum Disorder and Alzheimer's Disease

72. Khoury R, Rajamanickam J, Grossberg GT (2018) An update on the safety of current therapies for Alzheimer's disease: focus on rivastigmine. Ther Adv Drug Saf 9:171–178. https://doi.org/10.1177/2042098617750555

73. Kalia L V., Kalia SK, Lang AE (2015) Disease-modifying strategies for Parkinson's disease. Mov Disord 30:1442–1450. https://doi.org/10.1002/mds.26354

74. Frigerio R, Fujishiro H, Ahn TB, et al (2011) Incidental Lewy body disease: Do some cases represent a preclinical stage of dementia with Lewy bodies? Neurobiol Aging 32:857–863. https://doi.org/10.1016/j.neurobiolaging.2009.05.019

75. Singleton AB, Farrer M, Johnson J, et al (2003) α-Synuclein Locus Triplication Causes Parkinson's Disease. Science (80-) 302:841. https://doi.org/10.1126/science.1090278

76. Paisán-Ruíz C, Jain S, Evans EW, et al (2004) Cloning of the gene containing mutations that cause

PARK8-linked Parkinson's disease. Neuron 44:595–600. https://doi.org/10.1016/j.neuron.2004.10.023

77. Zimprich A, Biskup S, Leitner P, et al (2004) Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron 44:601–607. https://doi.org/10.1016/j.neuron.2004.11.005

78. Trinh J, Guella I, Farrer MJ (2014) Disease penetrance of late-onset parkinsonism: A meta-analysis. JAMA Neurol 71:1535–1539. https://doi.org/10.1001/jamaneurol.2014.1909

79. Lin X, Parisiadou L, Gu XL, et al (2009) Leucine-Rich Repeat Kinase 2 Regulates the Progression of Neuropathology Induced by Parkinson's-Disease-Related Mutant α-synuclein. Neuron 64:807–827. https://doi.org/10.1016/j.neuron.2009.11.006

80. Daher JPL, Abdelmotilib HA, Hu X, et al (2015) Leucine-rich repeat kinase 2 (LRRK2) pharmacological inhibition abates α-synuclein gene-induced neurodegeneration. J Biol Chem 290:19433–19444. https://doi.org/10.1074/jbc.M115.660001

81. Daher JPL, Volpicelli-Daley LA, Blackburn JP, et al (2014) Abrogation of α-synuclein -mediated dopaminergic neurodegeneration in LRRK2-deficient rats. Proc Natl Acad Sci U S A 111:9289–9294. https://doi.org/10.1073/pnas.1403215111

82. Volpicelli-Daley LA, Abdelmotilib H, Liu Z, et al (2016) G2019s-LRRK2 expression augments α-synuclein sequestration into inclusions in neurons. J Neurosci 36:7415–7427. https://doi.org/10.1523/JNEUROSCI.3642-15.2016

83. Denton T, Howard BD (1987) A Dopaminergic Cell Line Variant Resistant to the Neurotoxin 1-Methyl-4-Phenyl-1,2,3,6-Tetrahydropyridine. J Neurochem 49:622–630. https://doi.org/10.1111/j.1471-4159.1987.tb02909.x

84. Hospital PH, Turnbull DM, Wg C, et al Required Re-Operation. Methods Biochem Anal 000:550–550

85. Gray R, Ives N, Rick C, et al (2014) Long-term effectiveness of dopamine agonists and monoamine oxidase B inhibitors compared with levodopa as initial treatment for Parkinson's disease (PD MED): A large, open-label, pragmatic randomised trial. Lancet 384:1196–1205. https://doi.org/10.1016/S0140-6736(14)60683-8

86. Chen SY, Tsai ST (2010) The epidemiology of Parkinson's disease. Tzu Chi Med J 22:73–81. https://doi.org/10.1016/S1016-3190(10)60044-4

87. Poewe W, Antonini A (2015) Novel formulations and modes of delivery of levodopa. Mov Disord 30:114–120. https://doi.org/10.1002/mds.26078

88. Seppi K, Weintraub D, Coelho M, et al (2011) The movement disorder society evidence-based medicine review update: Treatments for the non-motor symptoms of Parkinson's disease. Mov Disord 26:42–80. https://doi.org/10.1002/mds.23884

89. Ferreira JJ, Lees A, Rocha JF, et al (2016) Opicapone as an adjunct to levodopa in patients with Parkinson's disease and end-of-dose motor fluctuations: A randomised, double-blind, controlled trial. Lancet Neurol 15:154–165. https://doi.org/10.1016/S1474-4422(15)00336-1

90. Schapira AHV (2011) Monoamine oxidase B inhibitors for the treatment of parkinsons disease: A review of symptomatic and potential disease-modifying effects. CNS Drugs 25:1061–1071. https://doi.org/10.2165/11596310-000000000-00000

91. Schapira AHV, Fox SH, Hauser RA, et al (2017) Assessment of safety and efficacy of safinamide as a levodopa adjunct in patients with Parkinson disease and motor fluctuations a randomized clinical trial. JAMA Neurol 74:216–224. https://doi.org/10.1001/jamaneurol.2016.4467

92. Connolly BS, Lang AE (2014) Pharmacological treatment of Parkinson disease: A review. JAMA - J Am Med Assoc 311:1670–1683. https://doi.org/10.1001/jama.2014.3654

93. Jankovic J, Poewe W (2012) Therapies in Parkinson's disease. Curr Opin Neurol 25:433–447. https://doi.org/10.1097/WCO.0b013e3283542fc2

94. Zesiewicz TA (2019) Parkinson Disease. Contin Lifelong Learn Neurol 25:896–918.

https://doi.org/10.1212/CON.0000000000000764

95. Frankel JP, Lees AJ, Kempster PA, Stern GM (1990) Subcutaneous apomorphine in the treatment of Parkinson's disease. J Neurol Neurosurg Psychiatry 53:96–101. https://doi.org/10.1136/jnnp.53.2.96

96. Khatib S, Nerya O, Musa R, et al (2005) Chalcones as potent tyrosinase inhibitors: The importance of a 2,4-substituted resorcinol moiety. Bioorganic Med Chem 13:433–441. https://doi.org/10.1016/j.bmc.2004.10.010

97. Nowakowska Z (2007) A review of anti-infective and anti-inflammatory chalcones. Eur J Med Chem 42:125–137. https://doi.org/10.1016/j.ejmech.2006.09.019

98. K. Sahu N, S. Balbhadra S, Choudhary J, V. Kohli D (2012) Exploring Pharmacological Significance of Chalcone Scaffold: A Review. Curr Med Chem 19:209–225. https://doi.org/10.2174/092986712803414132

99. Evranos Aksöz B, Ertan R (2011) Chemical and structural properties of chalcones I. Fabad J Pharm Sci 36:223–242

100. Kashman Y, Green D, Garcia C, Garcia Arevalos D (1991) Verrucoside, A new cytotoxic pregnane glycoside from a gorgonian eunicella verrucosa. J Nat Prod 54:1651–1655. https://doi.org/10.1021/np50078a026

101. Palleros DR (2004) Solvent-free synthesis of chalcones. J Chem Educ 81:1345–1347. https://doi.org/10.1021/ed081p1345

102. Toda F, Tanaka K, Hamai K (1990) Aldol condensations in the absence of solvent: Acceleration of the reaction and enhancement of the stereoselectivity. J Chem Soc Perkin Trans 1 3207–3209. https://doi.org/10.1039/p19900003207

103. Stork T, Answer E, Dilemma H, York N (1995) Book re v iews 513. 119:513–514

104. Singh P, Anand A, Kumar V (2014) Recent developments in biological activities of chalcones: A mini review. Eur J Med Chem 85:758–777. https://doi.org/10.1016/j.ejmech.2014.08.033

105. Lakshminarayanan B, Baek SC, Lee JP, et al (2019) Ethoxylated Head of Chalcones as a New Class of Multi-Targeted MAO Inhibitors. ChemistrySelect 4:6614–6619. https://doi.org/10.1002/slct.201901093

106. Rehuman NA, Oh JM, Nath LR, et al (2021) Halogenated Coumarin–Chalcones as Multifunctional Monoamine Oxidase-B and Butyrylcholinesterase Inhibitors. ACS Omega. https://doi.org/10.1021/acsomega.1c04252

107. Mathew B, Mathew GE, Uçar G, et al (2015) Development of fluorinated methoxylated chalcones as selective monoamine oxidase-B inhibitors: Synthesis, biochemistry and molecular docking studies. Bioorg Chem 62:22–29. https://doi.org/10.1016/j.bioorg.2015.07.001

108. Sasidharan R, Manju SL, Uçar G, et al (2016) Identification of Indole-Based Chalcones: Discovery of a Potent, Selective, and Reversible Class of MAO-B Inhibitors. Arch Pharm (Weinheim) 627–637. https://doi.org/10.1002/ardp.201600088

109. Mathew B, Uçar G, Rapheal C, et al (2017) Characterization of Thienylchalcones as hMAO-B Inhibitors: Synthesis, Biochemistry and Molecular Dynamics Studies. ChemistrySelect 2:11113–11119. https://doi.org/10.1002/slct.201702141

110. Robinson SJ, Petzer JP, Petzer A, et al (2013) Selected furanochalcones as inhibitors of monoamine oxidase. Bioorganic Med Chem Lett 23:4985–4989. https://doi.org/10.1016/j.bmcl.2013.06.050

111. Mathew B, Haridas A, Uçar G, et al (2016) Exploration of chlorinated thienyl chalcones: A new class of monoamine oxidase-B inhibitors. Int J Biol Macromol 91:680–695. https://doi.org/10.1016/j.ijbiomac.2016.05.110

112. Nair AS, Oh J, Koyiparambath VP, et al (2021) Development of Halogenated Pyrazolines as Selective. Mol Artic 26:1–17. https://doi.org/https://doi.org/10.3390/molecules26113264

113. Badr G, Gul HI, Yamali C, et al (2018) Curcumin analogue 1,5-bis(4-hydroxy-3-((4-methylpiperazin-1-yl)methyl)phenyl)penta-1,4-dien-3-one mediates growth arrest and apoptosis by targeting the

PI3K/AKT/mTOR and PKC-theta signaling pathways in human breast carcinoma cells. Bioorg Chem 78:46–57. https://doi.org/10.1016/j.bioorg.2018.03.006

114. Yamali C, Gul HI, Sakagami H, Supuran CT (2016) Synthesis and bioactivities of halogen bearing phenolic chalcones and their corresponding bis Mannich bases. J Enzyme Inhib Med Chem 31:125–131. https://doi.org/10.1080/14756366.2016.1221825

115. Krallinger M, Rabal O, Lourenço A, et al (2017) Information retrieval and text mining technologies for chemistry. Chem Rev 117:7673–7761. https://doi.org/10.1021/acs.chemrev.6b00851

116. Pan Y, Chen Y, Yu X, et al (2012) The synthesis of a novel chalcone and evaluation for anti-free radical activity and antagonizing the learning impairments in Alzheimer's model. Cell Physiol Biochem 29:949–958. https://doi.org/10.1159/000188336

117. Zhang X, Rakesh KP, Bukhari SNA, et al (2018) Multi-targetable chalcone analogs to treat deadly Alzheimer's disease: Current view and upcoming advice. Bioorg Chem 80:86–93. https://doi.org/10.1016/j.bioorg.2018.06.009

118. Kang L, Gao XH, Liu HR, et al (2018) Structure–activity relationship investigation of coumarin–chalcone hybrids with diverse side-chains as acetylcholinesterase and butyrylcholinesterase inhibitors. Mol Divers 22:893–906. https://doi.org/10.1007/s11030-018-9839-y

119. Shah MS, Khan SU, Ejaz SA, et al (2017) Cholinesterases inhibition and molecular modeling studies of piperidyl-thienyl and 2-pyrazoline derivatives of chalcones. Biochem Biophys Res Commun 482:615–624. https://doi.org/10.1016/j.bbrc.2016.11.082

120. Alipour M, Khoobi M, Foroumadi A, et al (2012) Novel coumarin derivatives bearing N-benzyl pyridinium moiety: Potent and dual binding site acetylcholinesterase inhibitors. Bioorganic Med Chem 20:7214–7222. https://doi.org/10.1016/j.bmc.2012.08.052

121. Koch G (2017) Medicinal Chemistry. Chimia (Aarau) 71:643. https://doi.org/10.2307/j.ctvnwc0d0.18

122. Paquin F, Rivnay J, Salleo A, et al (2015) Multi-phase semicrystalline microstructures drive exciton dissociation in neat plastic semiconductors. J Mater Chem C 3:10715–10722. https://doi.org/10.1039/b000000x

123. Kang JE, Cho JK, Curtis-Long MJ, et al (2013) Inhibitory evaluation of sulfonamide chalcones on β-secretase and acylcholinesterase. Molecules 18:140–153. https://doi.org/10.3390/molecules18010140

124. Liu HR, Zhou C, Fan HQ, et al (2015) Novel Potent and Selective Acetylcholinesterase Inhibitors as Potential Drugs for the Treatment of Alzheimer's Disease: Synthesis, Pharmacological Evaluation, and Molecular Modeling of Amino-Alkyl-Substituted Fluoro-Chalcones Derivatives. Chem Biol Drug Des 86:517–522. https://doi.org/10.1111/cbdd.12514

125. Liu H, Fan H, Gao X, et al (2016) Design, synthesis and preliminary structure-activity relationship investigation of nitrogen-containing chalcone derivatives as acetylcholinesterase and butyrylcholinesterase inhibitors: A further study based on Flavokawain B Mannich base derivatives. J Enzyme Inhib Med Chem 31:580–589. https://doi.org/10.3109/14756366.2015.1050009

126. Mathew B, Oh JM, Baty RS, et al (2021) Piperazine-substituted chalcones: a new class of MAO-B, AChE, and BACE-1 inhibitors for the treatment of neurological disorders. Environ Sci Pollut Res 28:38855–38866. https://doi.org/10.1007/s11356-021-13320-y

127. Wang L, Wang Y, Tian Y, et al (2017) Design, synthesis, biological evaluation, and molecular modeling studies of chalcone-rivastigmine hybrids as cholinesterase inhibitors. Bioorganic Med Chem 25:360–371. https://doi.org/10.1016/j.bmc.2016.11.002

128. Kwon YE, Park JY, No KT, et al (2007) Synthesis, in vitro assay, and molecular modeling of new piperidine derivatives having dual inhibitory potency against acetylcholinesterase and Aβ1-42 aggregation for Alzheimer's disease therapeutics. Bioorganic Med Chem 15:6596–6607. https://doi.org/10.1016/j.bmc.2007.07.003

129. López SN, Castelli M V., Zacchino SA, et al (2001) In vitro antifungal evaluation and structure-activity relationships of a new series of chalcone derivatives and synthetic analogues, with inhibitory properties against polymers of the fungal cell wall. Bioorganic Med Chem 9:1999–2013.

https://doi.org/10.1016/S0968-0896(01)00116-X

130. Boeck P, Leal PC, Yunes RA, et al (2005) Antifungal activity and studies on mode of action of novel xanthoxyline-derived chalcones. Arch Pharm (Weinheim) 338:87–95. https://doi.org/10.1002/ardp.200400929

131. Yayli N, Üçüncü O, Aydin E, et al (2005) Stereoselective photochemistry of heteroaryl chalcones in solution and the antioxidant activities. J Photochem Photobiol A Chem 169:229–234. https://doi.org/10.1016/j.jphotochem.2004.06.015

132. Cabrera M, Simoens M, Falchi G, et al (2007) Synthetic chalcones, flavanones, and flavones as antitumoral agents: Biological evaluation and structure-activity relationships. Bioorganic Med Chem 15:3356–3367. https://doi.org/10.1016/j.bmc.2007.03.031

133. Nakamura C, Kawasaki N, Miyataka H, et al (2002) Synthesis and biological activities of fluorinated chalcone derivatives. Bioorganic Med Chem 10:699–706. https://doi.org/10.1016/S0968-0896(01)00319-4

134. Tabata K, Motani K, Takayanagi N, et al (2005) Xanthoangelol, a major chalcone constituent of Angelica keiskei, induces apoptosis in neuroblastoma and leukemia cells. Biol Pharm Bull 28:1404–1407. https://doi.org/10.1248/bpb.28.1404

135. Chimenti F, Fioravanti R, Bolasco A, et al (2009) Chalcones: A valid scaffold for monoamine oxidases inhibitors. J Med Chem 52:2818–2824. https://doi.org/10.1021/jm801590u

136. Cheng ZJ, Lin CN, Hwang TL, Teng CM (2001) Broussochalcone A, a potent antioxidant and effective suppressor of inducible nitric oxide synthase in lipopolysaccharide-activated macrophages. Biochem Pharmacol 61:939–946. https://doi.org/10.1016/S0006-2952(01)00543-3

137. (2002) Synthesis and Inhibitory Activity of Dimethylamino-Chalcone Derivatives on the Induction of Nitric Oxide Synthase. Eur J Med Chem 37:46100. https://doi.org/10.1002/chin.200301095

138. Takahashi T, Takasuka N, Iigo M, et al (2004) Isoliquiritigenin, a flavonoid from licorice, reduces prostaglandin E2 and nitric oxide, causes apoptosis, and suppresses aberrant crypt foci development. Cancer Sci 95:448–453. https://doi.org/10.1111/j.1349-7006.2004.tb03230.x

139. Israf DA, Khaizurin TA, Syahida A, et al (2007) Cardamonin inhibits COX and iNOS expression via inhibition of p65NF- □ B nuclear translocation and I □ -B phosphorylation in RAW 264 . 7 macrophage cells. Mol Immunol 44:673–679. https://doi.org/10.1016/j.molimm.2006.04.025

140. Kim YH, Kim J, Park H, Kim HP (2007) Anti-inflammatory activity of the synthetic chalcone derivatives: Inhibition of inducible nitric oxide synthase-catalyzed nitric oxide production from lipopolysaccharide-treated RAW 264.7 cells. Biol Pharm Bull 30:1450–1455. https://doi.org/10.1248/bpb.30.1450

141. Ngameni B, Watchueng J, Boyom FF, et al (2007) Antimalarial prenylated chalcones from the twigs of Dorstenia barteri var. subtriangularis. Arkivoc 2007:116–123. https://doi.org/10.3998/ark.5550190.0008.d14

142. Santos L, Pedrosa RC, Correa R, et al (2006) Biological evaluation of chalcones and analogues as hypolipidemic agents. Arch Pharm (Weinheim) 339:541–546. https://doi.org/10.1002/ardp.200600034

143. Orallo F (2008) Trans-Resveratrol: A Magical Elixir of Eternal Youth? Curr Med Chem 15:1887–1898. https://doi.org/10.2174/092986708785132951

144. Conaway RC, Brower CS, Conaway JW, Ulrich HD (1999) Aging and Sirtuins. Trends Cell Biol Sci Eukaryot Cell Annu Rev Cell Dev Biol 9:107–112. https://doi.org/10.1038/nature01965.1.

145. DiMasi JA, Hansen RW, Grabowski HG (2003) The price of innovation: New estimates of drug development costs. J Health Econ 22:151–185. https://doi.org/10.1016/S0167-6296(02)00126-1

146. Hubbard RE (2005) 3D structure and the drug-discovery process. Mol Biosyst 1:391–406. https://doi.org/10.1039/b514814f

147. Hughes JP, Rees SS, Kalindjian SB, Philpott KL (2011) Principles of early drug discovery. Br J Pharmacol 162:1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

148. Murcko M, Nagaraj PH, Marshall G, et al (1998) Virtual screening-an overview Related papers Virt ual screening in drug discovery-a comput at ional perspect ive. Drug Discov Today 3:. https://doi.org/https://doi.org/10.1016/S1359-6446(97)01163-X

149. Enyedy IJ, Egan WJ (2008) Can we use docking and scoring for hit-to-lead optimization? J Comput Aided Mol Des 22:161–168. https://doi.org/10.1007/s10822-007-9165-4

150. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 64:4–17. https://doi.org/10.1016/j.addr.2012.09.019

151. Veber DF, Johnson SR, Cheng HY, et al (2002) Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 45:2615–2623. https://doi.org/10.1021/jm020017n

152. Johnson TW, Dress KR, Edwards M (2009) Using the Golden Triangle to optimize clearance and oral absorption. Bioorganic Med Chem Lett 19:5560–5564. https://doi.org/10.1016/j.bmcl.2009.08.045

153. Aparoy P, Kumar Reddy K, Reddanna P (2012) Structure and Ligand Based Drug Design Strategies in the Development of Novel 5- LOX Inhibitors. Curr Med Chem 19:3763–3778. https://doi.org/10.2174/092986712801661112

154. Speck-Planche A, Luan F, N.D.S. Cordeiro M (2012) Role of Ligand-Based Drug Design Methodologies toward the Discovery of New Anti- Alzheimer Agents: Futures Perspectives in Fragment-Based Ligand Design. Curr Med Chem 19:1635–1645. https://doi.org/10.2174/092986712799945058

155. Lavecchia A, Fominaya J, Mat eos G, et al (2013) Virtual screening strategies in drug discovery: a critical review Related papers Int egrat ion of virt ual and physical screening Eric Prossnit z Discovery of Mit ogen-Act ivat ed Prot ein Kinase-Int eract ing Kinase 1 Inhibit ors by a Comprehensive Fr… V. Curr Med Chem 20:2839–2860

156. Perez-Pineiro R, Burgos A, Jones DC, et al (2009) Development of a novel virtual screening cascade protocol to identify potential trypanothione reductase inhibitors. J Med Chem 52:1670–1680. https://doi.org/10.1021/jm801306g

157. Mason J, Good A, Martin E (2005) 3-D Pharmacophores in Drug Discovery. Curr Pharm Des 7:567–597. https://doi.org/10.2174/1381612013397843

158. Gao Q, Yang L, Zhu Y (2010) Pharmacophore Based Drug Design Approach as a Practical Process in Drug Discovery. Curr Comput Aided-Drug Des 6:37–49. https://doi.org/10.2174/157340910790980151

159. Muegge I, Mukherjee P (2016) An overview of molecular fingerprint similarity search in virtual screening. Expert Opin Drug Discov 11:137–148. https://doi.org/10.1517/17460441.2016.1117070

160. Cereto-Massagué A, Ojeda MJ, Valls C, et al (2015) Molecular fingerprint similarity search in virtual screening. Methods 71:58–63. https://doi.org/10.1016/j.ymeth.2014.08.005

161. Zhao D, Toba Y, Sugioka KI, Komori S (2006) New sea spray generation function for spume droplets. J Geophys Res Ocean 111:. https://doi.org/10.1029/2005JC002960

162. Vuorinen A, Schuster D (2015) Methods for generating and applying pharmacophore models as virtual screening filters and for bioactivity profiling. Methods 71:113–134. https://doi.org/10.1016/j.ymeth.2014.10.013

163. Langer T, Eder M, Hoffmann RD, et al (2004) Lead identification for modulators of multidrug resistance based on in silico screening with a pharmacophoric feature model. Arch Pharm (Weinheim) 337:317–327. https://doi.org/10.1002/ardp.200300817

164. Chen IJ, Foloppe N (2008) Conformational sampling of druglike molecules with MOE and catalyst: Implications for pharmacophore modeling and virtual screening. J Chem Inf Model 48:1773–1791. https://doi.org/10.1021/ci800130k

165. Wolber G, Dornhofer AA, Langer T (2006) Efficient overlay of small organic molecules using 3D pharmacophores. J Comput Aided Mol Des 20:773–788. https://doi.org/10.1007/s10822-006-9078-7

166. Wolber G, Langer T (2005) LigandScout: 3-D pharmacophores derived from protein-bound ligands and

their use as virtual screening filters. J Chem Inf Model 45:160–169. https://doi.org/10.1021/ci049885e

167. Tropsha A (2010) Best Practices for QSAR Model Development , Validation , and Exploitation. Mol Informatics, 29:476–488. https://doi.org/10.1002/minf.201000061

168. Jollife IT, Cadima J (2016) Principal component analysis: A review and recent developments. Philos Trans R Soc A Math Phys Eng Sci 374:. https://doi.org/10.1098/rsta.2015.0202

169. Topliss JG, Edwards RP (1979) Chance Factors in Studies of Quantitative Structure-Activity Relationships. J Med Chem 22:1238–1244. https://doi.org/10.1021/jm00196a017

170. Wold H (1966) Estimation of Principal Components and Related Models by Iterative Least squares In Multivariate Analysis. Multivar Anal 391–420

171. Höskuldsson A (1988) PLS regression methods. J Chemom 2:211–228. https://doi.org/10.1002/cem.1180020306

172. Wold S, Sjöström M, Eriksson L (2001) PLS-regression: A basic tool of chemometrics. Chemom Intell Lab Syst 58:109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

173. Mitra I, Saha A, Roy K (2010) Exploring quantitative structure-activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol Simul 36:1067–1079. https://doi.org/10.1080/08927022.2010.503326

174. Grandgirard J, Poinsot D, Krespi L, et al (2002) Costs of secondary parasitism in the facultative hyperparasitoid Pachycrepoideus dubius: Does host size matter? Entomol Exp Appl 103:239–248. https://doi.org/10.1023/A

175. King RD, Hirst JD, Sternberg MJE (1993) New approaches to QSAR: Neural networks and machine learning. Perspect Drug Discov Des 1:279–290. https://doi.org/10.1007/BF02174529

176. Huuskonen J, Salo M, Taskinen J (1997) Neural network modeling for estimation of the aqueous solubility of structurally related drugs. J Pharm Sci 86:450–454. https://doi.org/10.1021/js960358m

177. Livingstone DJ, Manallack DT, Tetko I V. (1997) Data modelling with neural networks: Advantages and limitations. J Comput Aided Mol Des 11:135–142. https://doi.org/10.1023/A:1008074223811

178. Guha R (2008) On the interpretation and interpretability of quantitative structure-activity relationship models. J Comput Aided Mol Des 22:857–871. https://doi.org/10.1007/s10822-008-9240-5

179. Eriksson L, Jaworska J, Worth AP, et al (2003) Regulatory Acceptance of (Q) SARs | Mini-Monograph Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification-and Regression-Based QSARs. 111:1361–1375. https://doi.org/10.1289/ehp.5758

180. Software BIA (2005) Bioinformatics applications note. Bioinformatics 21:1745–1746

181. El Aissouq A, Chedadi O, Kasmi R, et al (2021) Molecular Modeling Studies of C-Glycosylfavone Derivatives as GSK-3β Inhibitors Based on QSAR and Docking Analysis. J Solution Chem 50:808–822. https://doi.org/10.1007/s10953-021-01083-6

182. El Aissouq A, Toufik H (2019) QSAR study of isonicotinamides derivatives as Alzheimr ' s disease inhibitors using PLS-R and ANN methods. 2019 Int Conf Intell Syst Adv Comput Sci 1–7. https://doi.org/10.1109/ISACS48493.2019.9068919

183. Roy K (2007) On some aspects of validation of predictive quantitative structure–activity relationship models. Expert Opin Drug Discov 2:1567–1577. https://doi.org/10.1517/17460441.2.12.1567

184. Cramer RD, Patterson DE, Bunce JD (1988) Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. J Am Chem Soc 110:5959–5967. https://doi.org/10.1021/ja00226a005

185. Xue C Bin, Zhang L, Luo WC, et al (2007) 3D-QSAR and molecular docking studies of benzaldehyde thiosemicarbazone, benzaldehyde, benzoic acid, and their derivatives as phenoloxidase inhibitors. Bioorganic Med Chem 15:2006–2015. https://doi.org/10.1016/j.bmc.2006.12.038

186. Semerjian L, Ayoub GM (2003) High-pH-magnesium coagulation-flocculation in wastewater treatment.

Adv Environ Res 7:389–403. https://doi.org/10.1016/S1093-0191(02)00009-6

187. Tervo AJ, Nyrönen TH, Rönkkö T, Poso A (2004) Comparing the quality and predictiveness between 3D QSAR models obtained from manual and automated alignment. J Chem Inf Comput Sci 44:807–816. https://doi.org/10.1021/ci0342268

188. Wildman SA, Crippen GM (2003) Validation of DAPPER for 3D QSAR: Conformational search and chirality metric. J Chem Inf Comput Sci 43:629–636. https://doi.org/10.1021/ci0256081

189. Tan BJ, Klabunde KJ, Tanaka T, et al (1988) An EXAFS Study of Co-Mn/SiO2Bimetallic Solvated Metal Atom Dispersed (SMAD) Catalysts. J Am Chem Soc 110:5951–5958. https://doi.org/10.1021/ja00226a004

190. Montgomery JA, Niwas S, Rose JD, et al (1993) Structure-Based Design of Inhibitors of Purine Nucleoside Phosphorylase. 1. 9-(Arylmethyl) Derivatives of 9-Deazaguanine. J Med Chem 36:55–69. https://doi.org/10.1021/jm00053a008

191. Gohlke H, Klebe G (2002) Drugscore meets CoMFA: Adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein. J Med Chem 45:4153–4170. https://doi.org/10.1021/jm020808p

192. Clark M, Cramer RD (1993) The Probability of Chance Correlation Using Partial Least Squares (PLS). Quant Struct Relationships 12:137–145. https://doi.org/10.1002/qsar.19930120205

193. Xu Y, Goodacre R (2018) On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J Anal Test 2:249–262. https://doi.org/10.1007/s41664-018-0068-2

194. Teichmann SA, Chothia C, Gerstein M (1999) Advances in structural genomics. Curr Opin Struct Biol 9:390–399. https://doi.org/10.1016/S0959-440X(99)80053-0

195. Capener CE, Shrivastava IH, Ranatunga KM, et al (2000) Homology modeling and molecular dynamics simulation studies of an inward rectifier potassium channel. Biophys J 78:2929–2942. https://doi.org/10.1016/S0006-3495(00)76833-0

196. Floudas CA, Fung HK, McAllister SR, et al (2006) Advances in protein structure prediction and de novo protein design: A review. Chem Eng Sci 61:966–988. https://doi.org/10.1016/j.ces.2005.04.009

197. Lengauer T, Rarey M (1996) Computational methods for biomolecular docking. Curr Opin Struct Biol 6:402–406. https://doi.org/10.1016/S0959-440X(96)80061-3

198. Santana Azevedo L, Pretto Moraes F, Morrone Xavier M, et al (2013) Recent Progress of Molecular Docking Simulations Applied to Development of Drugs. Curr Bioinform 7:352–365. https://doi.org/10.2174/157489312803901063

199. Breda A, Basso L, Santos D, de Azevedo Jr. W (2008) Virtual Screening of Drugs: Score Functions, Docking, and Drug Design. Curr Comput Aided-Drug Des 4:265–272. https://doi.org/10.2174/157340908786786047

200. De Azevedo Jr. W (2010) MolDock Applied to Structure-Based Virtual Screening. Curr Drug Targets 11:327–334. https://doi.org/10.2174/138945010790711941

201. Allen WJ, Balius TE, Mukherjee S, et al (2015) DOCK 6: Impact of new features and current docking performance. J Comput Chem 36:1132–1156. https://doi.org/10.1002/jcc.23905

202. Allouche A (2012) Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares. J Comput Chem 32:174–182. https://doi.org/10.1002/jcc

203. Liebeschuetz JW, Cole JC, Korb O (2012) Pose prediction and virtual screening performance of GOLD scoring functions in a standardized test. J Comput Aided Mol Des 26:737–748. https://doi.org/10.1007/s10822-012-9551-4

204. Korb O, Stützle T, Exner TE (2009) Empirical scoring functions for advanced Protein-Ligand docking with PLANTS. J Chem Inf Model 49:84–96. https://doi.org/10.1021/ci800298z

205. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD (2015) Molecular docking and structure-based

drug design strategies. Molecules 20:13384–13421. https://doi.org/10.3390/molecules200713384

206. Cross JB, Thompson DC, Rai BK, et al (2009) Comparison of several molecular docking programs: Pose prediction and virtual screening accuracy. J Chem Inf Model 49:1455–1474. https://doi.org/10.1021/ci900056c

207. Fischer E (1894) Einfluss der Configuration auf die Wirkung der Enzyme. Berichte der Dtsch Chem Gesellschaft 27:2985–2993. https://doi.org/10.1002/cber.18940270364

208. ASANO N, TAKEUCHI M, KAMEDA Y, et al (1990) Trehalase inhibitors, validoxylamine A and related compounds as insecticides. J Antibiot (Tokyo) 43:722–726. https://doi.org/10.7164/antibiotics.43.722

209. Watson AA, Fleet GW, Asano N, et al (2001) Polyhydroxylated alkaloids -- natural occurrence and therapeutic applications. Phytochemistry 56:265–95. https://doi.org/https://doi.org/10.1016/S0031-9422(00)00451-9

210. Vogt AD, Di Cera E (2013) Conformational Selection Is a Dominant Mechanism of Ligand Binding. Biochemistry 52:5723–5729. https://doi.org/10.1021/bi400929b

211. Legler G, Jülich E (1984) Synthesis of 5-amino-5-deoxy-D-mannopyranose and 1,5-dideoxy-1,5-imino-D-mannitol, and inhibition of alpha- and beta-D-mannosidases. Carbohydr Res 128:61–72

212. Sousa SF, Fernandes PA, Ramos MJ (2006) Protein-ligand docking: Current status and future challenges. Proteins Struct Funct Bioinforma 65:15–26. https://doi.org/10.1002/prot.21082

213. Yuriev E, Agostino M, Ramsland PA (2011) Challenges and advances in computational docking: 2009 in review. J Mol Recognit 24:149–164. https://doi.org/10.1002/jmr.1077

214. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) DOCKING AND SCORING IN VIRTUAL SCREENING FOR DRUG DISCOVERY : METHODS AND APPLICATIONS. 3:. https://doi.org/10.1038/nrd1549

215. Brooijmans N, Kuntz ID (2003) Molecular Recognition and Docking Algorithms. Annu Rev Biophys Biomol Struct 32:335–373. https://doi.org/10.1146/annurev.biophys.32.110601.142532

216. Duffy BC, Zhu L, Decornez H, Kitchen DB (2012) Early phase drug discovery: Cheminformatics and computational techniques in identifying lead series. Bioorganic Med Chem 20:5324–5342. https://doi.org/10.1016/j.bmc.2012.04.062

217. Olivei A (1976) Surface imaging by scanning-laser-induced voltage. J Phys D Appl Phys 9:183–195. https://doi.org/10.1088/0022-3727/9/2/008

218. Cerutti DS, Case DA (2010) Multi-level ewald: A hybrid multigrid/fast fourier transform approach to the electrostatic particle-mesh problem. J Chem Theory Comput 6:443–458. https://doi.org/10.1021/ct900522g

219. Nosé S (1991) Molecular Dynamics Simulations at Constant Temperature and Pressure. Comput Simul Mater Sci 2384:21–41. https://doi.org/10.1007/978-94-011-3546-7_2

220. Berendsen HJC, Postma JPM, Van Gunsteren WF, et al (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690. https://doi.org/10.1063/1.448118

221. Ditchfield R, Hehre WJ, Pople JA (1971) Self-consistent molecular-orbital methods. IX. An extended gaussian-type basis for molecular-orbital studies of organic molecules. J Chem Phys 54:720–723. https://doi.org/10.1063/1.1674902

222. Zwanzig RW (1954) High†ï¿½Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. J Chem Phys 22:1420–1426. https://doi.org/10.1063/1.1740409

223. Lybrand TP, McCammon JA, Wipff G (1986) Theoretical calculation of relative binding affinity in host-guest systems. Proc Natl Acad Sci U S A 83:833–835. https://doi.org/10.1073/pnas.83.4.833

224. Kollman PA, Massova I, Reyes C, et al (2000) Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. Acc Chem Res 33:889–897. https://doi.org/10.1021/ar000033j

225. Wang J, Morin P, Wang W, Kollman PA (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J Am Chem Soc 123:5221–5230. https://doi.org/10.1021/ja003834q

226. Fortmeyer R (2007) The zero effect. Archit Rec 195:153. https://doi.org/10.1038/news070604-1

227. Drews J (2009) Drug Discovery : A Historical Perspective. 1960:. https://doi.org/10.1126/science.287.5460.1960

228. Siramshetty VB, Nickel J, Omieczynski C, et al (2016) WITHDRAWN — a resource for withdrawn and discontinued drugs. 44:1080–1086. https://doi.org/10.1093/nar/gkv1192

229. Keller TH, Pichota A, Yin Z (2006) A practical view of ' druggability .' Curr Opin Chem Biol 10:357–361. https://doi.org/10.1016/j.cbpa.2006.06.014

230. Ursu O, Rayan A, Goldblum A, Oprea T (2011) Understanding drug-likeness. https://doi.org/10.1002/wcms.52

231. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 64:3–25. https://doi.org/https://doi.org/10.1016/j.addr.2012.09.019

232. Ghose AK, Viswanadhan VN, Wendoloski JJ (1999) A Knowledge-Based Approach in Designing Combinatorial or Medicinal Chemistry Libraries for Drug Discovery . 1 . A Qualitative and Quantitative Characterization of Known Drug Databases. 55–68. https://doi.org/10.1021/cc9800071

233. Zernov V V, Balakin K V, Ivaschenko AA, et al (2003) Drug Discovery Using Support Vector Machines. The Case Studies of Drug-likeness, Agrochemical-likeness, and Enzyme Inhibition Predictions. J Chem Inf Comput Sci 43:2048–2056. https://doi.org/https://doi.org/10.1021/ci0340916

234. Byvatov E, Fechner U, Sadowski J, Schneider G (2003) Comparison of Support Vector Machine and Artificial Neural Network Systems for Drug / Nondrug Classification. J Chem Inf Comput Sci 46:1882–1889. https://doi.org/https://doi.org/10.1021/ci0341161

235. Sadowski J, Kubinyi H (1998) A Scoring Scheme for Discriminating between Drugs and Nondrugs. J Med Chem 2623:3325–3329. https://doi.org/https://doi.org/10.1021/jm9706776

236. Ertl P, Rohde B, Selzer P (2000) Fast Calculation of Molecular Polar Surface Area as a Sum of Fragment-Based Contributions and Its Application to the Prediction of Drug Transport Properties. J Med Chem 43:3714–3717. https://doi.org/https://doi.org/10.1021/jm000942e

237. Jancova P, Anzenbacher P, Anzenbacherova E (2010) Phase II drug metabolizing enzymes. Biomed Pap 154:103–116. https://doi.org/10.5507/bp.2010.017

238. Sinko PJ (2014) Predicting Fraction Dose Absorbed in Humans Using a Macroscopic Mass Balance Approach. Pharm Res 8:979–988. https://doi.org/10.1023/A

239. Jr LGV (2009) In silico toxicology for the pharmaceutical sciences ☆. Toxicol Appl Pharmacol 241:356–370. https://doi.org/10.1016/j.taap.2009.08.022

240. Zauhar RJ, Moyna G, Tian LF, et al (2003) Shape Signatures: A New Approach to Computer-Aided Ligand- and Receptor-Based Drug Design. J Med Chem 46:5674–5690. https://doi.org/10.1021/jm030242k

241. Ballester PJ, Finn PW, Richards WG (2009) Ultrafast shape recognition: Evaluating a new ligand-based virtual screening technology. J Mol Graph Model 27:836–845. https://doi.org/10.1016/j.jmgm.2009.01.001

242. Delage B, Dashwood RH (2008) Dietary manipulation of histone structure and function. Annu Rev Nutr 28:347–366. https://doi.org/10.1146/annurev.nutr.28.061807.155354

243. Porcu M, Chiarugi A (2005) The emerging therapeutic potential of sirtuin-interacting drugs: From cell death to lifespan extension. Trends Pharmacol Sci 26:94–103. https://doi.org/10.1016/j.tips.2004.12.009

244. Yang E Bin, Zhang K, Cheng LY, Mack P (1998) Butein, a specific protein tyrosine kinase inhibitor.

Biochem Biophys Res Commun 245:435–438. https://doi.org/10.1006/bbrc.1998.8452

245. Funakoshi-Tago M, Okamoto K, Izumi R, et al (2015) Anti-inflammatory activity of flavonoids in Nepalese propolis is attributed to inhibition of the IL-33 signaling pathway. Int Immunopharmacol 25:189–198. https://doi.org/10.1016/j.intimp.2015.01.012

246. Park JY, Ko JA, Kim DW, et al (2016) Chalcones isolated from Angelica keiskei inhibit cysteine proteases of SARS-CoV. J Enzyme Inhib Med Chem 31:23–30. https://doi.org/10.3109/14756366.2014.1003215

247. Baba K, Nakata K, Taniguchi M, et al (1990) Chalcones from Angelica keiskei. Phytochemistry 29:3907–3910. https://doi.org/10.1016/0031-9422(90)85357-L

248. Jamil S, Sirat HM, Jantan I, et al (2008) A new prenylated dihydrochalcone from the leaves of Artocarpus lowii. J Nat Med 62:321–324. https://doi.org/10.1007/s11418-008-0226-3

249. López JA, Barillas W, Gomez-Laurito J, et al (1998) Galiposin: A new β-hydroxychalcone from Galipea granulosa. Planta Med 64:76–77. https://doi.org/10.1055/s-2006-957373

250. Mayer R (1993) A β-hydroxychalcone from Leptospermum scoparium. Planta Med 59:269–271. https://doi.org/10.1055/s-2006-959667

251. Masesane IB, Yeboah SO, Liebscher J, et al (2000) A bichalcone from the twigs of Rhus pyroides. Phytochemistry 53:1005–1008. https://doi.org/10.1016/S0031-9422(99)00553-1

252. Williams PG, Luesch H, Yoshida WY, et al (2003) Continuing studies on the cyanobacterium Lyngbya sp.: Isolation and structure determination of 15-norlyngbyapeptin A and lyngbyabellin D. J Nat Prod 66:595–598. https://doi.org/10.1021/np030011g

253. Tih AE, Tih RG, Sondengam BL, et al (1999) A novel hexaflavonoid from Lophira alata. Tetrahedron Lett 40:4721–4724. https://doi.org/10.1016/S0040-4039(99)00832-1

254. Smet M, Liao LX, Dehaen W, McGrath D V. (2000) Photolabile dendrimers using o-nitrobenzyl ether linkages. Org Lett 2:511–512. https://doi.org/10.1021/ol991373b

255. Tasdemir D, Sticher O, Çalis I, Linden A (1997) Further labdane diterpenoids isolated from Leonurus persicus. J Nat Prod 60:874–879. https://doi.org/10.1021/np9700782

256. (1994) AND CHALCONES FROM MUNDULEA SERZCEA THAT INHIBIT PHORBOL ESTER-INDUCED ORNITHINE DECARBOXYLASE. Phytochemistry 36:1523–1526. https://doi.org/https://doi.org/10.1016/S0031-9422(00)89755-1

257. Stevens JF, Page JE (2004) Xanthohumol and related prenylflavonoids from hops and beer: To your good health! Phytochemistry 65:1317–1330. https://doi.org/10.1016/j.phytochem.2004.04.025

258. Wang Q, Ding ZH, Liu JK, Zheng YT (2004) Xanthohumol, a novel anti-HIV-1 agent purified from Hops Humulus lupulus. Antiviral Res 64:189–194. https://doi.org/10.1016/j.antiviral.2004.08.005

259. Wu W, Ye H, Wan L, et al (2013) Millepachine, a novel chalcone, induces G2/M arrest by inhibiting CDK1 activity and causing apoptosis via ROS-mitochondrial apoptotic pathway in human hepatocarcinoma cells in vitro and in vivo. Carcinogenesis 34:1636–1643. https://doi.org/10.1093/carcin/bgt087

260. Oh KY, Lee JH, Curtis-Long MJ, et al (2010) Glycosidase inhibitory phenolic compounds from the seed of Psoralea corylifolia. Food Chem 121:940–945. https://doi.org/10.1016/j.foodchem.2010.01.022

261. Clark AM (2001) Maclura. Planta Med 67:87–89. https://doi.org/10.1055/s-2001-10621

262. Yoon G, Lee W, Kim SN, Cheon SH (2009) Inhibitory effect of chalcones and their derivatives from Glycyrrhiza inflata on protein tyrosine phosphatase 1B. Bioorganic Med Chem Lett 19:5155–5157. https://doi.org/10.1016/j.bmcl.2009.07.054

263. Rubal JJ, Guerra FM, Moreno-Dorado FJ, et al (2006) Sesquiterpenes from Thapsia nitida var. meridionalis and Thapsia nitida var. nitida. J Nat Prod 69:1566–1571. https://doi.org/10.1021/np0603065

264. Narender T, Shweta, Tanvir K, et al (2005) Prenylated chalcones isolated from Crotalaria genus inhibits in vitro growth of the human malaria parasite Plasmodium falciparum. Bioorganic Med Chem Lett

15:2453–2455. https://doi.org/10.1016/j.bmcl.2005.03.081

265. Mustafa K, Kjaergaard HG, Perry NB, Weavers RT (2003) Hydrogen-bonded rotamers of 2′,4′,6′-trihydroxy-3′-formyldihydrochalcone, an intermediate in the synthesis of a dihydrochalcone from Leptospermum recurvum. Tetrahedron 59:6113–6120. https://doi.org/10.1016/S0040-4020(03)00940-2

266. Takeshi N (1977) NII-Electronic Library Service. Chem Pharm Bull 57:364–370

267. Yang Z, Ma X, Tan W, et al (2015) Two new chalcones from Shuteria sinensis. Nat Prod Res 29:1909–1913. https://doi.org/10.1080/14786419.2015.1012718

268. Hoang DM, Ngoc TM, Dat NT, et al (2009) Protein tyrosine phosphatase 1B inhibitors isolated from Morus bombycis. Bioorganic Med Chem Lett 19:6759–6761. https://doi.org/10.1016/j.bmcl.2009.09.102

269. Chen RM, Hu LH, An TY, et al (2002) Natural PTP1B inhibitors from Broussonetia papyrifera. Bioorganic Med Chem Lett 12:3387–3390. https://doi.org/10.1016/S0960-894X(02)00757-6

270. Agüero MB, Gonzalez M, Lima B, et al (2010) Argentinean propolis from Zuccagnia punctata cav. (Caesalpinieae) exudates: Phytochemical characterization and antifungal activity. J Agric Food Chem 58:194–201. https://doi.org/10.1021/jf902991t

271. Adityachaudhhury N, Das AK CA and DP (1976) Aurentiacin, a new chalcone from Didymocarpus aurentica [1976]. Phytochemistry 15:229–30

272. Star AE, Mabry TJ, Smith DM (1978) Triangularin, a new chalcone from Pityrogramma triangularis. Phytochemistry 17:586–587. https://doi.org/10.1016/S0031-9422(00)89386-3

273. Tanaka H, Ichino K, Ito K (1984) Dihydrochalcones from lindera umbellata. Phytochemistry 23:1198–1199. https://doi.org/10.1016/S0031-9422(00)82646-1

274. Aida K, Tawata M, Shindo H, et al (1990) Isoliquiritigenin: A new aldose reductase inhibitor from glycyrrhizae radix. Planta Med 56:254–258. https://doi.org/10.1055/s-2006-960950

275. Christensen LP, Lam J, Thomasen T (1990) A chalcone and other constituents of bidens tripartitus. Phytochemistry 29:3155–3156. https://doi.org/10.1016/0031-9422(90)80177-I

276. Bhatt P, Dayal R (1992) Stipulin, a prenylated chalcone from Dalbergia stipulacea. Phytochemistry 31:719–721. https://doi.org/10.1016/0031-9422(92)90074-Z

277. Koyama K, Takahashi K (1992) New triterpenes from. J Nat Prod 5:953–955. https://doi.org/https://doi.org/10.1021/np50085a017

278. Venkata R, Rajendra P (1992) Two chalcones from Tephrosia spinosa. Phytochemistry 31:2121–2122. https://doi.org/https://doi.org/10.1016/0031-9422(92)80376-P

279. Gurney A, Sheppard N (1993) Anacine, a new benzodiazepine metabolite of. 1:1993. https://doi.org/https://doi.org/10.1021/np50100a009

280. Ahsan M, Armstrong JA, Waterman PG (1994) Dihydrochalcones from the aerial parts of Boronia inconspicua. Phytochemistry 36:799–801. https://doi.org/10.1016/S0031-9422(00)89821-0

281. Zhang H, Sun H, Takeda Y (1995) Four new taxanes from the roots of taxus yunnanensis. J Nat Prod 58:1153–1159. https://doi.org/10.1021/np50122a001

282. Srivastava R, Shaw AK, Kulshreshtha DK (1995) Triterpenoids and chalcone from Syzygium samarangense. Phytochemistry 38:687–689. https://doi.org/10.1016/0031-9422(94)00739-G

283. Shin J, Seo Y, Cho KW (1998) Five new polyacetylenes from a sponge of the genus Petrosia. J Nat Prod 61:1268–1273. https://doi.org/10.1021/np9802015

284. Yenesew A, Midiwo JO, Waterman PG (1998) ergamon. Phytochemistry 47:295–300. https://doi.org/https://doi.org/10.1016/S0031-9422(97)00424-X

285. Chantrapromma K, Rat-A-Pa Y, Karalai C, et al (2000) A chalcone and a dihydrochalcone from Uvaria dulcis. Phytochemistry 53:511–513. https://doi.org/10.1016/S0031-9422(99)00477-X

286. Abegaz BM, Ngadjui BT, Dongo E, et al (2002) Chalcones and other constituents of Dorstenia prorepens

and Dorstenia zenkeri. Phytochemistry 59:877–883. https://doi.org/10.1016/S0031-9422(01)00483-6

287. Tuchinda P, Reutrakul V, Claeson P, et al (2002) Anti-inflammatory cyclohexenyl chalcone derivatives in Boesenbergia pandurata. Phytochemistry 59:169–173. https://doi.org/10.1016/S0031-9422(01)00451-4

288. Ersam T, Achmad SA, Ghisalberti EL, et al (2002) A new isoprenylated chalcone, artoindonesianin J, from the root and tree bark of Artocarpus bracteata. J Chem Res - Part S 2002:186–187. https://doi.org/10.3184/030823402103171618

289. De Andrade Cunha GM, Fontenele JB, Nobre Júnior H V., et al (2003) Cytotoxic activity of chalcones isolated from lonchocarpus sericeus (Pocr.) Kunth. Phyther Res 17:155–159. https://doi.org/10.1002/ptr.1096

290. Akihisa T, Tokuda H, Ukiya M, et al (2003) Chalcones, coumarins, and flavanones from the exudate of Angelica keiskei and their chemopreventive effects. Cancer Lett 201:133–137. https://doi.org/10.1016/S0304-3835(03)00466-X

291. Yang WM, Liu JK, Qin XD, et al (2004) Antioxidant activities of three dihydrochalcone glucosides from leaves of Lithocarpus pachyphyllus. Zeitschrift fur Naturforsch - Sect C J Biosci 59:481–484. https://doi.org/10.1515/znc-2004-7-805

292. Daikonya A, Katsuki S, Kitanaka S (2004) Antiallergic agents from natural sources 9. Inhibition of nitric oxide production by novel chalcone derivatives from Mallotus philippinensis (Euphorbiaceae). Chem Pharm Bull 52:1326–1329. https://doi.org/10.1248/cpb.52.1326

293. Cuca Suárez LE, Bañol Vargas ÓE (2005) nuevas chalconas de Beilschmiedia tovarensis. Rev Colomb Química 34:35–41

294. Usman H, Hakim EH, Harlim T, et al (2006) Cytotoxic chalcones and flavanones from the tree bark of Cryptocarya costata. Zeitschrift fur Naturforsch - Sect C J Biosci 61:184–188. https://doi.org/10.1515/znc-2006-3-405

295. Zhang C, Ondeyka JG, Zink DL, et al (2009) Isolation, structure and antibacterial activity of pleosporone from a pleosporalean ascomycete discovered by using antisense strategy. Bioorganic Med Chem 17:2162–2166. https://doi.org/10.1016/j.bmc.2008.04.018

296. Padhye S, Ahmad A, Oswal N, Sarkar FH (2009) Emerging role of Garcinol, the antioxidant chalcone from Garcinia indica Choisy and its synthetic analogs. J Hematol Oncol 2:1–13. https://doi.org/10.1186/1756-8722-2-38

297. Borges-Argáez R, Vela-Catzín T, Yam-Puc A, et al (2009) Antiprotozoal and cytotoxic studies on some isocordoin derivatives. Planta Med 75:1336–1338. https://doi.org/10.1055/s-0029-1185670

298. Suresh A, Sheela X, Kanmani R, Mani C, Easwaran L SA and RV (2010) Isolation and identification of a chalcone from Baccopa monnieri. Asian J Chem 22:965–970

299. Yang Y, Zhang T, Xiao L, et al (2010) Two new chalcones from leaves of Morus alba L. Fitoterapia 81:614–616. https://doi.org/10.1016/j.fitote.2010.03.005

300. Dao TT, Nguyen PH, Lee HS, et al (2011) Chalcones as novel influenza A (H1N1) neuraminidase inhibitors from Glycyrrhiza inflata. Bioorganic Med Chem Lett 21:294–298. https://doi.org/10.1016/j.bmcl.2010.11.016

301. Zhao ZX, Jin J, Lin CZ, et al (2011) Two new chalcone glycosides from the stems of Entada phaseoloides. Fitoterapia 82:1102–1105. https://doi.org/10.1016/j.fitote.2011.07.005

302. Su XH, Li CY, Zhong YJ, et al (2012) A new prenylated chalcone from the seeds of Millettia pachycarpa. Chin J Nat Med 10:222–225. https://doi.org/10.3724/SP.J.1009.2012.00222

303. Zhang XJ, Li LY, Wang SS, et al (2013) Oxyfadichalcones A-C: Three chalcone dimers fused through a cyclobutane ring from Tibetan medicine Oxytropis falcata Bunge. Tetrahedron 69:11074–11079. https://doi.org/10.1016/j.tet.2013.11.018

304. Wu YC, Sureshbabu M, Fang YC, et al (2013) Potent inhibition of human neutrophil activations by bractelactone, A novel chalcone from Fissistigma bracteolatum. Toxicol Appl Pharmacol 266:399–407.

https://doi.org/10.1016/j.taap.2012.11.021

305. Wang PC, Ran XH, Luo HR, et al (2013) Nardokanshone A, a new type of sesquieterpenoid-chalcone hybrid from Nardostachys chinensis. Tetrahedron Lett 54:4365–4368. https://doi.org/10.1016/j.tetlet.2013.05.127

306. Ramli F, Rahmani M, Kassim NK, et al (2013) New diprenylated dihyrochalcones from leaves of Artocarpus elasticus. Phytochem Lett 6:582–585. https://doi.org/10.1016/j.phytol.2013.07.009

307. Shang YF, Oidovsambuu S, Jeon JS, et al (2013) Chalcones from the flowers of coreopsis lanceolata and their in vitro antioxidative activity. Planta Med 79:295–300. https://doi.org/10.1055/s-0032-1328188

308. Yu L, Zhang F, Hu Z, et al (2014) Novel prenylated bichalcone and chalcone from Humulus lupulus and their quinone reductase induction activities. Fitoterapia 93:115–120. https://doi.org/10.1016/j.fitote.2013.12.019

309. De Castro CCB, Costa PS, Laktin GT, et al (2015) Cardamonin, a schistosomicidal chalcone from Piper aduncum L. (Piperaceae) that inhibits Schistosoma mansoni ATP diphosphohydrolase. Phytomedicine 22:921–928. https://doi.org/10.1016/j.phymed.2015.06.009

310. Escobar-Ramos A, Lobato-García CE, Zamilpa A, et al (2017) Homoisoflavonoids and chalcones isolated from haematoxylum campechianum L., with spasmolytic activity. Molecules 22:1–10. https://doi.org/10.3390/molecules22091405

311. Hailemariam A, Feyera M, Deyou T, Abdissa N (2018) Antimicrobial Chalcones from the Seeds of Persicaria lapathifolia. Biochem Pharmacol Open Access 07:1–4. https://doi.org/10.4172/2167-0501.1000237

Partie II: Travaux Réalisés

(Applications sur la Maladie de Parkinson)

Contexte:

Comme nous l'avons signalé en introduction générale, la partie II intitulée travaux réalisés est composée de cinq chapitres où on expose nos travaux notamment sous forme de 5 publications. Les 3 premières concernent la maladie de Parkinson tandis que la 4ème et 5ème publication traite les résultats relatifs à la maladie d'Alzheimer.

Les deux premières publications qui constituent le chapitre I et II s'intéressent à la monoamine oxydase (MAO). C'est une enzyme dépendante de la flavine adénine dinucléotide (FAD) impliquée dans l'oxydation des neurotransmetteurs monoamines tels que les catécholamines (l'épinéphrine, la norépinéphrine et la dopamine) et la 5 hydroxytryptamine (la sérotonine) en leurs aldéhydes correspondants avec la formation de peroxyde d'hydrogène (H_2O_2). Chez les mammifères, cette enzyme existe sous deux isoformes à savoir la MAO-A et la MAO-B. Les isoenzymes humaines MAO-A (527 aa) et MAO-B (520 aa) présentent un degré élevé d'homologie de séquence d'acides aminés (elles partagent 70% d'identité d'acides aminés). La MAO-A est une cible médicamenteuse potentielle pour le traitement des troubles du comportement comme la dépression et l'anxiété, tandis que la MAO-B, est une cible médicamenteuse potentielle pour la maladie de Parkinson, qui résulte d'un métabolisme excessif de la dopamine. L'inhibition sélective de la MAO-B améliore les symptômes de la maladie de Parkinson en réduisant les périodes d'arrêt. Les inhibiteurs de la MAO-B sont également indiqués dans la maladie d'Alzheimer pour réduire le métabolisme de la dopamine et l'excès de stress oxydatif causé par le peroxyde d'hydrogène généré dans le métabolisme des neurotransmetteurs monoamines.

Au cours de ces dernières années, plusieurs groupes de recherche ont permis de découvrir de nouvelles petites molécules appartenant à diverses classes chimiques et présentant un potentiel inhibiteur de la MAO-B. Ces classes comprennent les acétylphénols, les acétophénones, les nitriles, les propargylamines, les indanes, les indanones, les oxadiazoles, les pyridines, les chromones, les coumarines, les quinolones, les isoxazoles, les pyrazolines et les benzoxazoles, les benzothiazoles dont les structures générales sont présentées dans la figure a1.

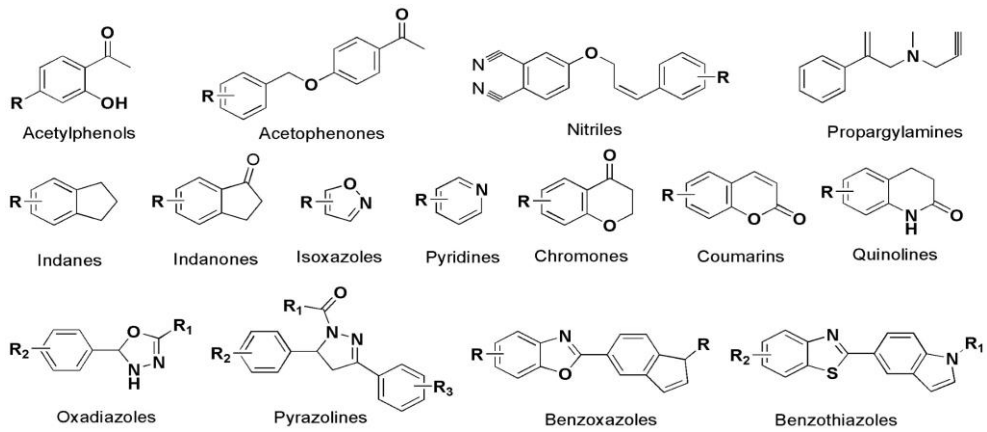
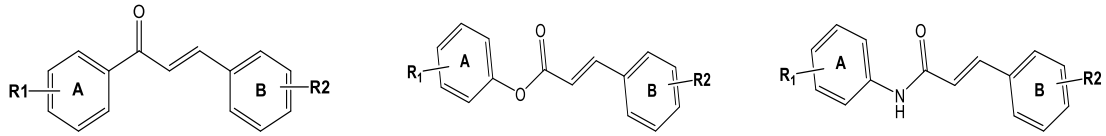


Figure a1. *Structure générale des squelettes chimiques possédant des propriétés inhibitrices de la MAO-B*

Une autre classe qui occupe actuellement une place importante dans le domaine de la chimie organique en tant qu'intermédiaires de synthèse et en thérapeutique en tant que produits biologiquement actifs, nous parlons de la chalcone et ses dérivés figure a2.



(1) **(2)** **(3)**

Figure a2. *Structures chimiques des chalcones (1), des dérivés esters α, β-insaturés (2), et des dérivés amides α, β-insaturés (3)*

Dans cette étude, nous avons sélectionnés une série de 25 molécules dérivées de chalcone et réalisé une étude *in silico* par les méthodes QSAR (2D/3D-QSAR), docking moléculaire, simulations par dynamique moléculaire et propriétés ADMET. Le premier objectif était de trouver une relation entre cette série de molécules et les descripteurs moléculaires calculés, et par conséquent la prédiction de nouveaux inhibiteurs de la MAO-B avec une efficacité inhibitrice élevée (Etude QSAR nD). Le second objectif était d'étudier le site de liaison et les interactions de ces composés (études de docking) dans le site actif de la MAO-B et par conséquent la prédiction de la stabilité (étude de dynamique moléculaire). Le troisième objectif était d'évaluer les propriétés pharmacocinétiques de ces composés. Pour atteindre ces objectifs ce travail a été divisé en deux grandes parties et a donné lieu à deux publications (Chapitre I et II).

Chapitre I : Étude computationnelle de dérivés carbonylés insaturés en tant qu'inhibiteurs de la MAO-B à l'aide de QSAR, ADME/Tox, docking moléculaire et simulations par dynamique moléculaire

Résumé :

La première publication consiste à étudier une série de 25 molécules à base de chalcone. Les composés sélectionnés sont connus comme agents inhibiteurs de la MAO-B et présentent une large gamme d'activité pIC_{50} (d'ordre 3 unités logarithmiques). Tout d'abord, les molécules ont été dessinées à l'aide du logiciel Marvin Sketch. Ensuite, les molécules ont été optimisées par le logiciel SYBYL-X 2 en utilisant la procédure de gradient conjugué basée sur le champ de force tripos avec une valeur de critère de convergence de 0,01 kcal/mol Å et un nombre maximum d'itérations de 2000.

Pour comprendre la relation entre les structures chimiques et leurs activités biologiques associées, une analyse 2D-QSAR a été réalisée, sur l'ensemble des molécules. Pour chaque composé de l'ensemble d'apprentissage (80% de l'ensemble des molécules), un nombre total de 354 descripteurs moléculaires a été calculé en utilisant le logiciel MOE (molecular operating environment). Après élimination des descripteurs corrélés entre eux et des descripteurs n'ayant aucune influence sur l'activité biologique étudiée, la méthode de régression des moindres carrés partiales (ou PLS) a été appliquée pour générer le modèle 2D-QSAR.

Le modèle développé a été évalué par le coefficient de détermination (r^2), le coefficient de détermination ajusté (r^2_a), l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne et la valeur de Fisher (F). La validation interne du modèle élaboré a été établie par la procédure leave-one-out (Q^2_{LOO}), le paramètre R^2_m (LOO) et le test de randomisation en Y. La validation externe a été réalisée en utilisant le reste de l'ensemble des molécules (20% de l'ensemble total). Les paramètres r^2 test, R^2_m (test) et Q^2 (F2) ainsi que les critères de Golbraikh et Tropsha sont utilisés pour vérifier la validation externe du modèle QSAR élaboré. Enfin, la validité du modèle développé a été gouvernée par le domaine d'applicabilité.

Les modes d'interaction ainsi que les conformations des ligands par rapport à la structure cible étudiée (MAO-B) ont été déterminés par la méthode de docking moléculaire en utilisant le programme autodock vina. Dans un premier temps, la structure 3D du récepteur MAO-B a été téléchargée de la banque de données des protéines (PDB ID : 2BK3). Le ligand a été extrait de la protéine et les molécules d'eau ont été éliminées. Les charges de Gast-Huck et l'hydrogène polaire ont été ajoutés. Ensuite, une grille centrée sur le site actif du récepteur MAO-B a été créée avec une dimension de 40 × 40 × 40 Å dans les directions x, y et z. Après avoir validé le site actif du récepteur par la méthode de re-docking, tous les composés ont été dockés dans la même grille sélectionnée. La stabilité des ligands dans le site actif du récepteur a été conférée par l'étude de la simulation par dynamique moléculaire. Tous les calculs ont été effectués à l'aide du programme GROMACS 5.1.4. Le champ de force tout-atome CHARMM36 et le champ de force général (CGenFF) ont été utilisés pour générer les paramètres de protéine et de ligand, respectivement. Les fichiers de topologie de protéine ont été créés à l'aide du module pdb2gmx. Ensuite, les complexes protéine-ligands ont été construits dans une boîte périodique cubique et solvatés via un modèle d'eau à trois points (TIP3). La charge des systèmes a été neutralisée en ajoutant des ions sodium (Na^+) ou chlorure (Cl^-). Tous les systèmes ont ensuite été minimisés en énergie en utilisant la méthode de descente avec un maximum de 50 000 pas et une force maximale de 10.0 KJ/mol. Les simulations ont été effectuées dans un ensemble à

nombre constant de molécules, pression constante et température constante (NPT) en utilisant le thermostat de Berendsen avec une température et une pression couplées à 310 K et 1 bar, respectivement. Enfin, le système protéine-ligand a été évalué par la méthode de la dynamique moléculaire

Les propriétés pharmacocinétiques des composés étudiés ont été analysées par la prédiction ADMET (adsorption, distribution, métabolisme, excrétion et toxicité) en utilisant les serveurs web pkCSM et SWISSadmet. Tous les composés ont été évalués en fonction de leurs adsorptions, distributions, métabolismes, excrétions et toxicités dans le corps humain.

Résultats et discussion

Les résultats des études 2D-QSAR ont révélé l'importance des descripteurs BCUT_SMR_2, logP (o/w), SlogP_VSA4 et vsurf_IW3 dans la description des valeurs d'activité biologique étudiée. Tous les descripteurs sélectionnés montrent une faible corrélation entre eux et une forte corrélation avec l'activité. L'équation linéaire reliant les descripteurs sélectionnés à l'activité inhibitrice de la MAO-B est présentée ci-dessous :

$$\boldsymbol{pIC_{50}} = -3.754 + 8{,}469 * \mathbf{BCUT_SMR_2} + 0.636 * \mathbf{logP}\left(\frac{\mathbf{o}}{\mathbf{w}}\right) + 0.057 * \mathbf{SlogP_VSA4} - 0.442 * \mathbf{vsurf_IW3}$$

Le meilleur modèle généré explique 88 % (r^2 = 0.88) de la variance totale de l'ensemble d'apprentissage avec des faibles valeurs d'erreur (RMSE = 0.28 et MAE = 0.2). La valeur élevée du test de Fisher (F = 27.76) indique que le modèle généré est statistiquement significatif. De même, la signification de chaque descripteur a été vérifiée en calculant le facteur de variation de l'inflation (VIF) et la valeur de p-value. Tous les descripteurs du modèle généré ont des valeurs acceptables de VIF et de p-value (VIF < 5 et p-value < 0.05), ce qui montre qu'il existe une relation significative entre la réponse modélisée (pIC_{50}) et les descripteurs sélectionnés.

La performance de la validation interne a été déterminée par la validation croisée en utilisant la procédure « leave-one out ». Les résultats montrent une bonne stabilité interne du modèle élaboré (Q^2_{LOO} = 0.81, $RMSE_{LOO}$ = 0.31 et MAE_{LOO} = 0.27). La qualité de Q^2_{LOO} a également été vérifiée en calculant le paramètre R^2_m (R^2_m = 0.57 > 0.5). Pour assurer la robustesse du modèle développé, un test de randomisation en Y a été appliqué. Plusieurs randomisations aléatoires de pIC_{50} ont été effectuées. Les valeurs faibles de R^2 et Q^2 (R^2= 0.15 et Q^2 = -0.52) et la valeur élevée de cRp^2 (cRp^2 = 0.81 > 0.5), indiquent que les bons résultats de notre modèle original ne sont pas dus à une corrélation fortuite. De plus, les résultats de la validation externe (r^2_{test} = 0.71) indiquent que le modèle développé est capable de prédire l'activité de nouveaux composés non testés. En outre, la qualité de la validation externe a été vérifiée en calculant le paramètre Q^2 (F2), qui a donné une valeur supérieure à 0.5 (Q^2 (F2) = 0.52).

Afin de prédire des nouveaux agents inhibiteurs de MAO-B, il est nécessaire d'expliquer la signification de chaque descripteur sélectionné. Le descripteur le plus important dans le modèle 2D-QSAR établi est SlogP_VSA4. Ce descripteur représente la somme des surfaces de van der Waals accessibles telles que le logP pour l'atome i est dans la gamme (0.1 à 0.15). La

contribution positive (+ 0.057) et la forte corrélation (r = 0.72) avec l'activité inhibitrice, indiquent que les composés ayant une surface de van der Waals accessible élevée pourraient augmenter l'activité inhibitrice de l'enzyme MAO-B. Une analyse approfondie des structures chimiques et de leur activité expérimentale indique que le descripteur SlogP_VSA4 est lié au nombre d'atome de fluor (F) dans les positions R1 et R2. À partir de l'ensemble de données expérimentales, les composés avec R1 = F (16a, 16b, 16c, 17a et 18a) ont la même valeur de SlogP_VSA4 (SlogP_VSA4 = 6.3716) et les composés avec R2 = CF_3 ont la même valeur de SlogP_VSA4 (SlogP_VSA4 = 33.4189). Pour s'assurer que le descripteur SlogP_VSA4 est lié au nombre d'atome de fluor, nous avons comparé les composés 10b et 11b, qui ont le même substituant en position R2 (R2 = CF_3) mais diffèrent en position R1 (10b : R1 = -OCH_3, 11b : R1 = OH). La valeur du descripteur SlogP_VSA4 est la même dans les deux composés (SlogP_VSA4 = 33.4189).

L'affinité de ces composés dans le site de liaison de MAO-B (PDB ID : 2BK3) a été déterminée par la méthode de docking moléculaire. Les résultats montrent que ces composés ont une bonne affinité avec le site actif de la MAO-B, ce qui pourrait conduire à une inhibition de son activité enzymatique. De plus, le meilleur composé en terme d'activité anti-MAO-B de la série étudiée (composé 10b) a été évalué par la méthode de dynamique moléculaire en utilisant le logiciel GROMACS. Les résultats montrent l'importance de la liaison hydrogène dans la stabilité du complexe (10b-2BK3) durant la période de simulations. Enfin, les propriétés pharmaceutiques en termes d'absorption, distribution, métabolisme, excrétion et toxicité ont été déterminées à l'aide d'analyse ADMET. Les résultats positifs de l'analyse ADMET indiquent que les dérivés cétoniques étudiés ont un potentiel prometteur en tant qu'inhibiteurs de la MAO-B.

Turkish Journal of Chemistry
http://journals.tubitak.gov.tr/chem/
Research Article

Turk J Chem
(2022) 46: 687-703
© TÜBİTAK
10.55730/1300-0527.3360

Computational investigation of unsaturated ketone derivatives as MAO-B inhibitors by using QSAR, ADME/Tox, molecular docking, and molecular dynamics simulations

Abdellah EL AISSOUQ[1,*], Mohammed BOUACHRINE[3,4], Abdelkrim OUAMMOU[2], Fouad KHALIL[1]

[1] LPME Laboratory, Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[2] LIMOME Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[3] MCNS Laboratory, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco
[4] EST Khenifra, Sultan Moulay Sliman University, Morocco

Received: 31.07.2021 **Accepted/Published Online:** 18.12.2021 **Final Version:** 16.06.2022

Abstract: Unsaturated ketone derivatives are known as monoamine oxidase B (MAO-B) inhibitors, a potential drug target for Parkinson's disease. Here, molecular modeling studies, including 2D-QSAR, ADMET prediction, molecular docking, and MD simulation, were performed on a new series of MAO-B inhibitors. The objective is to identify new MAO-B inhibitors with high inhibitory efficacy. The developed 2D-QSAR model was based on the descriptors of MOE software. The most appropriate model, using the partial least squares regression (PLS regression) method, yielded 0.88 for the determination coefficient (r^2), 0.28 for the root-mean-square error (RMSE), and 0.2 for the mean absolute error (MAE). The predictive capacity of the generated model was evaluated by internal and external validations, which gave the Q^2 and R^2_{test} values of 0.81 and 0.71, respectively. The ability of a compound to be orally active was determined using the drug-likeness and ADMET prediction. The results indicate that most of the compounds have moderate pharmacokinetic characteristics without any side effects. Furthermore, the affinity of the ligands (unsaturated ketone derivatives) to the MAO-B receptor was determined using molecular docking. The top conformers were then subjected to MD simulation. This research may pave the way for the development of novel unsaturated ketone derivatives capable of inhibiting the MAO-B enzyme.

Key words: MAO-B inhibitors, Parkinson's disease, ligand-based drug design, structure-based drug design

2. 1. Introduction

Monoamine oxidase (MAO) is a flavin adenine dinucleotide (FAD)-dependent enzyme involved in the oxidation of monoamine neurotransmitters such as catecholamines (i.e. epinephrine, norepinephrine, and dopamine) and 5-hydroxytryptamine (i.e. serotonin, to their corresponding aldehydes with the formation of hydrogen peroxide (H_2O_2) [1] (figure 1). In mammals, this enzyme exists in two isoforms, MAO-A and MAO-B, encoded by two genes located on the X chromosome [2]. Human isoenzymes MAO-A (527 AA) and MAO-B (520 AA) display a high degree of amino acid sequence homology (share 70% amino acid identity) [3]. Each isoenzyme is the target of affective and neurogenerative disorders [4, 5]. MAO-A inhibitors are used to treat behavioral disorders, particularly depression [6, 7]. In contrast, MAO-B inhibitors such as selegiline and rasagiline are used in the treatment of Parkinson's disease (PD). MAO-B inhibitors are also involved in the treatment of Alzheimer's disease (AD) [8, 9] and other neurodegenerative disorders [10]. However, MAO-B drugs are irreversible and cause pharmacological side effects in the long-term treatment of PD [11]. Therefore, the development of reversible MAO-B inhibitors, such as the recent anti-Parkinson drug safinamide [12], is required.

* Correspondence: abdellah.elaissouq@usmba.ac.ma
This work is licensed under a Creative Commons Attribution 4.0 International License.

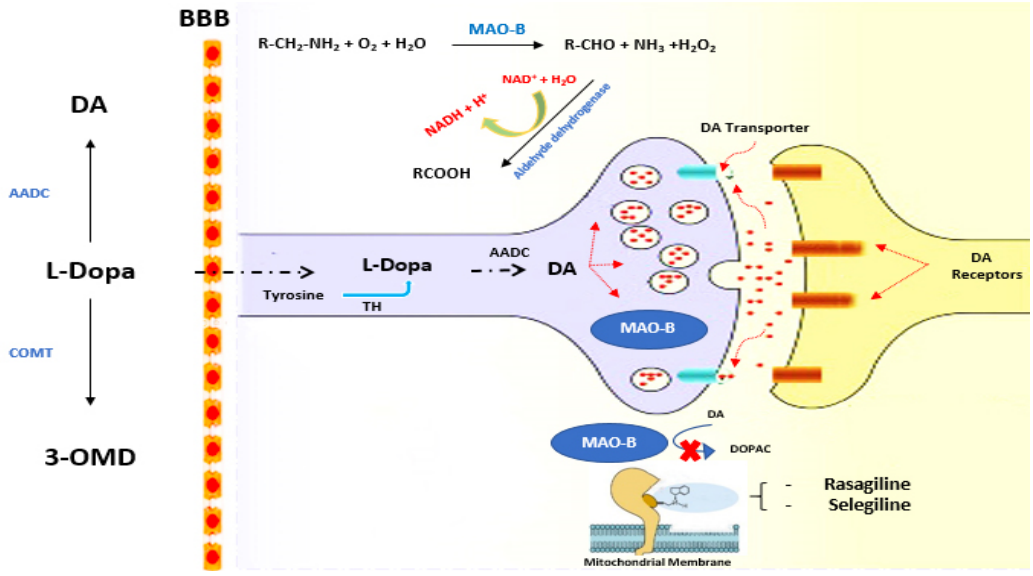
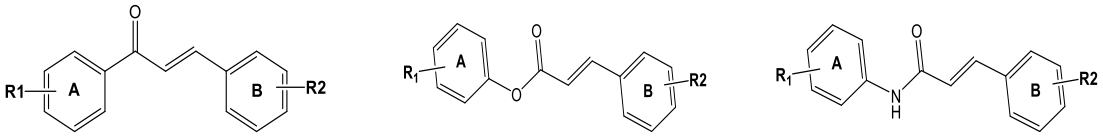


Figure 1. *Mechanism of interaction of monoamine oxidase-B (MAO-B) inhibitors. AD = dopamine, L-Dopa = levodopa, AADC = aromatic L-amino acid decarboxylase, 3-OMD= 3-O-methyldopa, COMT= catechol-O-methyltransferase, TH= tyrosine hydroxylase, and DOPAC = 3, 4-dihydroxyphenylacetic acid*

To search for new MAO-B inhibitors, chalcone scaffolds are selected. Chalcones are a class of polyphenolic derivatives belonging to the flavonoid family, their structure consisting of two aromatic cycles linked by a three-carbon α, β-unsaturated carbonyl system (Figure 2). They exhibit a wide range of pharmacological activities, including anticancer [13], anti-inflammatory [14], antidiabetic [15], antioxidants [16, 17], antimalarial [18], anti-HIV [19] and anti-MAO-B [20].



(1) **(2)** **(3)**

Figure 2. *Chemical structures of chalcones (1), α, β-unsaturated ester derivatives (2), and α, β-unsaturated amide derivatives (3)*

Computer-aided drug design (CADD) approaches have recently become essential in drug discovery. These techniques proved to be effective in various stages of the drug development process, reducing both the cost and time required to develop a drug compared to conventional methods [21]. In general, CADD approaches are classified into two types: structure-based drug design (SBDD) and ligand-based drug design (LBDD). When the 3D structure of the protein is available, SBDD methods such as homology modeling, molecular docking, and molecular dynamics simulation are used [22]. On the other hand, LBDD methods such as quantitative structure-activity relationship (QSAR),

similarity search, and pharmacophore modeling are applied when the 3D structure of the protein is not available. With the help of these methods, several approved drugs have been developed [23-28]. For example, the discovery of amprenavir as a potential inhibitor of the human immunodeficiency virus (HIV) protease using protein modeling and MD simulations [29], imatinib as an inhibitor of the tyrosine kinase using SBDD [30], amprenavir as a potential inhibitor of the human immunodeficiency virus (HIV) protease using protein modeling and MD simulations[31, 32], thymidylate synthase inhibitor, raltitrexed against HIV using the SBDD approach [33], Norfloxacin as an inhibitor of topoisomerase II, IV using QSAR modeling [32] and darolutamide as an inhibitor of androgen receptor using docking and MD simulations[34].

In this work, a new series of chalcones and their derivatives were synthesized by Choi and co-workers [35]. The synthesized compounds are known as monoamine oxidase inhibitors (MAOI). They have a wide range of activity ($0.016 < IC_{50}$ (µM) < 8.39). To understand the relationship between the chemical structures and their related activities, 2D-QSAR was performed, based on the descriptors of molecular operating environment (MOE) software. The pharmacokinetic properties of compounds were analyzed using the drug-likeness and the ADMET (adsorption, distribution, metabolism, excretion, and toxicity) prediction. Docking study was also displayed to find out the binding modes of the unsaturated ketone derivatives in the active site of monoamine oxidase- B (MAO-B) receptor. Finally, the dynamic behavior and stability of ligand-receptor complexes were evaluated using molecular dynamics (MD) simulation. In this study, the combination of ligand and structure-based drug design may be helpful to develop and design new unsaturated ketone candidates as MAO-B inhibitors.

I. Materials and Methods

1. Dataset preparation and structure optimization

A dataset of unsaturated ketone derivatives synthetized by Choi and co-workers was selected [35]. The IC_{50} values were converted into corresponding pIC_{50} values (pIC_{50} = log ($1/IC_{50}$.) and then used as a dependent variable in this study. The selected compounds are known as monoamine oxidase inhibitors and display a wide range of activity (3 log units). The dataset was randomly split into a training set (80% of the dataset) for 2D-QSAR model generation and a test set (20% of the dataset) for testing the predictive ability of the generated model. The core substructure of these compounds is shown in **Figure 3**.

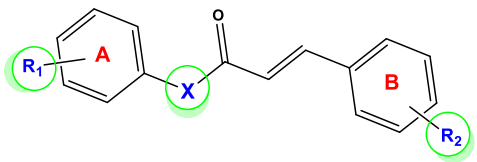
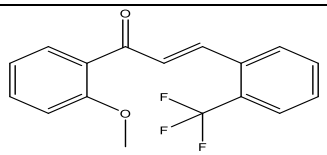
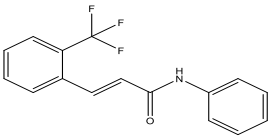
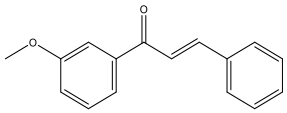
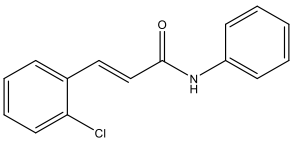
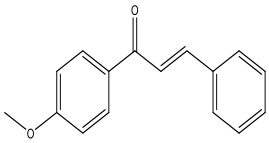
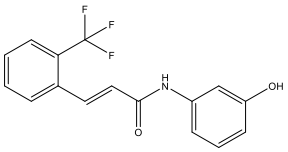
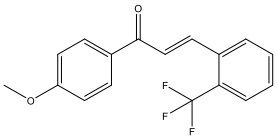
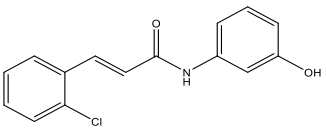
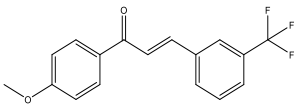
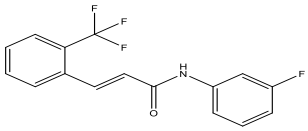
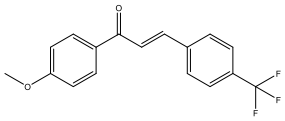
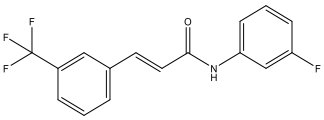
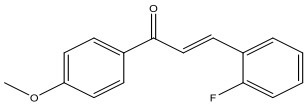
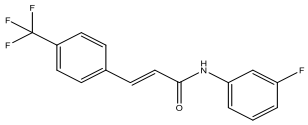
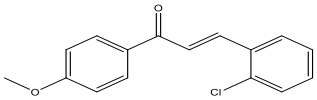
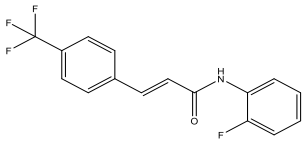
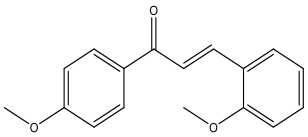
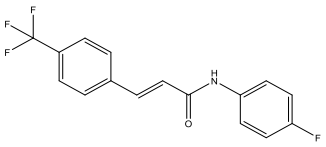
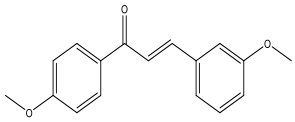
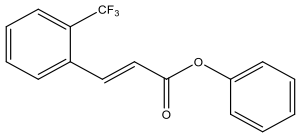
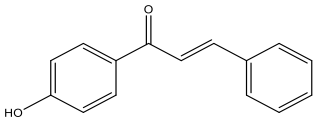
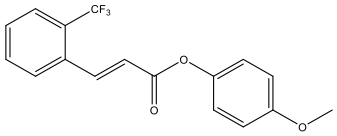
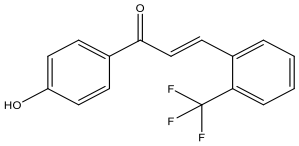
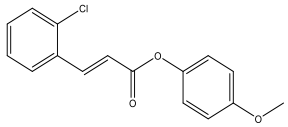
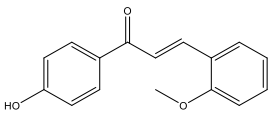


Figure 3. *Core sub-structure of unsaturated ketone derivatives*

Before modeling, the 3D chemical structures were designed using Marvin Sketch software. Energy minimization and calculations were performed using SYBYL-X 2.1 software [36]. All compounds were energy minimized using the conjugate gradient procedure based on the tripos force field with a convergence criterion value set to 0.01 kcal/mol Å [37]. The maximum number of iterations was set to 2000. Partial atomic charges were added, using Gasteiger-Hücke charges [38]. The chemical structures and their corresponding activity are presented in **Table 1**.

Table 1. *Chemical structures of unsaturated ketone derivatives and their corresponding experimental activity*

No	Structure	pIC_{50}	No	Structure	pIC_{50}
8b		6.5969	12a^{T}		5.0762
9a		6.1007	12b		5.5498
10a		6.3270	15a		5.2233
10b		7.7959	15b		5.2832
10c		6.8386	16a		5.4724
10d^{T}		6.7011	16b		5.3107

10e		7.2518	16c		6.5482
10f		7.1612	17a		6.2366
10g		6.3778	18a		6.6517
10h[T]		5.8551	19a		5.2306
11a[T]		6.7399	20a		5.5867
11b[T]		7.2441	20b		6.0283
11g		6.2048			

[T] indicates test set compounds

2. 2D-QSAR analysis

2.1.Molecular descriptors calculation

The 2D-QSAR model was constructed using MOE descriptors. For each compound, a total of 354 descriptors were calculated. Then, the descriptors with a zero or equal variance predictor and the descriptors with a constant value for all observations were deleted. The number of descriptors was reduced by removing descriptors with a low

correlation with the pIC_{50} value. The multicollinearity of the selected descriptors was verified by calculating the variation inflation factors (VIF), which can be calculated as follows:

$$VIF = \frac{1}{1-r^2} \quad \textbf{Eq. 1}$$

Where r^2 is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. If the VIF value is greater than 5 (VIF > 5), multicollinearity is very high [39].

3. Model generation

After selecting the most appropriate descriptors, PLS method was used to build the linear 2D-QSAR model. The MOE descriptors were used as independent variables (X variables), while MAO-B inhibitory activity (pIC_{50}) was employed as the dependent variable (Y variables). Both variables are related by the following equation below:

$$Y = a_0 + a_1X_1 + a_2X_2 + \cdots + a_nX_n \quad \textbf{Eq.2}$$

The developed QSAR model is evaluated by the squared correlation coefficient (r^2), the adjusted squared correlation coefficient (r^2_a), the root-mean-square error (RMSE), the mean absolute error and the fisher value (F).

$$r^2 = 1 - \left[\frac{\sum_i (Y_i \text{ obs} - Y_i \text{pred})^2}{\sum_i (Y_i \text{ obs} - \overline{Y}_i \, mean)^2}\right] = 1 - \frac{SSE}{SST} \quad \textbf{Eq.3}$$

$$r_a^2 = \frac{(n-1) \times r^2 - p}{n-1-p} \quad \textbf{Eq.4}$$

$$RMSE = \sqrt{\frac{\sum_{I=1}^{n} (\text{Yi obs} - \text{Yi pred})}{n}} \quad \textbf{Eq.5}$$

$$MAE = \frac{1}{n}\sum_{I=1}^{n} |\text{Yi obs} - \text{Yi pred}| \quad \textbf{Eq.6}$$

$$F = \frac{\frac{\sum_i (Y_i \text{ pred} - \overline{Y}_i \, mean)^2}{p}}{\frac{\sum_i (Y_i \text{ obs} - \text{Yi pred})^2}{N-P-1}} \quad \textbf{Eq.7}$$

In the above equations, Yi obs and Yi pred are the observed and predicted activities (pIC_{50} obs and pIC_{50} pred) for i^{th} compound in the training set, $\overline{Y}_i \, mean$ is the average activity ($\overline{pIC_5}o$ mean) of the compounds in the training set, SSE is the residual of squares, SST is the total sum of squares, n is the total number of compounds in the training set, and p is the number of descriptors in the generated model. For a good model, the value of r^2

should be closed to 1, the values of RMSE and MAE should be closed to 0, and the value of F should be high. The generated model is used to understand how the activity changes when any one of the descriptors is varied.

2.2.QSAR Model validation

2.2.1. Internal validation

Leave-One-Out cross validation

In order to judge the quality and goodness of the generated QSAR model, the Leave-One-Out (LOO) cross validation process is performed [40, 41]. In this process, one compound is primarily eliminated from the training set. Then, the QSAR model is built based on the remaining compounds (n-1), and the activity of the deleted compound is predicted by the established QSAR model. This process is repeated until all the compounds have been removed once. The performance of the QSAR model is measured by the cross validated correlation coefficient (Q^2_{LOO}), which is calculated by the following equation below [42]:

$$Q^2_{LOO} = 1 - \left[\frac{\sum_i (Y_i \text{ obs} - Y_i \text{pred})^2}{\sum_i (Y_i \text{ obs} - \overline{Y}_i \, mean)^2}\right] \qquad \textbf{Eq.8}$$

In the above equation, Yi obs and Yi pred are the observed and predicted activities (pIC_{50} obs and pIC_{50} pred) for i^{th} compound in the training set, based on the LOO cross validation method, $\overline{Y}_i \, mean$ is the average activity ($\overline{pIC50} \, mean$) of the compounds in the training set. The value of Q^2_{LOO} reflects the quality of the model and it should be > 0.5 (Q^2_{LOO} > 0.5) [43].

R2m (LOO) parameter

A high value of Q^2_{LOO} does not indicates that the observed and predicted activities are close to each other. To solve this problem, and to better indicate the internal validation of the QSAR model, $R^2_{m\,(LOO)}$ and $R'^2_{m\,(LOO)}$ metrics are calculated [44].

$$R_m^2 = \boldsymbol{R^2} \times \left(\mathbf{1} - \sqrt{(\boldsymbol{R^2} - \mathbf{R_0^2})}\right) \qquad \textbf{Eq.9}$$

$$R'_m{}^2 = \boldsymbol{R^2} \times \left(\mathbf{1} - \sqrt{(\boldsymbol{R^2} - \mathbf{R_0'^2})}\right) \qquad \textbf{Eq.10}$$

Where, R^2 and R_0^2 are the squared correlation coefficient values between the observed and calculated (Leave-One-Out) activities with and without intercept, respectively. The parameter $R_0'^2$ has the same meaning as R_0 but with the axes reversed. The value of $R^2_{m\,(LOO)}$ and $R'^2_{m\,(LOO)}$ should be more than 0.5.

Y-randomization test

To ensure the robustness of the developed QSAR model, a Y-randomization test is performed[45]. The values of pIC_{50} are randomly permuted and new a QSAR model is generated using the original descriptors [46]. The new QSAR models are expected to have significant low R^2 and Q^2 values for several trials, which confirm that the generated QSAR model is robust and not due to a chance correlation. Another parameter, ${}^{c}R_p^2$ is also calculated by the following equation below:

$$ {}^{c}R_p^2 = R * \sqrt{R^2 - (Average\ R_r)^2} \qquad \textbf{Eq.9} $$

Where, R_r is the average 'R' of random models. The ${}^{c}R_p^2$ value should be more than 0.5 to pass this test.

2.2.2. External validation

Test set (r2 test)

The predictability of the elaborated QSAR model is evaluated by the external validation or test set. This process consists in keeping a set of compounds not included in the model generation and their activity values are predicted by the generated QSAR model. The performance of the external validation is evaluated by the squared correlation coefficient of prediction (R^2_{test}), which is calculated by the following equation [47].

$$ r^2_{test} = 1 - \left[\frac{\sum_i (Y_i \text{ obs (test)} - Y_i \text{pred (test)})^2}{\sum_i (Y_i \text{ obs (test)} - \overline{Y}_i \ mean\ (train))^2}\right] \qquad \textbf{Eq.10} $$

Where, $Y_{i\ obs\ (test)}$ and $Y_{i\ pred\ (test)}$ represent the observed and predicted activity ($pIC_{50\ obs}$ (test) and $pIC_{50\ pred}$ (test)) of the i^{th} compound in the test set, respectively and $\overline{Y}_i$ mean (train) ($\overline{pIC50}\ mean$) represents the mean activity of the compounds in the training set. The value of the R^2_{test} should be more than 0.5, for a power predictability.

Golbraikh and Tropsh's criteria

The external predictability of the generated QSAR model is also evaluated by several parameters. According to Golbraikh and Tropsha [48], a QSAR model is considered satisfactory if all the following conditions are satisfied:

1. $Q^2 > 0.5$
2. $R^2_{test} > 0.6$
3. $\frac{r^2 - r_0^2}{r^2} < 0.1$ and $0.85 < k < 1.15$ or $\frac{r^2 - r_0'^2}{r^2} < 0.1$ and $0.85 < k' < 1.15$
4. $\left|r^2 - r_0'^2\right| < 0.3$

Where k is the slop of the plot of the observed and predicted values of compounds via the origin and $k'$is the reversed axes intercept.

R^2_m (test) parameter

A high value of R^2_{test} does not indicate that the observed and predicted activities are close to each other. To solve this problem, and to better indicate the external validation of the QSAR model, $R^2_{m\ (test)}$ and R'^2_m (test) metrics, similar to $R^2_{m\ (LOO)}$ and $R'^2_{m\ (LOO)}$ are calculated [44]. For an acceptable prediction, the value of ΔR^2_m (test) should be lower than 0.2, provided that the value of $\overline{R^2_m}$ is more than 0.5 [44].

$$\overline{R^2_m} = \frac{(R^2_m + R'^2_m)}{2} \qquad \textbf{Eq.11}$$

$$\Delta R_m = |R^2_m - R'^2_m| \qquad \textbf{Eq.12}$$

Q^2 (F2) metric

The Q^2 (F2) metric can be calculated by the following equation below [49]:

$$Q^2_{(F2)} = 1 - \frac{\sum_i (Y_i \text{ obs (test)} - Y_i \text{pred (test)})^2}{\sum_i (Y_i \text{ obs } (train) - \overline{Y}_i\ test)^2} \qquad \textbf{Eq.13}$$

Where, $\overline{Y}_i$ test represents the mean observed data of the test set compounds. The value of Q^2 (F2) should be more than 0.5.

2.3. Applicability Domain

The applicability domain is defined as a theoretical region in chemical space, including both the model variables and modeled response [50, 51]. The applicability domain is based on the calculation of the leverage value (h_i) of each compound i, for which a QSAR model is used to predict its activity:

$$h_i = x_i^T (X^T X)^{-1} x_i \qquad (i = 1, \ldots, n), \qquad \textbf{Eq.14}$$

In **Eq.14**, x_i is the descriptor row-vector of a query compound, and X is the n*(k-1) matrix of k descriptor values for n data set compounds. A compound i was considered outside the applicability domain when the leverage value (h_i) of this compound is greater than the critical value (h*) (h* = 3(d+1)/n, where d is the number of variables and n is the number of compounds in the training set). Conversely, a compound i is considered inside the applicability domain when h_i is lower than the h*.

3. Drug likeness and ADMET analysis

Drug likeness and ADMET (absorption, distribution, metabolism, excretion, and toxicity) analysis were performed using pkCSM [52] and SWISSadmet [53] web servers. The drug likeness of all compounds was verified using the Lipinski's rule of five [54], veber [55], Egan, Muegge [55], and Ghose [56] rules. The pharmacokinetic properties such as blood-brain barrier (BBB) permeability, human intestinal absorption (HIA), water solubility (log mol/L), CYP450 substrate and its inhibitor (CYP2D6, CYP3A4, CYP1A2, CYP2C19 and CYP3A4), CaCo-2 permeability and toxicity were evaluated using ADMET analysis. Some other physicochemical properties such as molecular weight (MW), octanol-water partition (LogP), number of hydrogen bond acceptors (HBA), number of hydrogen bond donors (HBD), and topological polar surface area were also verified.

4. Docking study

Molecular docking is one of the most virtual screening methods, especially when the 3D structure of the receptor is available [57-59]. Here, the X-ray diffraction structure of monoamine oxidase B (MAO-B) was downloaded from the protein database bank (PDB ID: 2BK3) [60]. All preparations were carried out using the AutoDock tool [61]. The ligand was extracted from the protein and the water molecules were removed. Gast-Huck charges and polar hydrogen were added to the crystal structure. Then, a grid box centered on the catalytic site of the MAO-B receptor was created with a dimension of 40 x 40 x 40 Å in x, y, and z directions, respectively. Finally, the optimized ligands were converted to the pdbqt format and docked into the binding site of MAO-B using AutoDock vina [62]. The results were analyzed and visualized using PyMol [63] and Discovery Studio 2017 R2 [64] software.

5. Molecular Dynamics simulations

The molecular dynamics simulations were performed using GROMACS 5.1.4 package [65]. The topology files of the ligand and the protein were generated using the CHARMM General Force Field (CGenFF) server and the 'pdb2gmx' script, respectively [66]. The simulations were run using the CHARMM36 all-atom (March, 2019) force field [67] in a triclinic box with a distance of 1.0 nm and a TIP3P water model solvated system [68]. The neutralization of the system was performed by adding sodium (Na^+) or Chlorure (Cl^-) ions. The energy minimization system was subjected to 50,000 steps using the steepest descent algorithm. Then, the production MD simulations were run for 20 ns for each simulation at a temperature of 300 k, a pressure of 1 bar and a time step of 2 fs.

II. Results and discussion

1. 2D-QSAR analysis

1.1. Model generation

2D-QSAR model was built based on the training set compounds. The PLS method was used to generate the 2D-QSAR model by establishing a linear correlation between the most relevant descriptors (BCUT_SMR_2, LogP (o/w), SlogP_VSA4, and vsurf_IW3) and the inhibitory activity of MAO-B enzyme (pIC50). The linear equation connecting the selected descriptors to MAO-B inhibitory activity is presented below.

$$pIC50 = -3{,}754 + 8{,}469 * BCUT_SMR_2 + 0{,}636 * logP\left(\frac{o}{w}\right) + 0{,}057 * SlogP_VSA4 - 0{,}442 * vsurf_IW3 \quad \textbf{Eq. 15}$$

$N = 20$, $RMSE = 0.28$, $r^2 = 0.88$, $r_{Adjusted}{}^2 = 0.84$, $F\ model = 27.76$, $p < 0.0001$, $\alpha = 5\%$, $r^2_{cv} = 0.80$, $RMSE_{cv} = 0.31$, $r^2_{test} = 0.71$

In the above equation, N is the number of compounds in the training set, RMSE is the root mean square error, and F is the Fisher value.
The best generated model explains 88% ($r^2 = 0.88$) of the total variance in the training set with small values of RMSE and MAE (RMSE = 0.28 and MAE = 0.2). The high value of the F-test (F = 27.76) indicates that the generated model is statistically significant. Also, the significance of each descriptor in **Eq.15** was verified by calculating the value of VIF and p (**Table 2**). All descriptors in the generated model have acceptable values of VIF and p (VIF < 5 and p < 0.05), indicating a good significant relationship between the modeled response and the selected descriptors. The correlation matrix of the selected descriptors is shown in **Table 3**. **Table 4**, shows the values of selected descriptors, experimental activity and predicted activity.

***Table 2**. VIF and p values of the selected descriptors*

Statistique	BCUT_SMR_2	logP(o/w)	SlogP_VSA4	Vsurf_IW3
Tolérance	0.5117	0.6216	0.5946	0.6546
VIF	1.9545	1.6087	1.6817	1.5277
p-value	0.0012	0.0092	<0.0001	0.0007

***Table 3**. Correlation matrix of the selected descriptors*

Variables	BCUT_SMR_2	logP(o/w)	SlogP_VSA4	vsurf_IW3	pIC_{50}
BCUT_SMR_2	1.0000				
logP(o/w)	-0.4640	1.0000			
SlogP_VSA4	-0.5413	0.0986	1.0000		
vsurf_IW3	0.2441	-0.5042	0.1923	1.0000	
pIC_{50}	-0.3386	0.4451	0.7195	-0.3148	1.0000

***Table 4**. Values of selected descriptors, experimental activity and predicted activity*

No	Molecular descriptors				pIC_{50}_Obs	2D-QSAR	
	logP(o/w)	SlogP_VSA4	vsurf_IW3	BCUT_SMR_2		pIC_{50}_Pred	pIC_{50}_Loo
8b	4.9518	33.4189	2.4626	0.7751	6,5968	6.7780	6.8746
9a	4.0580	30.2334	2.8303	0.8189	6.1007	6.2360	6.2645
10a	4.0210	30.2334	2.1943	0.8309	6,3269	6.5958	6.6497
10b	4.9538	33.4189	0.4323	0.7648	7.7958	7.5893	7.4689
10c	4.9928	33.4189	2.0546	0.7876	6.8386	7.0902	7.1913
10e	4.1720	30.2334	1.8544	0.8309	7.2518	6.8420	6.7629
10f	4.6110	30.2334	2.3715	0.8311	7.1611	6.8939	6.8292
10g	3.9750	30.2334	1.8080	0.7969	6.3777	6.4492	6.4735
11g	3.7110	30.2334	2.8445	0.8335	6.2048	6.1328	6.1054
12b	4.0930	3.1856	1.5152	0.8826	5.5497	5.8354	5.8976
15a	4.1648	6.3712	1.8161	0.7838	5.2232	5.0928	5.0162
15b	3.8220	3.1856	3.0879	0.9392	5.2831	5.4470	5.5726
16a	4.6258	6.3712	1.5711	0.7690	5.4723	5.3691	5.3122
16b	4.6648	6.3712	1.2923	0.8091	5.3106	5.8565	5.9500
16c	4.6278	6.3712	0.7767	0.8543	6.5482	6.4438	6.4168
17a	4.5888	6.3712	1.2236	0.8543	6.2365	6.2213	6.2190
18a	4.5908	6.3712	0.2897	0.8543	6.6516	6.6357	6.6276
19a	4.1450	3.1856	2.0563	0.8543	5.2306	5.3892	5.4115
20a	4.1010	3.1856	2.5898	0.8543	5.5867	5.1253	5.0201
20b	4.6910	3.1856	2.3541	0.8718	6.0282	5.7528	5.6333
10d*	4.9558	33.4189	1.4270	0.8270	6.7014	7.6778	-
10h*	4.0140	30.2334	2.1155	0.7972	5.8551	6.3405	-
11a*	3.7570	30.2334	2.9518	0.8543	6.7399	6.2909	-
11b*	4.6898	33.4189	1.6233	0.8175	7.2441	7.3417	-
12a*	4.4358	6.3712	1.5940	0.7690	5.0762	5.2381	-

* Represents compounds of test set, pIC_{50}_Pred represent the value of activity predicted by the created model and pIC_{50_LOO} represent the value of activity predicted by Leave-one-out cross-validation method.

1.2. 2D-QSAR model validation

The predictability of the generated model was verified by internal and external validations. The statistical parameters are listed in **Table 5**. The performance of internal validation was determined by leave-one out cross validation. This approach was used to predict the activity (pIC_{50_LOO}) of each compound in the data set using the model established by (n-1) compounds. Our developed model is predictive, as evidenced by the

results (Q^2_{LOO} =0.81, RMSE $_{LOO}$ =0.31 and MAE_{LOO} = 0.27). The quality of Q^2_{LOO} was also verified by calculating the R^2_m (R^2_m = 0.57 > 0.5) parameter. To ensure the robustness of the developed model, a y- randomization test was applied. Several random shuffles of pIC_{50} were performed. The results are summarized in **Table S1** (see supplementary data). The low average R^2 and Q^2 (R^2 = 0.15 and Q^2 = -0.52) and the high value of $^cR_p^2$ ($^cR_p^2$ = 0.81 > 0.5), indicate that the good results in our original model are not due to a chance correlation. The external validation of the developed model was verified by a test set of 5 compounds. The results (r^2_{test} = 0.71) indicate that the developed model is capable of predicting the activity of new untested compounds. Also, the goodness of external validation was verified by calculating the $Q^2_{(F2)}$ parameter, which gave a value greater than 0.5 ($Q^2_{(F2)}$ = 0.52). The plot of the experimental versus the predicted activities for the internal and the external validations is shown in **Figure 4**.

Table 5. *Accepted 2D-QSAR model validation tools*

Parameters	Interpretation	Acceptable value	Generated 2D-QSAR model
r^2	Coefficient of determination	≥ 0.6	0.88
$r^2_{adjusted}$	Adjusted R-squared	> 0.6	0.84
Q^2_{LOO}	Coefficient of determination for internal validation	> 0.5	0.81
$R^2_{m\ cv}$	R^2_m cross validation parameter	> 0.5	0.57
$^cR_p^2$	Coefficient of determination for *Y*-randomization	> 0.5	0.81
R^2_{test}	Coefficient of determination of external validation	> 0.6	0.71
$Q^2_{(F)}$	The regression slope passing through the origin	> 0.5	0.52
K	The regression slope passing through the origin (plot of experimental versus predicted activities)	0.85≤ k ≤1.15	0.96
K’	The regression slope passing through the origin (plot of predicted versus experimental activities)	0.85≤ k’ ≤1.15	1.04

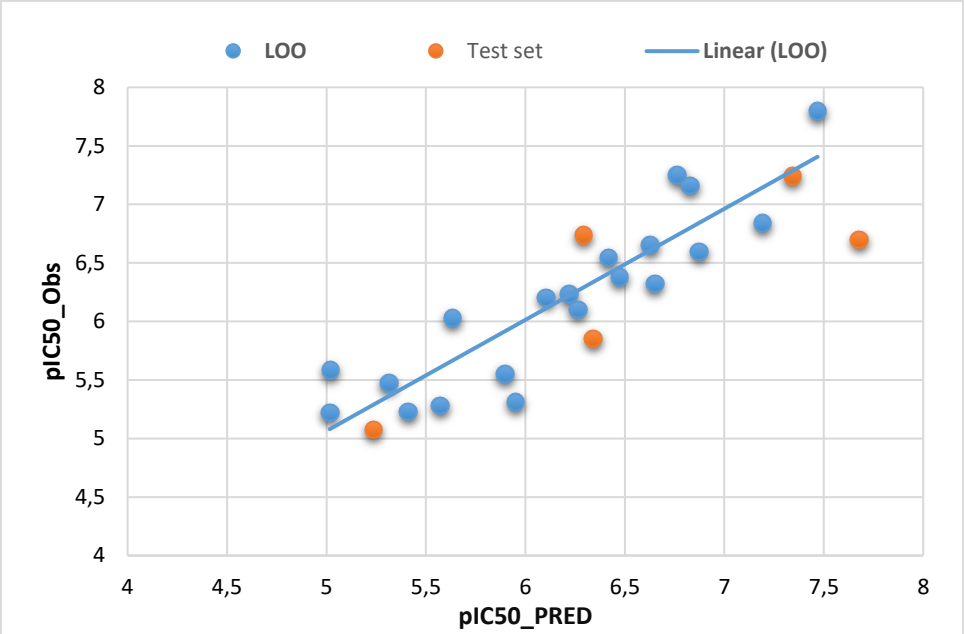


Figure 4. *Plot of predicted and observed activity. The blue dots denote internal validation (LOO), whereas the orange dots represent external validation*

1.3.2D-QSAR descriptors interpretation

In order to understand the relationship between BCUT_SMR_2, LogP (o/w), SlogP_VSA4, and vsurf_IW3 descriptors and the inhibitory activity of the MAO-B enzyme, it's necessary to explain the meaning of each descriptor in **Eq. 15**. **Figure 5**, shows the contribution of each descriptor in **Eq.15**.

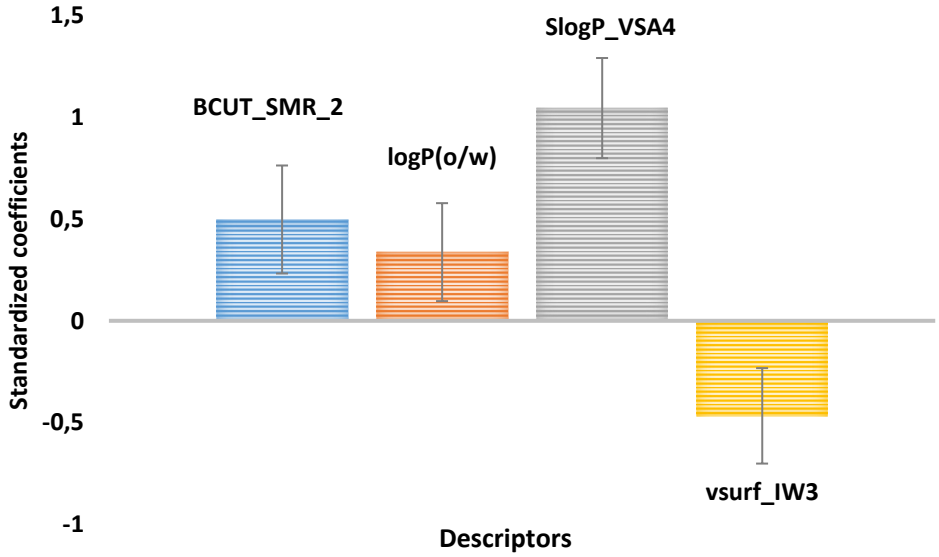


Figure 5. *Contribution of descriptors in the generated 2D-QSAR model*

As described in Figure 5, the most important descriptor in the established 2D-QSAR model is SlogP_VSA4. This descriptor, represents the sum of approximate accessible van

der Waal's surface area i such that logP for atom i is in the range (0.1 to 0.15]. The positive contribution (+ 0.057) and high correlation (r = 0.72) with the inhibitory activity, indicates that compounds with high accessible van der Waals surface area could increase the inhibitory activity of the MAO-B enzyme. A deep analysis of chemical structures and their experimental activity, indicate that the SlogP_VSA4 descriptor is related to the number of fluorine (F) in R1 and R2 positions. From the experimental dataset (**Table 1**), compounds with R1 = F (**16a**, **16b**, **16c**, **17a**, and **18a**) have the same value of SlogP_VSA4 (SlogP_VSA4 = 6.3716) and compounds with R2 = CF3 have the same value of SlogP_VSA4 (SlogP_VSA4 = 33.4189). To ensure that the SlogP_VSA4 descriptor is related to the number of fluorine, we compared compounds **10b** and **11b**, which have the same substituent in the R2 position (R2 = CF3) but differ in the R1 position (**10b**: R1 = -OCH3, **11b**: R1 = OH). The value of the SlogP_VSA4 descriptor is the same in both compounds (SlogP_VSA4 = 33.4189).

The second most important descriptor in the generated 2D-QSAR model is BCUT_SMR_1. The BCUT descriptor using atomic contribution to molar refractivity (using the Wildman and Crippen SMR method) instead of partial charge [69]. The positive influence (+ 8.469) of this descriptor in **Eq. 15**, reflects the importance of atomic properties that govern the intermolecular interactions (atomic charge, atomic polarizability, and atomic hydrogen bonding ability) on the inhibitory activity of MAO-B enzyme [70].

The next descriptor in the generated 2D-QSAR model is Vsurf_IW3, which describes the hydrophilic integy moment, calculated at -0.6 kcal/mol [71]. The integy moment of this descriptor measures the unbalance between the center of mass of a compound and the barycenter of specific regions of the surface [72]. The negative contribution (-0.442) of this descriptor in **Eq. 15**, indicates that the integy moment of compounds should be smaller. To reduce the integy moment, the polar moieties should be close to the center of mass or at the opposite ends of the compound. This may explain the significant increase in activity of compounds **10d** (pIC_{50} = 6.701, ortho-R2 = CF3), 10c (pIC_{50} = 6.8386, meta-R2 = CF3), and **10b** (pIC_{50} =7.7959, para-R2 = CF3). Moreover, compounds with only one hydrated region (**12a**, **12b** and **19a**), show high values of Vsurf_IW3 and consequently, low activity.

The LogP (o/w) (Log of the octanol/water partition coefficient) is a descriptor for measuring the overall hydrophobicity of compounds [73]. The positive contribution (+ 0.636) of this descriptor in **Eq. 15**, indicates that a decrease in the lipophilicity of compounds could increase the inhibitory activity of the MAO-B enzyme.

1.4. Applicability Domain

The applicability domain of the developed 2D-QSAR model was generated, using William's plot (**Figure 6**), taken by MINITAB. 17 software [74]. As described in this figure, all compounds are in the applicability domain, except compound **10d**, which has a standardized residual value greater than $\pm 2\sigma$. Although this compound is considered outside the applicability domain.

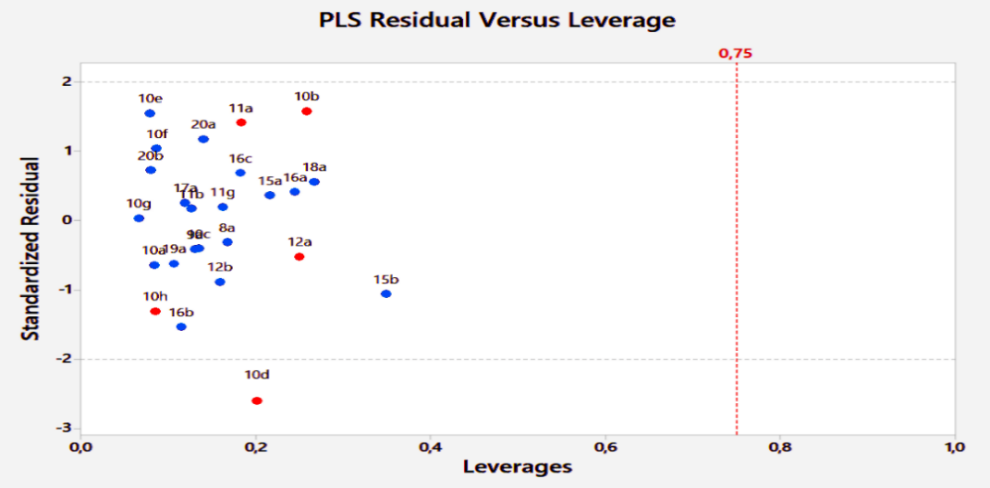


***Figure 6**. Applicability domain of the generated 2D-QSAR model. The blue dots represent the training set compounds and the red dots represent the test set compounds*

2. Physicochemical properties and drug-likeness

Physicochemical properties and drug-likeness results are summarized in **Table S2** (see supplementary data). All compounds show MW values less than 500 Da (224.25 <MW<309.26), LogP values less than 5 (3.28<LogP<4.61), HBA less than 10 (2<HBA<5), HBD less than 5 (0<HBD<2), and TPSA less than 140 Å (26.30 Å <TPSA< 49.33 Å). This result revealed that these compounds are very likely to be orally active. Also, the results of drug likeness showed that all compounds respect the Lipinski rule of five, Veber, Egan, and Muegge rules without any violations. However, compounds **8b**, **9b**, **10b**, **10c**, **16a**, **16b**, **16c**, **17a** and **18a** showed one violation for the Ghose rule (WLOGP>5.6).

3. ADMET prediction

The pharmacokinetic (PK) properties and toxicity were envaulted using ADMET analysis. The results are summarized in **Table S3** (see supplementary data). The moderate values of LogBB reveal that these compounds may be effective for treating the neurodegenerative disease. The values of human intestinal absorption (HIA) (HIA > 30%) indicate that these compounds are highly absorbed. Also, the absorption rate of the studied compounds was evaluated by the Caco-2 cell permeability parameter. The results indicate that all compounds show high Caco-2 cell

permeability (Caco-2 > 0.90). In addition, the metabolism of the studied compounds was verified by the inhibitory or substrate behavior of the cytochrome P450 enzymes (CYPs). This enzyme plays a major role in the oxidation process and facilitates the excretion of foreign organic compounds, including drugs. All the compounds were found to be a substrate of 3A4, while no compounds were found to be a substrate of 2D6. Moreover, results of inhibition studies indicate that no compounds were found to inhibit 2D6, whereas, all of them were found to inhibit 1A2, 2C19, and 2C9, except compounds **11a**, **11g** and **19a**. The toxicity analysis, including AMES toxicity and hepatoxicity indicates that the predicted compounds are not harmful. However, the maximum tolerated dose in human was in the range of 0.26-1.23 mg/kg/day.

4. Molecular docking analysis

4.1. Docking validation

In order to validate the binding site of the crystal structure, docking validation (or re-docking) was applied. First, the native ligand was extracted from its PBD structure (PDB ID: 2BK3). The docking parameters were generated using AutoDock tools. Then, the native ligand was re-docked into the same binding site of the crystal structure, using AutoDock 4.1. The best obtained pose gave the root mean square deviation (RMSD) value of 1.57, using 50 iterations. The superposition between the native and the re-docked ligands is shown in **Figure S1** (see supplementary data).

4.2. Binding modes interactions and affinity of the studied compounds

After docking validation, the dataset of compounds was used to generate the docking study. The results are summarized in **Table S4** (see supplementary data). The values of binding free energy revealed that unsaturated ketone derivatives showed significant stability in the binding site of the MAO-B enzyme. In addition, the correlation between the binding free energy and the pIC_{50} values ($r^2 = 0.73$) showed that the docking results are in agreement with the activity values. **Figure 7**, illustrates the correlation between the binding free energy and pIC_{50} values.

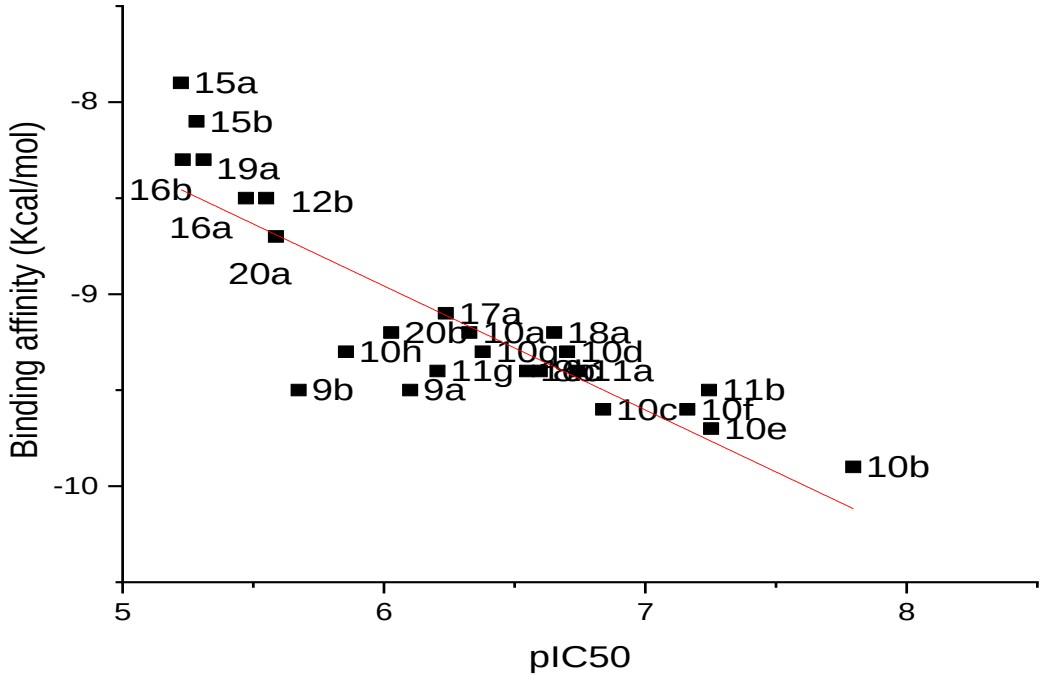


Figure 7. *Correlation between the binding free energy values and the pIC50 values*

In order to predict whether and how unsaturated ketone derivatives bind to the MAO-B active site, compounds **10b** and **10e** (the most potent inhibitors in the dataset) were used. The results are shown in **Figures 8** and **9**. As described in **Figure 8**, compound **10b** was fixed in the binding pocket of the MAO-B enzyme by several types of interactions. The most important are the conventional hydrogen bond interaction with CYS 172, the halogen interaction with GLY 204, and hydrophobic interactions (Pi-Pi stacked, Pi-Pi T-shaped, Alkyl, Pi-alkyl, and Pi-sigma) with TYR 398, PHE 343, TYR 326, CYS 172, ILE 199, and LEU 171. These types of interactions may confirm why this compound displays a low binding affinity (-9.9 Kcal/mol). The low binding affinity of compound **10e** (-9.7 Kcal/mol) was also related to the type of interactions that it has with the MAO-B receptor. As described in **Figure 9**, compound **10e** was fixed in the binding pocket of the MAO-B receptor with the same type of interactions (halogen interactions with the fluorine at the ortho-R1 position, conventional hydrogen bond interaction with the ketone group of α, β-unsaturated, and hydrophobic interactions at the R2 position) as that of compound **10b**. The values of binding affinity of the studied compounds are related to the number of interactions, the types of interactions, and the distance between the ligand and a particular amino acid.

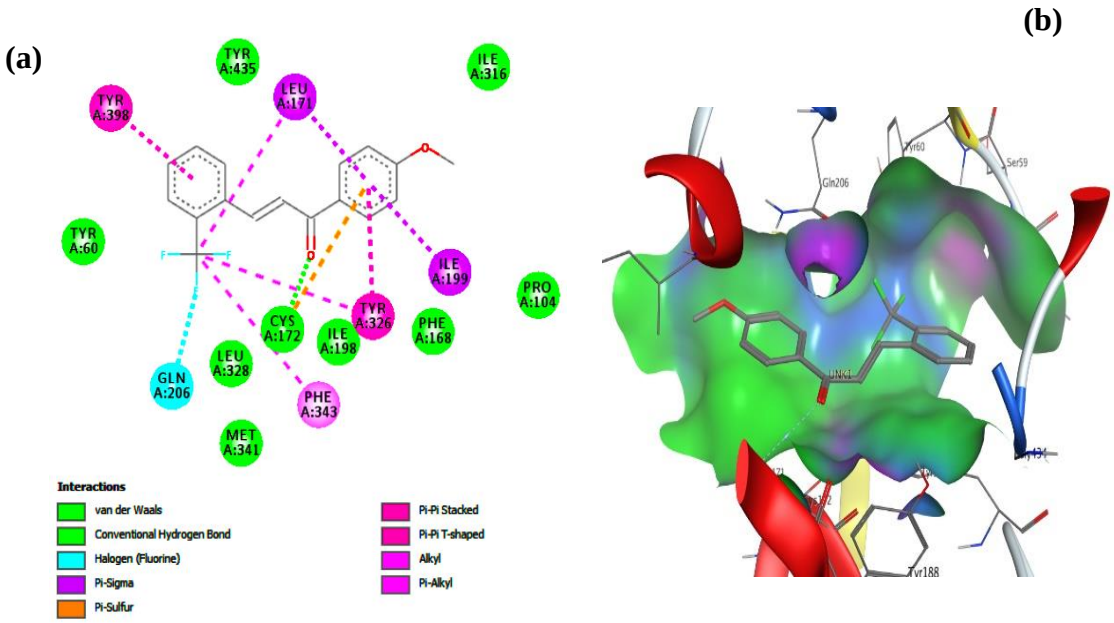


Figure 8. *Docking analysis of compound 10b. (a) 2D view of binding site interactions, (b) 3D view of the binding conformation*

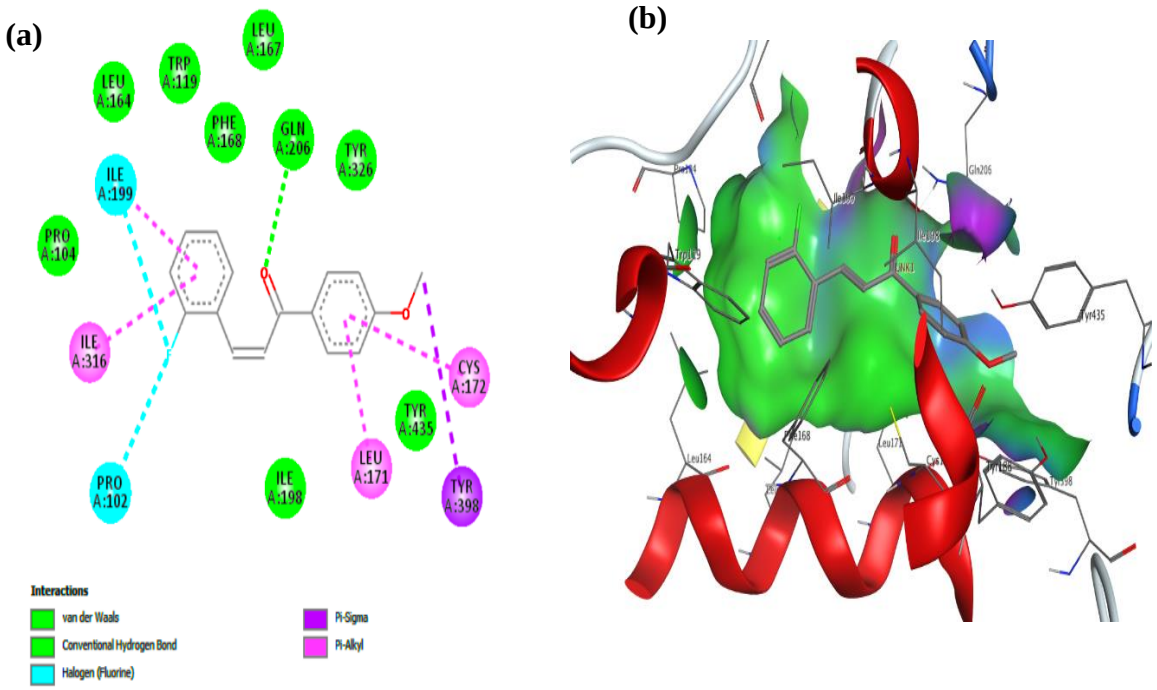


Figure 9. *Docking analysis of compound **10e**. **(a)** 2D view of binding site interactions, **(b)** 3D view of the binding conformation*

5. Molecular dynamics simulation

In order to validate the results of molecular docking and to affirm the stability of docked compounds in the binding pocket of the MAO-B receptor, the most active compound in the dataset (compound 10b) was subjected to MD simulation. The system was employed for 20 ns time scale simulation. The results are shown in **Figure 10**. The plot of the root means square deviation (RMSD) (**Figure 10a**) indicates that the 10b_2BK3 complex attained stability at 7 ns. Then, it remained constant during the MD simulation. The average RMSD value is 0.26 nm with a maximum value of 0.49 nm and a minimum value of 0.37 nm. The Root Means Square Fluctuation (RMSF) of the 10b_2BK3 complex is also determined (**Figure 10b)**. This parameter is an indicator of residual flexibility. From the results of the RMSF, we conclude that all amino acid residues had RMSF values less than 0.1 nm except for residues between 470-500, which had RMSF values more than 0.1nm. The stability of the system was also verified by plotting the radius of gyration (Rg) (**Figure 10c**). The Rg value of the complex is in the range of 5.15 – 5.2 nm suggesting that these complexes are stable and compact during the 20 ns of the MD simulation. In addition, and from **Figure 10d,** we can see that the main secondary structural elements of the ligand and the protein in the 10b_2BK3 complex remained close to its initial structure before the MD simulation.

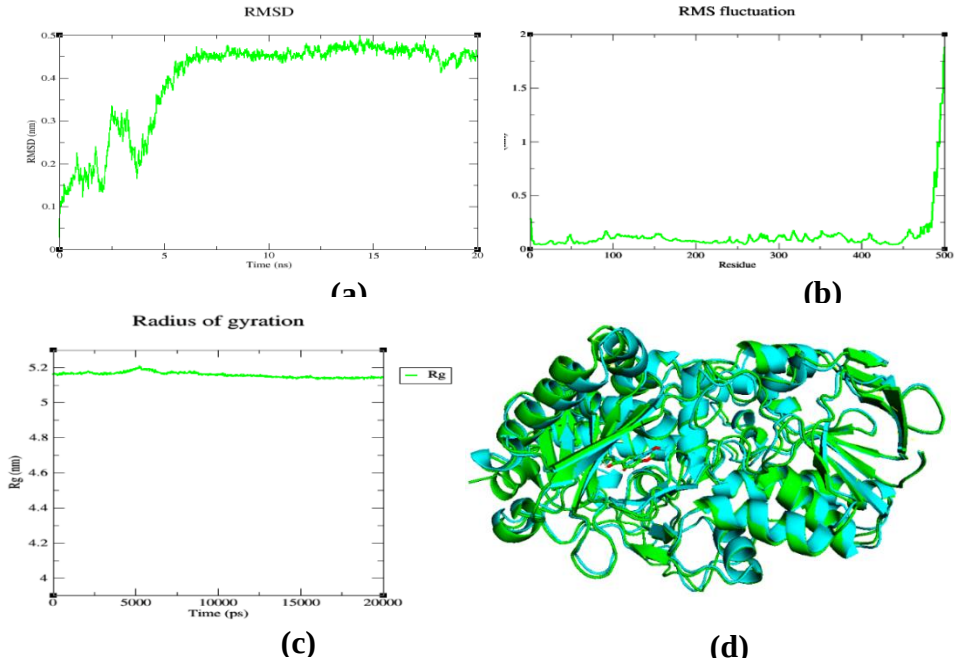


***Figure 10**. Molecular dynamics simulations of the 10b_2BK3 complex. **(a)** represents the RMSD plot, **(b)** represents the RMSF plot, **(c)** represents the Rg plot, and **(d)** represents the superposition of the final complex structure after MD simulation of 20 ns (green color) and initial complex structures before MD simulation (blue color)*

III. Conclusion

In this paper, molecular modeling studies including 2D-QSAR, ADMET, molecular docking, and molecular dynamics simulation were applied on a series of unsaturated ketone derivatives as MAO-B inhibitors. The PLS technique was utilized to build the 2D QSAR model, which employed the pIC50 values as a dependent variable and the MOE descriptors as independent variables. The results revealed the importance of BCUT_SMR_2, logP (o/w), SlogP_VSA4 and vsurf_IW3 descriptors in describing the activity values. The pharmacokinetic characteristics of compounds were also studied using drug likeness and ADMET prediction. Molecular docking was carried out to investigate the binding site interactions between ligands and the MAO-B receptor. The activity data exhibited a greater correlation ($r^2 = 0.73$) with the binding free energy values. Finally, the active compounds were evaluated using MD simulation. Finally, the active compounds were evaluated using MD simulation. The findings of this work clearly demonstrate the importance of unsaturated ketone derivatives in inhibiting MAO-B enzymatic activity, and they may open the way for the development of other inhibitory derivatives as Parkinson's disease possible treatments.

Acknowledgment

The authors want to thank the Moroccan Association of Theoretical Chemistry (MATC) for its relevant help regarding to the Software.

Conflict of interest

All authors declare that they have no conflict of interest in this work.

IV. References

1. Shih J C, Chen K, Ridd M J. MONOAMINE OXIDASE : From Genes to Behavior. Annu Rev Neurosci. 1999;22:197–217. Doi: https://doi.org/10.1146/annurev.neuro.22.1.197.

2. Denney R M. Assignment of Genes for Human Monoamine Oxidases A and B to the X Chromosome. J Neurosci Res. 1986;616:601–16. Doi: https://doi.org/10.1016/B978-0-7020-3373-5.00010-1.

3. Bach Alfred W J, Lant Nancy C, Johnsont Deborah L, et al. cDNA cloning of human liver monoamine oxidase A and B : Molecular basis of differences in enzymatic properties. Proc Nat! Acad Sci. 1988;85(15):4934–8. Doi: https://doi.org/10.1073/pnas.85.13.4934.

4. Cesura By Andrea M, Pletscher Alfred. The new generation of monoamine oxidase inhibitors. Prog Drug Res / Fortschritte Der Arzneimittelforsch / Progrès Des Rech Pharm. 1992;38:224–5. Doi: https://doi.org/10.1007/978-3-0348-7141-9_3.

5. Jegham Samir, George Pascal, Recherche Synthélabo, Snc Direction De Recherche, Carrières Rue. Monoamine oxidase A and B inhibitors. Expert Opin Ther Patents Jegham. 1998;8(9):1143–50. Doi: https://doi.org/10.1517/13543776.8.9.1143.

6. Academy Swiss, Sciences Medical. The discovery of antidepressants: A winding path. A Pletscher. 1991;47. Doi: https://doi.org/10.1007/BF02041242.

7. D Sidney Zisook M. A clinical overview of monoamine oxidase inhibitors. Psychosomatics.

1985;26(3):240-246,251. Doi: 10.1016/S0033-3182(85)72877-0.

8. Marin Deborah B, Bierera Linda M, Lawlorb Brian A, et al. L-Deprenyl and physostigmine for the treatment of Alzheimer ' s disease. Psychiatry Res. 1995;58(3). Doi: https://doi.org/10.1016/0165-1781(95)02714-8.

9. Jo Seonmi, Yarishkin Oleg, Hwang Yu Jin, et al. ARTICLES GABA from reactive astrocytes impairs memory in mouse models of Alzheimer ' s disease. Nat Publ Gr. 2014;20:886–896. Doi: 10.1038/nm.3639.

10. Waibel Stefan, Reuter Alexander, Ludolph Albert C. Rasagiline alone and in combination with riluzole prolongs survival in an ALS mouse model. J Neurol. 2004;251:1080–4. Doi: 10.1007/s00415-004-0481-5.

11. Halgren Thomas A, Murphy Robert B, Friesner Richard A, et al. Glide : A New Approach for Rapid , Accurate Docking and Scoring . 2 . Enrichment Factors in Database Screening. J Med Chem. 2004;47(7):1750–9. Doi: https://doi.org/10.1021/jm0310885.

12. Fabbri Margherita, Rosa Mario M, Abreu Daisy, Ferreira Joaquim J. Clinical pharmacology review of safinamide for the treatment of Parkinson ' s disease. Futur Med. 2015;5. Doi: https://doi.org/10.2217/nmt.15.46.

13. Salum Lívia B, Altei Wanessa F, Chiaradia Louise D, et al. European Journal of Medicinal Chemistry migration inhibitors. Eur J Med Chem. 2013;63:501–10. Doi: 10.1016/j.ejmech.2013.02.037.

14. Israf D A, Khaizurin T A, Syahida A, Lajis N H, Khozirah S. Cardamonin inhibits COX and iNOS expression via inhibition of p65NF- □ B nuclear translocation and I □ -B phosphorylation in RAW 264 . 7 macrophage cells. Mol Immunol. 2007;44:673–9. Doi: 10.1016/j.molimm.2006.04.025.

15. Mahapatra Debarshi Kar, Asati Vivek, Bharti Sanjay Kumar. European Journal of Medicinal Chemistry Chalcones and their therapeutic targets for the management of diabetes : Structural and pharmacological perspectives. Eur J Med Chem. 2015;92:839–65. Doi: 10.1016/j.ejmech.2015.01.051.

16. Mallikharjuna Rao Lambu, Suresh Kumar, Syed Khalid Yousuf, Deepak K Sharma, Altaf hussain, Ajay Kumar, Fayaz Malik and Debaraj Mukherjee. Medicinal chemistry of dihydropyran based medium ring macrolides related to aspergillides : selective inhibition of PI3K α ! J Med Chem. 2013;56(15):6136–6145. Doi: 10.1021/jm400515c.

17. Ramadan Mohamed, Sayed El, Hamadah Hamadah, El Abd, Fodah Razek. European Journal of Medicinal Chemistry Antiobesity , antioxidant and cytotoxicity activities of newly synthesized chalcone derivatives and their metal complexes. Eur J Med Chem. 2014;76:517–30. Doi: 10.1016/j.ejmech.2014.02.021.

18. Sinha Shweta, Medhi Bikash, Sehgal Rakesh. Chalcones as an Emerging Lead Molecule for Antimalarial Therapy : A Review. J Mod Med Chem. 2013;1:64–77.

19. Sharma Horrick, Patil Shivaputra, Sanchez Tino W, Neamati Nouri, Schinazi Raymond F, Buolamwini John K. Bioorganic & Medicinal Chemistry Synthesis , biological evaluation and 3D-QSAR studies of 3-keto salicylic acid chalcones and related amides as novel HIV-1 integrase inhibitors. Bioorg Med Chem. 2011;19(6):2030–45. Doi: 10.1016/j.bmc.2011.01.047.

20. Arasappan Ashok, Padilla Angela I, Jao Edwin, et al. Toward Second Generation Hepatitis C Virus NS3 Serine Protease Inhibitors : Discovery of Novel P4 Modified Analogues with Improved Potency and Pharmacokinetic Profile. J Med Chem. 2009;52:2806–17. Doi: https://doi.org/10.1021/jm801590u.

21. Newman David J., Cragg Gordon M. Natural products as sources of new drugs over the last 25 years. J Nat Prod. 2007;70(3):461–77. Doi: 10.1021/np068054v.

22. Scapin Giovanna. Structural Biology and Drug Discovery. Curr Pharm Des. 2006;12(17):2087–97. Doi: 10.2174/138161206777585201.

23. Drie John H. Computer-aided drug design: The next 20 years. J Comput Aided Mol Des. 2007;21(10–11):591–601. Doi: 10.1007/s10822-007-9142-y.

24. Zhou Haibin, Liu Liu, Huang Jing, et al. Structure-based design of high-affinity macrocyclic peptidomimetics to block the menin-mixed lineage leukemia 1 (MLL1) protein-protein interaction. J Med Chem. 2013;56(3):1113–23. Doi: 10.1021/jm3015298.

25. Fung Ho Yee Joyce, Chook Yuh Min. Atomic basis of CRM1-cargo recognition, release and inhibition. Semin Cancer Biol. 2014;27:52–61. Doi: 10.1016/j.semcancer.2014.03.002.

26. Ran Xu, Liu Liu, Yang Chao Yie, et al. Design of High-Affinity Stapled Peptides to Target the Repressor Activator Protein 1 (RAP1)/Telomeric Repeat-Binding Factor 2 (TRF2) Protein-Protein Interaction in the Shelterin Complex. J Med Chem. 2016;59(1):328–34. Doi: 10.1021/acs.jmedchem.5b01465.

27. Kargbo Robert B. Novel Triterpenone for Treatment of Viral Diseases-HIV Inhibitors. ACS Med Chem Lett. 2018;9(4):298–9. Doi: 10.1021/acsmedchemlett.8b00113.

28. Nagasaka Misako, Ge Yubin, Sukari Ammar, Kukreja Geetika, Ou Sai Hong Ignatius. A user's guide to lorlatinib. Crit Rev Oncol Hematol. 2020;151(April):102969. Doi: 10.1016/j.critrevonc.2020.102969.

29. Huang Hung Jin, Lee Kuei Jen, Yu Hsin Wei, et al. Structure-based and ligand-based drug design for her 2 receptor. J Biomol Struct Dyn. 2010;28(1):23–37. Doi: 10.1080/07391102.2010.10507341.

30. Druker B.J. Lydon N.B. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. J Clin Invest. 2000;105(1):3–7. Doi: 10.1172/JCI9083.

31. Wlodawer Alexander, Vondrasek Jiri. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. Annu Rev Biophys Biomol Struct. 1998;27:249–84. Doi: 10.1146/annurev.biophys.27.1.249.

32. Clark David E. What has computer-aided molecular design ever done for drug discovery? Expert Opin Drug Discov. 2006;1(2):103–10. Doi: 10.1517/17460441.1.2.103.

33. Depeursinge Adrien, Racoceanu Daniel, Iavindrasana Jimison, et al. Fusing Visual and Clinical Information for Lung Tissue Classification in HRCT Data. Artif Intell Med. 2010;10:ARTMED1118. Doi: 10.1016/j.

34. Liu Bo, Tsutsui Makusu, Taniguchi Masateru. Measuring single-molecule conductance at an Ultra-Low molecular concentration in vacuum. Micromachines. 2018;9(6). Doi: 10.3390/mi9060282.

35. Choi Ji Won, Jang Bo Ko, Cho Nam-chul, et al. Synthesis of a Series of Unsaturated Ketone Derivatives as Selective and Reversible Monoamine Oxidase Inhibitors. Bioorg Med Chem. 2015;23(19):6486–96. Doi: 10.1016/j.bmc.2015.08.012.

36. Tong Jianbo, Zhan Pei, Wang Xiang Simon, Wu Yingji. Quionolone carboxylic acid derivatives as HIV-1 integrase inhibitors: Docking-based HQSAR and topomer CoMFA analyses. J Chemom. 2017;31(12):1–13. Doi: 10.1002/cem.2934.

37. Clark Matthew, Cramer Richard D., Van Opdenbosch Nicole. Validation of the general purpose tripos 5.2 force field. J Comput Chem. 1989;10(8):982–1012. Doi: 10.1002/jcc.540100804.

38. Caballero J., Saavedra M., Fernández M., González-Nilo F.D. Quantitative structure-activity relationship of rubiscolin analogues as δ opioid peptides using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). J Agric Food Chem. 2007;55(20):8101–4. Doi: 10.1021/jf071031h.

39. El Aissouq Abdellah, Toufik Hamid. QSAR study of isonicotinamides derivatives as Alzheimr ' s disease inhibitors using PLS-R and ANN methods. 2019 Int Conf Intell Syst Adv Comput Sci. 2019:1–7. Doi: 10.1109/ISACS48493.2019.9068919.

40. El Aissouq Abdellah, Toufik Hamid, Stitou Mourad, Ouammou Abdelkrim, Lamchouri Fatima. In

Silico Design of Novel Tetra-Substituted Pyridinylimidazoles Derivatives as c-Jun N-Terminal Kinase-3 Inhibitors, Using 2D/3D-QSAR Studies, Molecular Docking and ADMET Prediction. Int J Pept Res Ther. 2020;26:1335–1351. Doi: 10.1007/s10989-019-09939-8.

41. El Aissouq Abdellah, Chedadi Oussama, Kasmi Rania, et al. Molecular Modeling Studies of C-Glycosylfavone Derivatives as GSK-3β Inhibitors Based on QSAR and Docking Analysis. J Solution Chem. 2021;50(5):808–22. Doi: 10.1007/s10953-021-01083-6.

42. Zhihua Lin, Yuzhang Wu, Bo Zhu, Bing Ni, Li Wang. Toward the Quantitative Prediction of T-Cell Epitopes: QSAR Studies on Peptides Having Affinity with the Class I MHC Molecular HLA-A*0201. J Comput Biol. 2005;11(4):683–94. Doi: 10.1089/cmb.2004.11.683.

43. Kunal Roy, Supratik Kar, Rudra, Narayan Das. A Primer on QSAR/QSPR Modeling. New York: Springer International Publishing. 2015.

44. Roy Partha Pratim, Roy Kunal. On Some Aspects of Variable Selection for Partial Least Squares Regression Models. QSAR Comb Sci. 2008;27(3):302–13. Doi: 10.1002/qsar.200710043.

45. Shukla Aparna, Tyagi Rekha, Meena Sanjeev, Datta Dipak, Srivastava Santosh Kumar, Khan Feroz. 2D- and 3D-QSAR modelling, molecular docking and in vitro evaluation studies on 18β-glycyrrhetinic acid derivatives against triple-negative breast cancer cell line. J Biomol Struct Dyn. 2019;38(1):168–85. Doi: 10.1080/07391102.2019.1570868.

46. Tropsha Alexander. Best Practices for QSAR Model Development , Validation , and Exploitation. Mol Informatics,. 2010;29:476–88. Doi: 10.1002/minf.201000061.

47. K. Roy S. Kar and R. A Primer on QSAR/QSPR Modeling: Fundamental Concepts. Boston: Academic Press. 2015.

48. Golbraikh Alexander, Tropsha Alexander. Beware of q^2! J Mol Graph Model. 2002;20(4):269–76. Doi: 10.1016/S1093-3263(01)00123-1.

49. Schu Gerrit, Ebert Ralf-uwe, Chen Jingwen, Wang Bin, Ku Ralph. External Validation and Prediction Employing the Predictive Squared Correlation Coefficient. J Chem Inf Model. 2008;48(11):2140–5.

50. Gramatica Paola. Principles of QSAR models validation: Internal and external. QSAR Comb Sci. 2007;26(5):694–701. Doi: 10.1002/qsar.200610151.

51. Goudzal Amina, Aissouq Abdellah El, Hamdani Hicham El, Ouammou Abdelkrim. Materials Today : Proceedings QSAR modeling , molecular docking studies and ADMET prediction on a series of phenylaminopyrimidine- (thio) urea derivatives as CK2 inhibitors. Mater Today Proc. 2020;(xxxx). Doi: 10.1016/j.matpr.2020.08.044.

52. Pires Douglas E.V., Blundell Tom L., Ascher David B. pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. J Med Chem. 2015;58(9):4066–72. Doi: 10.1021/acs.jmedchem.5b00104.

53. Daina Antoine, Michielin Olivier, Zoete Vincent. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7(October 2016):1–13. Doi: 10.1038/srep42717.

54. Dong Xin, Zhang Ze Ming, Liu Feng, et al. Metamorphic rocks of the causes of southeastern Lhasa terrane and the Mesozoic - Cenozoic orogeny. Acta Petrol Sin. 2012;28(6):1765–84.

55. Dalvit Claudio, Fasolini Marina, Flocco Maria, Knapp Stefan, Pevarello Paolo, Veronesi Marina. Spectroscopy Experiments : Detection of High-Affinity Ligands. J Med Chem. 2002;45:2610–4.

56. Yan Bing, Gremlich Hans Ulrich, Moss Serge, Coppola Gary M., Sun Qun, Liu Lina. A comparison of various FTIR and FT Raman methods: Applications in the reaction optimization stage of combinatorial chemistry. J Comb Chem. 1999;1(1):46–54. Doi: 10.1021/cc980003w.

57. El Aissouq Abdellah, Chedadi Oussama, Bouachrine Mohammed, Ouammou Abdelkrim. Identification

of Novel SARS-CoV-2 Inhibitors: A Structure-Based Virtual Screening Approach. J Chem. 2021;2021. Doi: 10.1155/2021/1901484.

58. Kasmi Rania, Hadaji Elghalia, Chedadi Oussama, El Aissouq Abdellah, Bouachrine Mohammed, Ouammou Abdelkrim. of CDK (anticancer activity) with an application of the molecular docking method. Heliyon. 2020;6(March):e04514. Doi: 10.1016/j.heliyon.2020.e04514.

59. El Mchichi L., El Aissouq A., Kasmi R., et al. In silico design of novel Pyrazole derivatives containing thiourea skeleton as anti-cancer agents using: 3D QSAR, Drug-Likeness studies, ADMET prediction and molecular docking. Mater Today Proc. 2021;(xxxx). Doi: 10.1016/j.matpr.2021.03.152.

60. Binda Claudia, Khalil Ashraf, Li Min, Mattevi Andrea, Castagnoli Neal, Edmondson Dale E. Demonstration of Isoleucine 199 as a Structural Determinant for the Selective Inhibition of Human Monoamine Oxidase B by Specific Reversible Inhibitors *. J Biol Chem. 2005;280(16):15761–6. Doi: 10.1074/jbc.M500949200.

61. Morris Garrett M, Huey Ruth, Lindstrom William, et al. Software News and Updates AutoDock4 and AutoDockTools4 : Automated Docking with Selective Receptor Flexibility. J Comput Chem. 2009;30(16):2785–91. Doi: 10.1002/jcc.

62. Trott Oleg, Olson Arthur J. Software News and Update AutoDock Vina : Improving the Speed and Accuracy of Docking with a New Scoring Function , Efficient Optimization , and Multithreading. J Comput Chem. 2009;31(2):455–61. Doi: 10.1002/jcc.

63. DeLano WL (2008) The PyMOL Molecular Graphics System, Version 2.0 Schrödinger LLC. PyMOL | pymol.org. Available at: https://pymol.org/2/. Accessed January 17, 2019, n.d.

64. Discovery Studio Predictive Science Application | Dassault Systèmes BIOVIA. Available at: https://www.3dsbiovia.com n.d.

65. Abraham Mark James, Murtola Teemu, Schulz Roland, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX. 2015;1–2:19–25. Doi: 10.1016/j.softx.2015.06.001.

66. K. VANOMMESLAEGHE, E. HATCHER, C. ACHARYA, S. KUNDU, S. ZHONG, J. SHIM, E. DARIAN, O. GUVENCH, P. LOPES, I. VOROBYOV A. D. MACKERELL JR. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. J Comput Chem. 2010;31(4):671–90. Doi: https://doi.org/10.1002/jcc.21367.

67. Best Robert B, Zhu Xiao, Shim Jihyun, et al. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ 1 and χ 2 Dihedral Angles. J Chem Theory Comput. 2012;8(9):3257–73. Doi: https://doi.org/10.1021/ct300400x.

68. Jorgensen William L, Chandrasekhar Jayaraman, Madura Jeffry D, et al. Comparison of simple potential functions for simulating liquid water Comparison of simple potential functions for simulating liquid water. J Chem Phys. 1983;926(79). Doi: 10.1063/1.445869.

69. Wildman Scott A, Crippen Gordon M. Three-dimensional molecular descriptors and a novel QSAR method. J Mol Graph Model. 2002;21(February):161–70. Doi: https://doi.org/10.1016/S1093-3263(02)00147-X.

70. Introduction A N, Chemoinformatics T O. AN INTRODUCTION TO CHEMOINFORMATICS. springer. 2007.

71. Sulistyo Bambang, Sudarmanto A R I, Yuswanto Agustinus, Susidarti Asmah, Noegrohati S R I. Mole cu la r Mod eling of H um an 3 β-Hy droxy steroid Dehydrogenase Type 2 : Combined Homology Modeling , Docking and QSAR Approach (Pemodelan Molekular Enzim 3β-Hydroxysteroid Dehydrogenase Tipe 2 : Pemodelan Kombinasi Homologi , Docking dan Pendekatan. J ILMU KEFARMASIAN Indones. 2017;15(1):7–16.

72. Meng-lund Helena, Kasten Georgia, Tarp Katrine, Poso Antti, Pantsar Tatu. European Journal of Pharmaceutical Sciences The use of molecular descriptors in the development of co-amorphous formulations. Eur J Pharm Sci. 2018;119(February):31–8. Doi: 10.1016/j.ejps.2018.04.014.

73. Shahapurkar S, Pandya T, Kawathekar N, Chaturvedi S C. Quantitative structure activity relationship studies of diaryl furanones as selective COX-2 inhibitors. Eur J Med Chem. 2004;39(4):383–8. Doi: 10.1016/j.ejmech.2003.12.007.

74. Http://www.minitab.com. Outils de statistiques, d'amélioration de processus et d'analyse de données

Minitab. Available at: http://www.minitab.com/fr-fr/. Accessed July 3, 2019, n.d.

Chapitre II: Développement de nouveaux inhibiteurs de la monoamine oxydase B (MAO-B) par la combinaison de docking, du 3D-QSAR, de la prédiction ADMET, de la simulation par dynamique moléculaire et de l'énergie libre de liaison MM_GBSA

Résumé :

Cette deuxième publication consiste à étudier la même série de molécule précédente (étudiée en publication 1) en utilisant cette fois la méthode d'alignement basé sur le docking, la méthode 3D-QSAR, l'approche ADMET et la simulation par dynamique moléculaire. L'objectif est de trouver un modèle prédictif qui permettrait d'expliquer la variation de l'activité expérimentale et, par conséquent, la prédiction de nouvelles molécules avec des propriétés pharmaceutiques et biologiques élevées.

Tout d'abord, le docking moléculaire a été appliqué pour générer les meilleures orientations de chaque molécule dans le site de liaison de récepteur MAO-B. Ensuite, toutes les molécules ont été alignées sur leur structure commune en utilisant la molécule 10b comme référence (template). Les modèles 3D-QSAR ont été construits en utilisant les méthodes CoMFA (comparative molecular field analysis) et CoMSIA (comparative molecular similarity index analysis). De plus, les molécules prédites ont été dockées dans le site de liaison de MAO-B, et leur stabilité a été évaluée par la méthode de dynamique moléculaire. Enfin, la biodisponibilité orale des molécules prédites a été examinée par l'analyse des propriétés ADMET.

Les résultats du docking moléculaire ont indiqué que les composés ayant une forte efficacité inhibitrice avaient également une affinité de liaison élevée. Les études 3D-QSAR ont montré l'importance des champs stériques, électrostatiques et des accepteurs de liaison hydrogène sur l'activité inhibitrice de la MAO-B. En se basant sur les résultats de cette publication et ceux de la publication précédente, une série de quatre molécules (T1, T2, T3 et T4), inhibiteurs de la MAO-B a été prédite (figure a1).

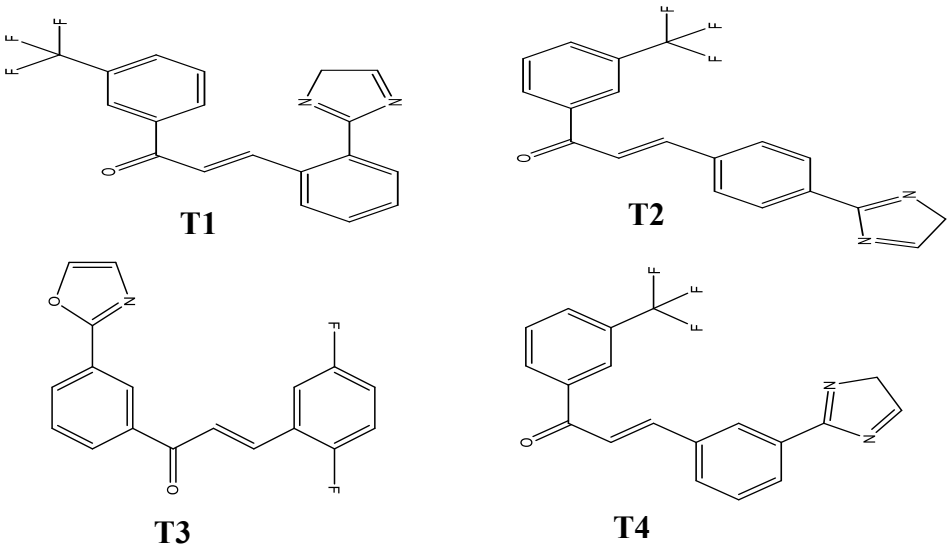


Figure a1. *Molécules prédites*

Toutes les molécules prédites ont montré une activité anti-MAO-B significative par rapport à la molécule la plus active de la série (10b). De plus, la stabilité de ces composés prédits a été évaluée en utilisant la simulation par dynamique moléculaire. Les résultats ont montré que les dérivés cétoniques insaturés sont stables et compacts pendant les 100 ns de la simulation MD. Enfin, l'analyse ADMET montre que ces composés ont des propriétés pharmaceutiques prometteuses.

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
https://doi.org/10.1080/07391102.2022.2071134

Taylor & Francis
Taylor & Francis Group

Check for updates

Development of novel monoamine oxidase B (MAO-B) inhibitors by combined application of docking-based alignment, 3D-QSAR, ADMET prediction, molecular dynamics simulation, and MM_GBSA binding free energy

Abdellah El Aissouq[a], Oussama Chedadi[b], Mohammed Bouachrine[c], Abdelkrim Ouammou[b] and Fouad Khalil[a]

[a]LPME Laboratory, Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco;

[b]LIMOME Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco;

[c]MCNS Laboratory, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco

Communicated by Ramaswamy H. Sarma

ABSTRACT

Unsaturated ketone derivatives are known as inhibitors of monoamine oxidase B (MAO-B), a potential drug target of Parkinson's disease. Here, docking-based alignment, 3 D-QSAR (three-dimensional quantitative structure-activity relationship) studies, ADMET (absorption, distribution, metabolism, excretion, and toxicity) prediction, molecular dynamics (MD) simulation, and MM_GBSA binding free energy were performed on a novel series of MAO-B inhibitors. The objective is to predict new MAO-B inhibitors with high potency activity. The 3 D-QSAR models were created using comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA). Molecular docking findings indicated that compounds with strong inhibitory efficacy also had a high binding affinity. 3 D-QSAR studies showed the importance of steric, electrostatic, and H-bond acceptor fields on the inhibitory activity of MAO-B. Based on the appropriate 3 D-QSAR model, a new series of MAO-B inhibitors were predicted and their pharmacokinetic characteristics were evaluated using in silico ADMET prediction. All screened compounds show good oral bioavailability without any side effects. Moreover, the dynamic behavior and stability of the most active compounds were evaluated using MD simulations. The results showed that unsaturated ketone derivatives are stable and compact during the 100 ns of MD simulation. Finally, the binding free energy of complexes was determined using the MM_GBSA method; the findings indicated that the T1 compound is more stable ($DG_{binding}$ ¼409.506 KJ/mol) than the data set's highest active compound ($DG_{binding}$ ¼31.883 KJ/mol).

ARTICLE HISTORY
Received 18 January 2022
Accepted 23 April 2022

KEYWORDS
MAO enzyme; unsaturated ketone derivatives; dockingbased alignment; ADMET; MD simulations

I. Introduction

Monoamine oxidases (MAOs) are mammalian flavoenzymes that convert monoamine neurotransmitters such as dopamine, serotonin, epinephrine, and norepinephrine into their corresponding aldehydes with the formation of hydrogen peroxide (H_2O_2) and ammonia (NH3) (Scrutton, 2004)(Shih et al., 1999) (Figure 1). In mammals, two isoforms of MAO (MAO-A and MAO-B) have been identified. They share 70% of amino acid identity (Bach et al., 1988), but differ in tissue localization, relative expression, inhibitor affinity, and substrate specificities (Tipton, 2018). Human MAO-A (527 aa) predominates in the cytosol and the outer mitochondrial membrane, whereas MAO-B (520 aa) predominates in glial cells, platelets, and serotonergic neurons (Saura et al., 1996). Each isoenzyme is the target of affective and neurogenerative disorders (Cesura & Pletscher, 1992) (Jegham et al., 1998). MAO-A inhibitors are used in the treatment of behavioral disorders, particularly depression (Academy & Sciences, 1991) (D, 1985), while MAO-B inhibitors, such as selegiline and rasagiline are used in the treatment of Parkinson's disease (PD) (Clarke et al., 2003)(Peretz et al., 2016). Previous studies have also shown the role of MAO-B inhibitors in the treatment of Alzheimer's disease (AD) (Marin et al., 1995)(Jo et al., 2014) and other neurodegenerative diseases (Waibel et al., 2004). However, MAO-B drugs are irreversible and cause pharmacological side effects in the long-term treatment of PD (Halgren et al., 2004). Therefore, the development of reversible MAO-B inhibitors, such as the recent antiParkinson drug safinamide (Fabbri et al., 2015), is required.

CONTACT Abdellah El Aissouq abdellah.elaissouq@usmba.ac.ma LPME Laboratory, Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco

Supplemental data for this article can be accessed online at http://dx.doi.org/10.1080/07391102.2022.2071341

Informa UK Limited, trading as Taylor & Francis Group

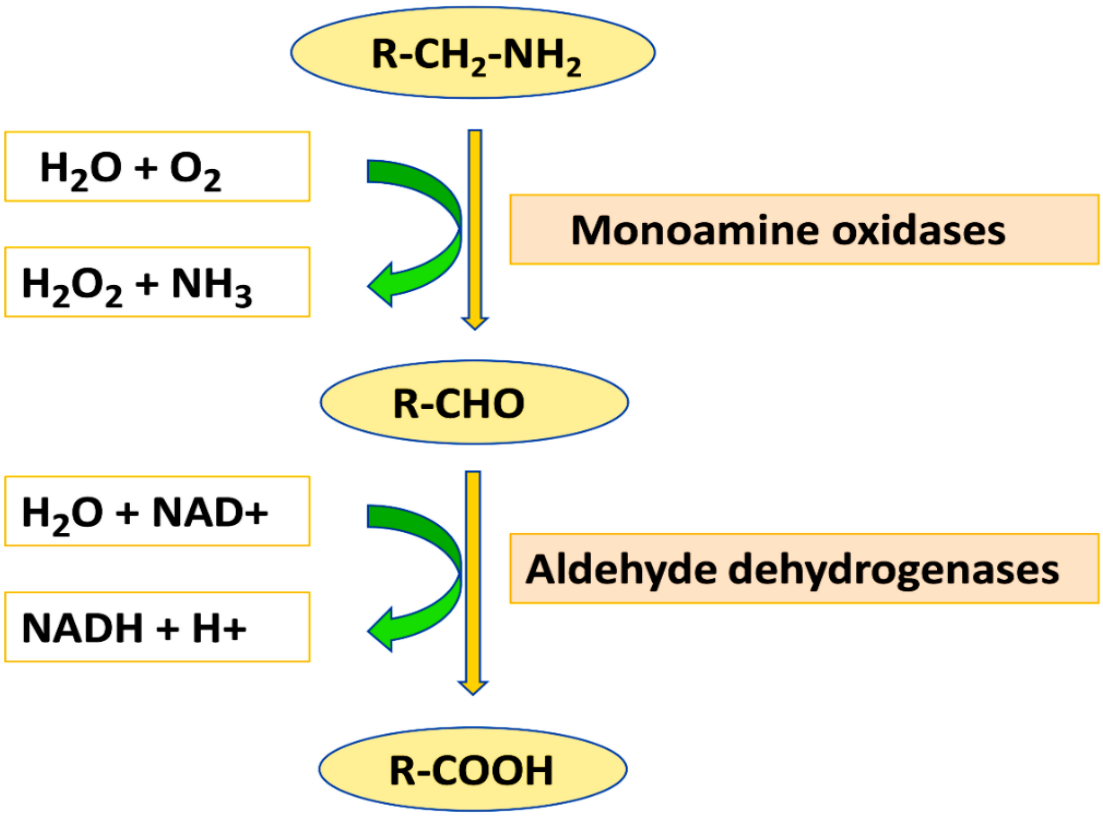


Figure 1. *Mechanism of amine oxidation by monoamine oxidases (MAOs)*

Chalcone (**Figure 2**) and its derivatives (α, β-unsaturated ester derivatives, and α, β-unsaturated amide derivatives) are of great significance due to their important roles in biological systems. Chalcones are a class of polyphenolic derivatives belonging to the flavonoid family, their structure consisting of two aromatic rings joined by a three-carbon, α, β-unsaturated carbonyl system. They exhibit a broad spectrum of pharmacological activities, including anticancer (Salum et al., 2013), anti-inflammatory (Israf et al., 2007), anti-diabetic (Mahapatra et al., 2015), anti-oxidants (Mallikharjuna Rao Lambu, Suresh Kumar, Syed Khalid Yousuf, Deepak K Sharma, Altaf Hussain, Ajay Kumar, Fayaz Malik, 2013) (Ramadan et al., 2014), anti-malarial (Sinha et al., 2013), anti-HIV (Sharma et al., 2011) and anti-MAO-B (Arasappan et al., 2009).

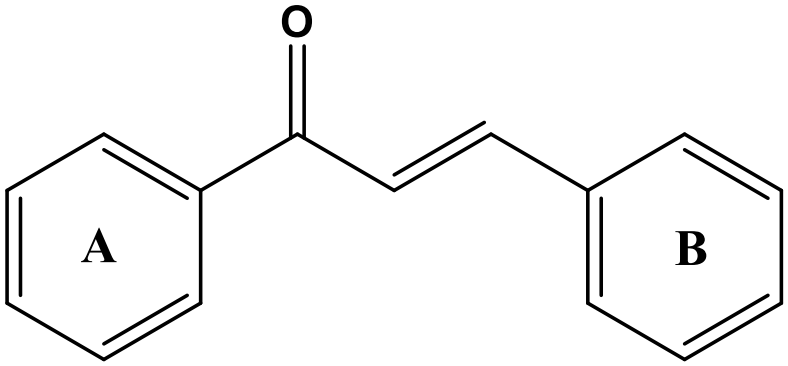


Figure 2*. Chemical structure of chalcone (1,3-diphenyl-2-propan-1-one)*

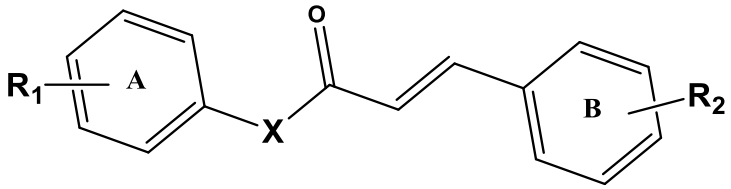
Recently, a series of chalcones and their derivatives were synthesized by Choi and co-workers (Choi et al., 2015). The synthesized compounds are known as monoamine oxidase inhibitors (MAOI) and display a wide range of activity ($0.016 < IC_{50}$ (µM) < 8.39). To understand the relationship between the chemical structures and their corresponding activities, docking-based 3D-QSAR (three-dimensional quantitative structure-activity relationship) was performed. The docking analysis was first applied to generate the most appropriate drug-like conformations and to investigate the binding site interactions between the studied compounds (ligands) and the MAO-B enzyme (receptor) (El Aissouq, Chedadi, Kasmi, et al., 2021) (El Aissouq, Chedadi, Bouachrine, et al., 2021). The 3D-QSAR models were built by using the methods of comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA), based on the most suitable conformations (obtained by docking analysis) (El Aissouq et al., 2020). The CoMFA method relates steric (S) and electrostatic (E) fields with their structural proprieties. As an extension of the CoMFA method, the CoMSIA method connects steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D), and hydrogen bond acceptor (A) fields with the structural pharmacophore of compounds (Chu et al., 2020). The oral bioavailability and toxicity of compounds were also evaluated by using computational ADMET (absorption, distribution, metabolism, excretion, and toxicity). Moreover, the dynamic behavior and stability of compounds were verified by molecular dynamics (MD) simulation (Oussama, 2022). Finally, the binding free energy of complexes was evaluated using the MM_GBSA method. In this study, the combination of ligand-based and structure-based drug design may be helpful to develop and design new unsaturated ketone candidates as reversible MAOI.

II. Materials and Methods

1. Dataset preparation

In this study, a series of unsaturated ketone derivatives which have been synthesized by Choi and co-workers were selected (Choi et al., 2015). The selected compounds are known as selective and reversible monoamine oxidase inhibitors and display a wide range of activity (0.016 µM to 8.39 µM). The IC_{50} values were converted into pIC_{50} values ($pIC_{50} = \log (1/IC_{50})$) and then used as a dependent variable in the 3D-QSAR study. The set of compounds was arbitrarily divided into a training set (about 80% of the dataset) for the generation of 3D-QSAR models and a test set (about 20% of the dataset) for testing the predictive ability of the model's generation. The chemical structures and their corresponding activity are presented in **Table 1**.

Table 1. *Dataset of unsaturated ketone derivatives and their corresponding experimental activity*



No	R1	R2	X	IC_{50} (µM)	pIC_{50}
8a	2-OCH3	H	Bond	0.253	6.5969
9a	3-OCH3	H	Bond	0.793	6.1007
10a	4-OCH3	H	Bond	0.471	6.3270
10b	4-OCH3	2'-CF3	Bond	0.016	7.7959
10c	4-OCH3	3'-CF3	Bond	0.145	6.8386
10d*	4-OCH3	4'-CF3	Bond	0.199	6.7011
10e	4-OCH3	2'-F	Bond	0.056	7.2518
10f	4-OCH3	2'-Cl	Bond	0.069	7.1612
10g	4-OCH3	2'-OCH3	Bond	0.419	6.3778
10h*	4-OCH3	3'-OCH3	Bond	1.396	5.8551
11a*	4-OH	H	Bond	0.182	6.7399
11b*	4-OH	2'-CF3	Bond	0.057	7.2441
11g	4-OH	2'-OCH3	Bond	0.624	6.2048
12a*	H	2'-CF3	N	8.39	5.0762
12b	H	2'-Cl	N	2.82	5.5498
15a	3-OH	2'-CF3	N	5.98	5.2233
15b	3-OH	2'-Cl	N	5.21	5.2832
16a	3-F	2'-CF3	N	3.37	5.4724
16b	3-F	3'-CF3	N	4.89	5.3107
16c	3-F	4'-CF3	N	0.283	6.5482
17a	2-F	4'-CF3	N	0.58	6.2366
18a	4-F	4'-CF3	N	0.223	6.6517
19a	H	2'-CF3	O	5.88	5.2306
20a	4-OCH3	2'-CF3	O	2.59	5.5867
20b	4-OCH3	2'-Cl	O	0.937	6.0283

* Denotes compounds used in the test set

2. **Dataset optimization**

Before modeling, the 3D chemical structures were designed using Marvin Sketch software. Energy minimization and calculations were performed using the SYBYL-X 2.1 program package (Tong et al., 2017). All compounds were energy minimized using the conjugate gradient procedure based on the tripos force field with a convergence criterion value set to 0.01 kcal/mol Å (Clark et al., 1989). The maximum number of iterations was set to 2000. Partial atomic charges were added using Gasteiger-Hücke charges (Caballero et al., 2007).

3. Docking study

Molecular docking analysis was carried out to generate the optimal conformation of each compound (for docking-based alignment) and to identify the binding mode of interactions between the unsaturated ketone derivatives (ligands) and the MAO-B enzyme (receptor). First, the x-ray diffraction structure of the MAO-B enzyme was selected from the RCSB Protein Data Bank (PDB ID: 2BK3) (Binda et al., 2005). Then, the ligand and water molecules were removed from the crystal structure. Polar hydrogen and Gast- Huck charges were added to the protein. The AutoDock tool (Morris et al., 2009) was used to define the docking area by selecting a cubic grid box (X= 40, Y= 40, and Z= 40 at 0.375 Å) centered at the catalytic site of the MAO-B enzyme. Finally, the docking validation was performed by docking the native ligand into the active site of the MAO-B enzyme using AutoDock 4.2 (Trott & Olson, 2009). The best-obtained pose gave the root mean square deviation (RMSD) value of 1.57 Å, using 50 iterations. The superposition between co-crystallized and redocked ligands is shown in **Figure 3**.

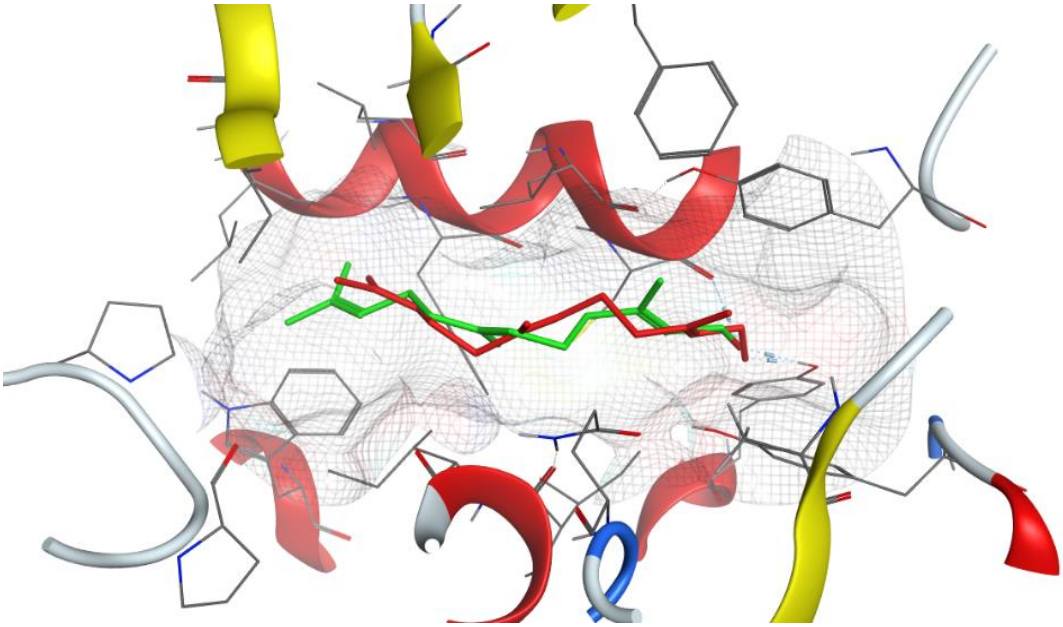


Figure 3. *Superposition of the native ligand (green color) and re-docked ligand (red color) in the binding pocket of MAO-B enzyme*

Following this procedure, all compounds were docked in the binding pocket of the MAO-B enzyme, and the conformations with the lowest energy were utilized to construct the docking-based alignment. The results were analyzed and visualized using PyMol (www.pymol.org) and Discovery Studio 2017 R2 (www.3dsbiovia.com) softwares.

4. 3D-QSAR analysis

Docking-Based alignment

To solve the problems related to molecular conformations and alignment in 3D-QSAR studies, docking-based alignment was performed. Compound 10b was used as a template for aligning the training and test sets. The optimal conformations were aligned to their common structure (core) using the database alignment option in SYBYL software. The alignment of the dataset compounds inside the binding pocket of the MAO-B enzyme is shown in **Figure 4.**

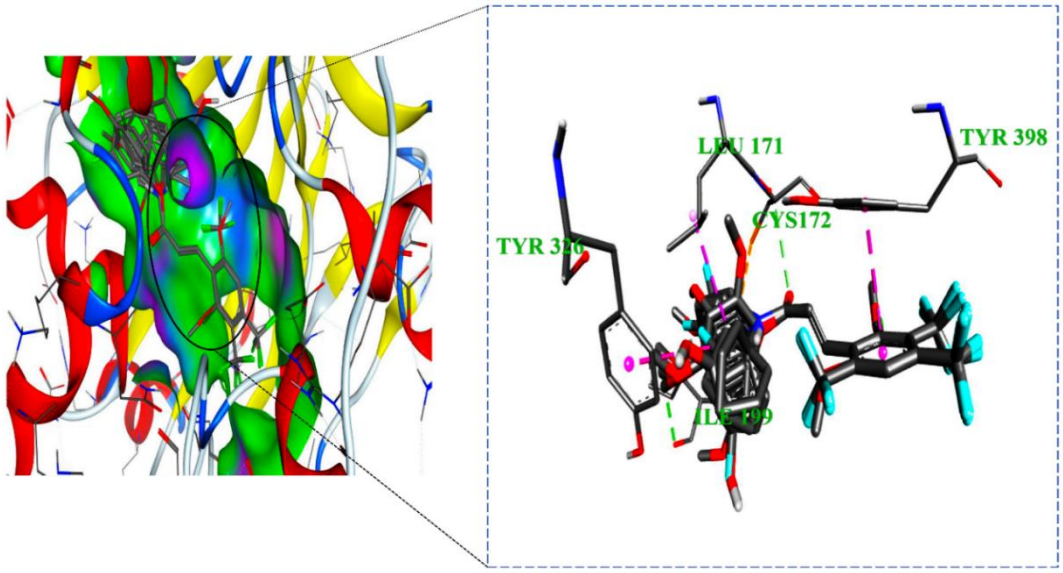


Figure 4. *Alignment of optimal conformations inside binding pocket of MAO-B enzyme*

CoMFA and CoMSIA analysis

To generate a predictive 3D-QSAR model. Comparative molecular field analysis (CoMFA) (Cramer et al., 1988) and comparative molecular similarity index analysis (CoMSIA) (G Klebe et al., 1994) were performed using the SYBYL-X 2.1 package. The CoMFA and CoMSIA descriptors were obtained by a 3D lattice with a grid spacing of 2.0 Å in all cartesian directions. In CoMFA analysis, steric and electrostatic fields were probed on an sp3 hybridized carbon atom with a charge of +1 and a van der Waal's radius of 1.52 Å. The Tripos force field with a distance-dependent dielectric was used to calculate steric and electrostatic interaction. The energy cut-off was set to be 30 kcal/mol by default (Wu et al., 2014). In CoMSIA analysis, steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor were carried out using a probe sp3 carbon atom with a van der Waal's radius of 1 Å, a charge of +1,

hydrophobicity of +1, and hydrogen bond acceptor and donor properties of +1. The attenuation factor was set to the default value of 0.3 (Gerhard Klebe & Abraham, 1999).
After CoMFA and CoMSIA descriptors calculation, partial-least-squares (PLS) (Bush & Nachbar, 1993) analysis was used to generate the linear relationship between the calculated descriptors and the pIC_{50} values of the best-selected conformer (obtained by the docking analysis). The leave-one-out (LOO) procedure was used to generate the cross-validated coefficient (q^2) and the optimum number of compounds (N) (Worachartcheewan et al., 2013). The final models were constructed by a non-cross-validation PLS to yield the conventional determination coefficient (R^2), standard error of estimate (SEE), and F values. The predictive abilities of the generated 3D-QSAR models were verified by the external validation (test set) (Waller et al., 1993) (Stitou et al., 2020).

5. Drug likeness and ADMET prediction

To minimize the number of predicted molecules in 3D-QSAR research, drug-likeness and ADMET prediction (adsorption, distribution, metabolism, excretion, and toxicity) were used (Goudzal et al., 2020)(El Mchichi et al., 2021). All calculations were performed using pkCSM (Pires et al., 2015) and SWISSadmet (Daina et al., 2017) web servers. The drug-likeness was assessed using the Lipinski, Ghose, Veber, Egan, and Muegge criteria, as well as the bioavailability score. The pharmacokinetic characteristics of the predicted compounds were then evaluated via ADMET screening. Compounds having low drug-likeness and weak ADMET characteristics will not be considered in this study.

6. Molecular Dynamics simulations

Molecular dynamics (MD) simulations were performed using the GROMACS 5.1.4 (M. J. Abraham et al., 2015) software package with the CHARMM36 all-atom force field (March, 2019) (Best et al., 2012). The topological parameters of the receptor (2BK3) and ligands were generated using the 'pdb2gmx' script and the CHARMM general force field (CGenFF) server, respectively (K. VANOMMESLAEGHE, E. HATCHER, C. ACHARYA, S. KUNDU, S. ZHONG, J. SHIM, E. DARIAN, O. GUVENCH, P. LOPES, I. VOROBYOV, 2010). The systems (ligands-receptor) were immersed in a triclinic box of TIP3P water molecules with a distance of 1.0 nm (Jorgensen et al., 1983). The solvated systems were neutralized by adding sodium (Na+) or chloride (Cl^-) ions. Then, the systems were subjected to 50,000 steps of energy minimization using the steepest descent algorithm. MD simulations were run with a constant number of particles, constant pressure, and constant temperature criteria (NPT) (Nosé, 1991) using the Berendsen thermostat (Berendsen et al., 1984) with coupled temperature and pressure at 300 K and 1 bar, respectively. Finally, the ligand-receptor complexes were subjected to 20 ns MD simulations at a temperature of 300 k, a pressure of 1 bar, and a time step of 2 fs. 7.

7. MM-GBSA binding free energy

The free energy calculations were performed using the molecular mechanics generalized born surface area (MM-GBSA) method (Šali & Blundell, 1993). The g_mmpbsa tools were designed

to work with GROMACS MD trajectories (Toledo Warshaviak et al., 2014). The binding free energy was calculated using the following equation:

$$\Delta G_{binding} = G_{complex} - (G_{ligand} + G_{receptor}) \quad \text{Eq. 1}$$

$$\Delta G_{binding} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S \quad \text{Eq.2}$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdw} \quad \text{Eq.3}$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SA} \quad \text{Eq.4}$$

Where $\Delta G_{binding}$ signifies the total change in free energy between bound-state ($G_{complex}$) and unbound-state ($G_{ligand} + G_{receptor}$) systems, ΔE_{MM} is the complex's gas-phase interaction energy; which includes the internal energy ($\Delta E_{internal}$), electrostatic energy ($\Delta E_{electrostatic}$), and van der Waals energy (ΔE_{vdW}) terms, ΔG_{solv} denotes the solvation free energy, wich includes the polar ($\Delta G_{PB/GB}$) and non-polar (ΔG_{SA}) terms.

III. Results and discussion

1. Molecular Docking Analysis

Molecular docking analysis was first applied to explore the bioactive conformations adopted within the binding pocket of the MAO-B enzyme. All compounds in the dataset were docked into the binding pocket of MAO-B and their binding affinity was calculated using AutoDock 4.2 software. The bioactive conformation of each compound was selected based on the binding affinity scores (see **Table S1** in the supplementary materials). The correlation between the experimental pIC_{50} values and the binding affinity values of compounds is shown in **Figure 5**.

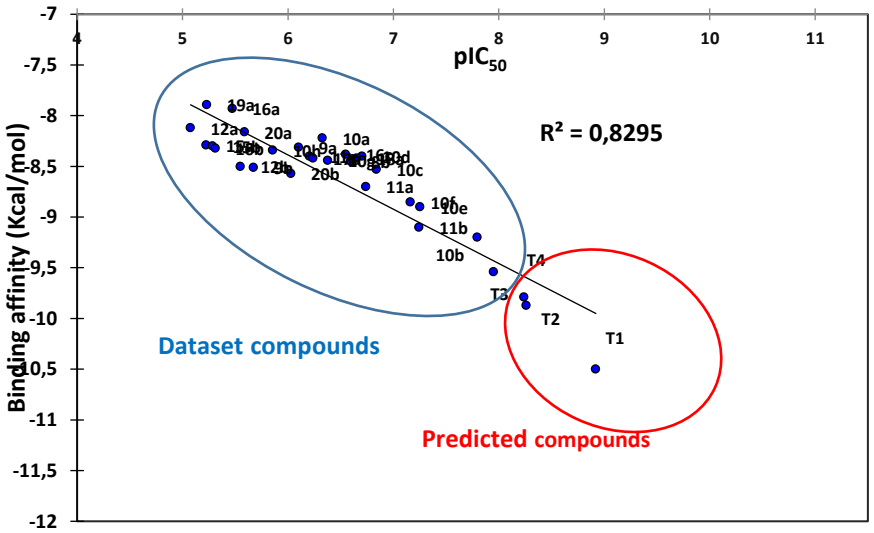


Figure 5. *Correlation between experimental pIC50 values and binding affinity values of compounds. Blue and red colors represent respectively the dataset and predicted compounds*

As described in Figure 5, the high correlation coefficient (R^2 = 0.83) between experimental pIC_{50} values and binding affinities, revealed that compounds with significant inhibitory efficacy also had a high binding affinity. The best poses were considered in the 3D-QSAR studies.

2. 3D-QSAR studies

2..1. CoMFA model analysis

The CoMFA technique was used to investigate the connection between steric and electrostatic fields and the inhibitory action of a variety of unsaturated ketone derivatives. The results are listed in **Table 2**.

Table 2. *Statistical parameters of CoMFA and CoMSIA models using PLS analysis*

Models	r^2_{cv}	r^2_{ncv}	SEE	N	F	r^2_{test}	Fields
CoMFA	0.74	0.98	0.15	8	57.143	0.73	S-E
CoMSIA	**0.74**	**0.99**	**0.12**	**5**	**78.516**	**0.84**	**S-E-A**

r_{cv}^2 = cross- validated correlation coefficient using the leave-one-out method; r_{ncv}^2 = non-cross-validated correlation coefficient; SEE = standard error of estimate; N= optimum number of components; F = Fisher value; r^2_{test} = predicted correlation coefficient; S = steric field; E = electrostatic field; H = hydrophobic field; D = hydrogen bond donor; A = hydrogen bond acceptor.

The CoMFA model yielded a leave-one-out cross-validation coefficient (r^2cv) of 0.74 and an N of 8 optimal components. PLS analysis with a non-cross validation coefficient yielded an r^2_{ncv} value of 0.98, an F value of 57.143 and an SEE value of 0.15. A set of 5 compounds was used to validate the predictability of the generated model, yielding an r^2_{test} value of 0.73, suggesting that the predicted pIC_{50} values are in good agreement with the experimental ones. The plot of predicted versus experimental activities is shown in **Figure 6**.

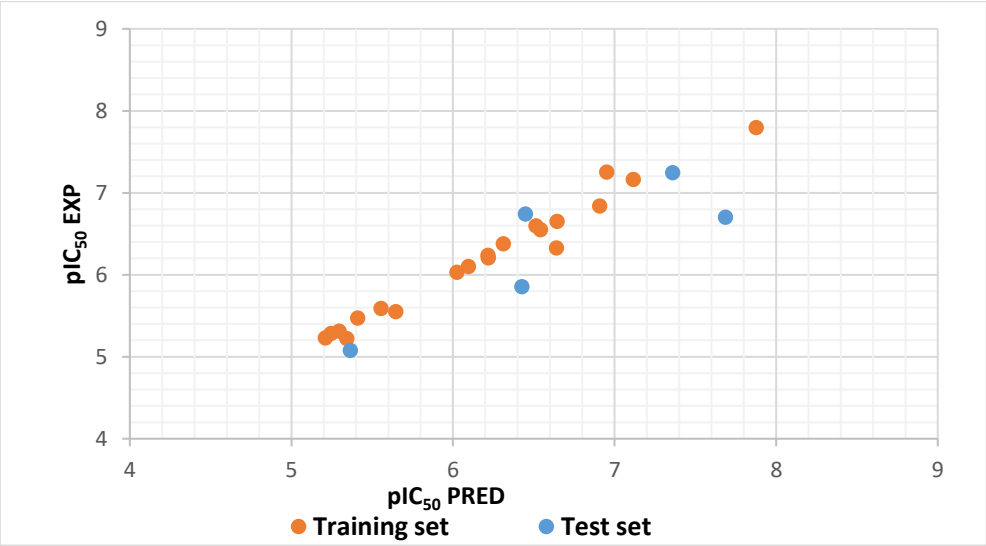


Figure 6. *Observed and predicted pIC_{50} values. The orange and blue dots, respectively, indicate the training and set compounds*

2.2.CoMSIA model analysis

As an extension of CoMFA, CoMSIA analysis was used to explain the effects of steric (S), electrostatic (E), hydrophobic (H), hydrogen bond donor (D), and hydrogen bond acceptor (A) fields on the inhibitory activity of unsaturated ketone derivatives. The developed CoMSIA models were obtained by various combinations of the 5 fields. The best-selected model was obtained by combining the S, E, and A fields. The statistical parameters of the best CoMSIA model are summarized in **Table 2**. The plot of predicted versus experimental activities is shown in **Figure 7**. **Table S2** shows the interaction between all CoMSIA descriptors.

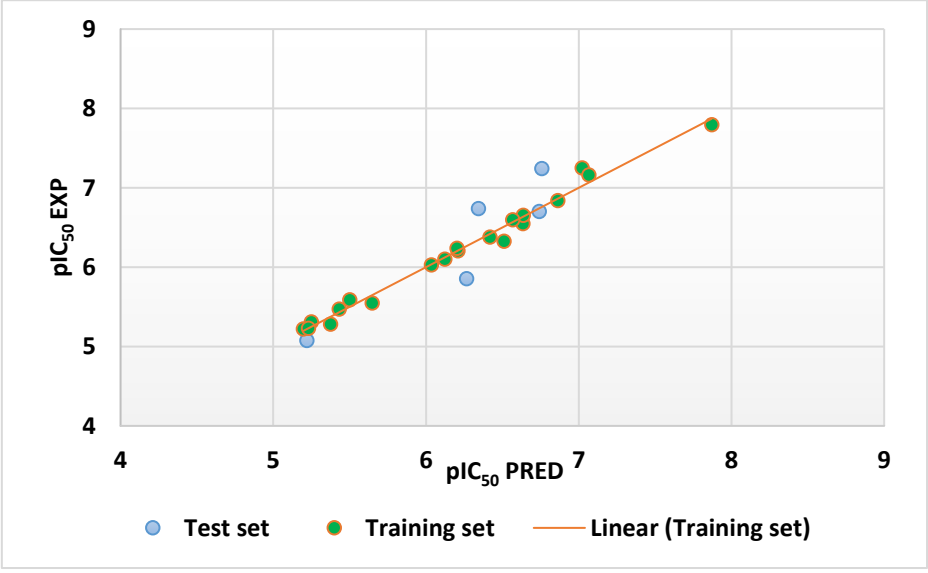


Figure 7. *Plot of observed versus predicted activities. Blue dots represent the training set compounds and green dots represent the test set compounds calculated by the CoMSIA method*

The constructed CoMSIA model revealed an r^2_{cv} value of 0.74 with a number of optimum components (N) of 5. PLS analysis with a non-cross validation coefficient yielded an r^2_{ncv} value of 0.99, an F value of 78.516 and an SEE value of 0.12. The predictability of the developed model was verified by a set of 5 compounds, which gave an r^2_{test} value of 0.84, indicating that predicted pIC_{50} values are in line with experimental ones. **Table 2** shows the statistical analysis of the best CoMSIA model. The observed and predicted activities of compounds, as well as their residuals (the difference between the observed and predicted activities) for the CoMFA and CoMSIA techniques, are listed in **Table 3.**

Table 3. *Observed activity, predicted activity, and residual values of the CoMFA and CoMSIA methods*

No	Sets	pIC_{50} obs	CoMFA		CoMSIA	
			pIC_{50} pred	Residual	pIC_{50} pred	Residual
10a	Training	6.327	6.641	-0.314	6.51	-0.183
10b		7.7959	7.878	-0.0821	7.871	-0.0751
10c		6.8386	6.908	-0.0694	6.863	-0.0244
10e		7.2518	6.951	0.3008	7.021	0.2308
10f		7.1612	7.116	0.0452	7.066	0.0952
10g		6.3778	6.312	0.0658	6.417	-0.0392
11g		6.2048	6.22	-0.0152	6.209	-0.0042
12b		5.5498	5.646	-0.0962	5.646	-0.0962
15a		5.2233	5.344	-0.1207	5.197	0.0263
15b		5.2832	5.247	0.0362	5.375	-0.0918
16a		5.4724	5.41	0.0624	5.43	0.0424
16b		5.3107	5.296	0.0147	5.249	0.0617
16c		6.5482	6.542	0.0062	6.633	-0.0848
17a		6.2366	6.218	0.0186	6.202	0.0346
18a		6.6517	6.644	0.0077	6.635	0.0167
19a		5.2306	5.211	0.0196	5.229	0.0016
20a		5.5867	5.556	0.0307	5.501	0.0857
20b		6.0283	6.024	0.0043	6.035	-0.0067
8b		6.5969	6.515	0.0819	6.566	0.0309
9a		6.1007	6.096	0.0047	6.122	-0.0213
10d	Test	6.7011	7.687	-0.9859	6.74	-0.8579
10h		5.8551	6.428	-0.5728	6.266	-0.7129
11a		6.7399	6.449	0.29093	6.342	1.0779
12a		5.0762	5.366	-0.1160	5.22	-0.1498
11b		7.2440	7.36	-0.2898	6.758	0.7890

2.3.CoMFA and CoMSIA contours map analysis

The CoMFA and CoMSIA contour maps are presented in **Figures 8**, **9**, and **10**. Compound **10b** was chosen as the reference structure because it had the highest inhibitory activity. The contour maps were generated using the StDevTCoefficient field type. The favorable and unfavorable contour maps represent 80% and 20% level contributions, respectively.

Steric contour maps

The steric contour maps of CoMFA and CoMSIA models are presented in **Figures 8a** and **8b** respectively. The green color (steri) represents areas where a bulky group increases activity, whereas the yellow color (sterically-unfavorable) represents areas where a bulky group decreases activity. A large yellow area around the meta-R1 position suggests that bulky groups in these regions are unfavorable for increasing pIC_{50} values. However, the green area around the para-R1 position indicates that bulky substituents are favorable in this position. According to the dataset, compounds with bulky groups at para-R1 (e.g., **10b**,**10e,** and **10f**) exhibit more potency activity, compared to compounds with bulky groups at meta-R1 (e.g., **15a**, **15b**, **16a**, and **16b**).

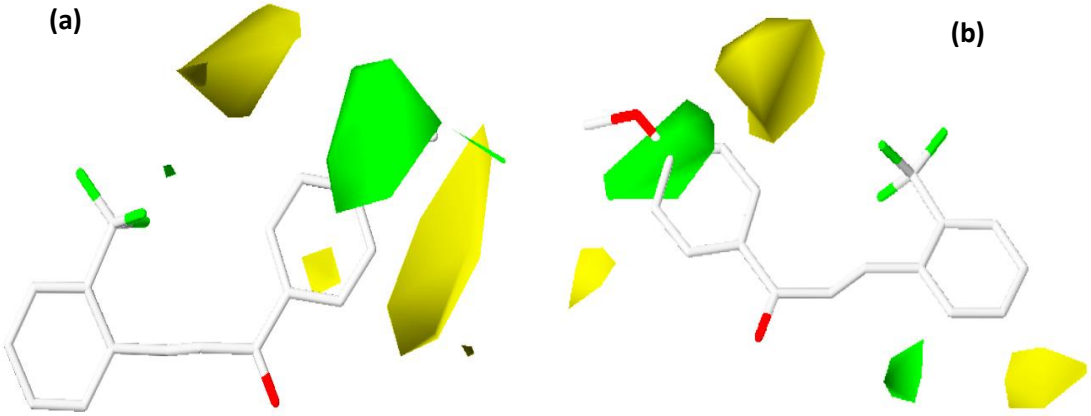


Figure 8. *Steric contour maps of CoMFA* ***(a)*** *and CoMSIA* ***(b)*** *based on compound* ***10b****. The green contours represent regions where bulky groups are favored and yellow contours represent regions where bulky groups are disfavored*

Electrostatic contour maps

The electrostatic contour maps of CoMFA and CoMSIA models are shown in **Figures 9a** and **9b**, respectively. The red color indicates the regions where the addition of electronegative groups increases activity, while the blue color indicates the regions where the addition of electropositive groups increases activity. A red contour is found around the para-R1 position, suggesting that electronegative groups in this position would increase the value of pIC_{50}. However, the large blue areas around the ortho/meta-R1 and the ortho/meta-R2 positions, suggest that electronegative groups in these regions would decrease the value of pIC_{50}. According to the experimental data, compounds with electronegative substitutions at the ortho

and meta positions (e.g., **16a**: meta-R1 = F, ortho-R2 = CF3. **16b**: meta-R1 = F, meta-R2 = CF3), showed low pIC_{50} values (**16a**: pIC_{50} = 5.4725, **16b**: pIC_{50} = 5.3107) compared to compounds with electronegative substitutions at para positions (e.g., **18a**: para-R1 = F and pIC_{50} = 6.6517).

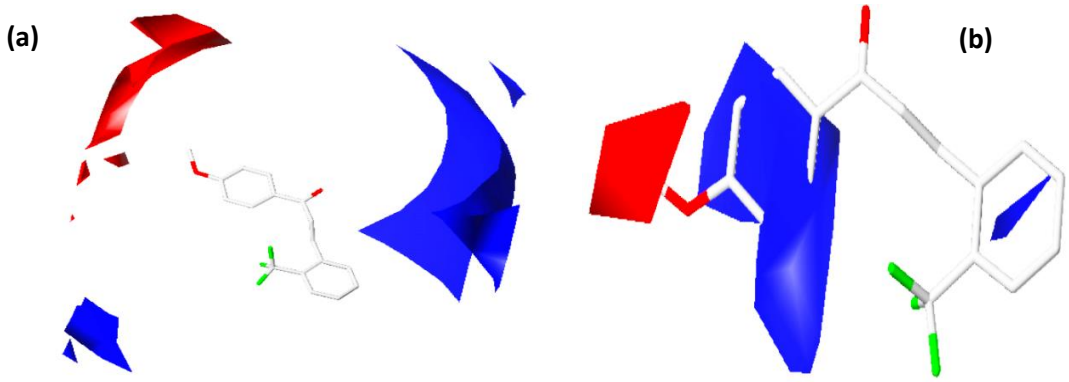


***Figure 9**. Electrostatic contour maps of CoMFA **(a)** and CoMSIA **(b)** based on compound 10b. The red color represents regions where electronegative groups are favored and the blue color represents regions where electronegative groups are disfavored*

H-Bond acceptor contour maps

The H-bond acceptors' contour maps of the CoMSIA model are shown in **Figure 10**. The red color represents the regions where the addition of H-bond acceptor groups decreases activity, while magenta contours indicate regions where the addition of H-bond acceptor groups increases activity. The large magenta areas around the X and meta-R1 positions suggest that H-bond acceptor substitutions are favorable in these regions. According to the experimental data, compounds with acceptors of the H-bond in these regions show moderate values of activity.

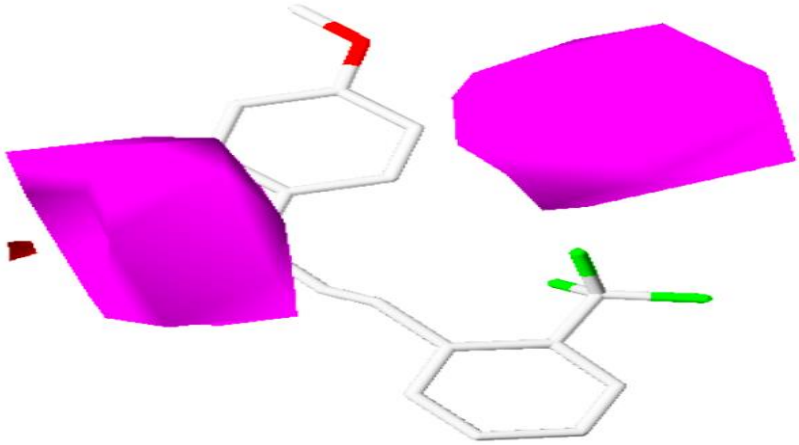


***Figure 10.** H-bond acceptor contour maps of CoMSIA model, based on compound **10b**. The red color represents regions where the addition of H-bond acceptor groups decrease activity and the magenta color represents regions where the addition of H-bond groups increase activity*

3. Design of new compounds

QSAR and molecular docking studies not only provide information regarding structural features responsible for biological activity but they can also be used to design and predict new potent compounds. Based on the information acquired from the best selected 3D-QSAR model (**Figure 11**), we suggested a novel series of unsaturated ketone derivatives by changing some structural characteristics. The chemical structures and their predicted activities as well as the values of binding affinity score are listed in **Table 4**.

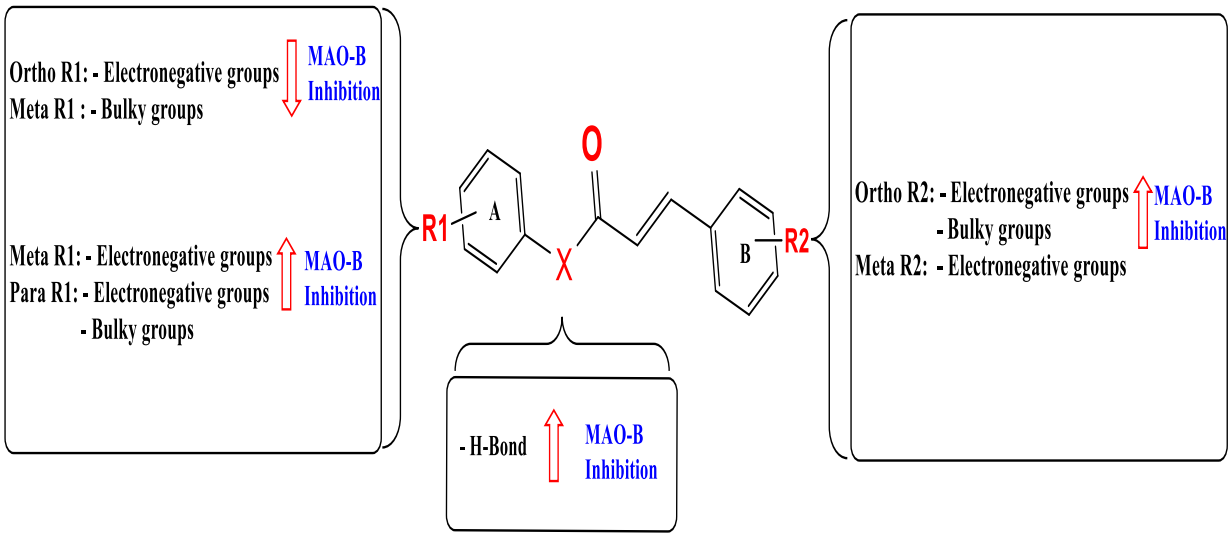
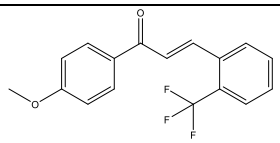
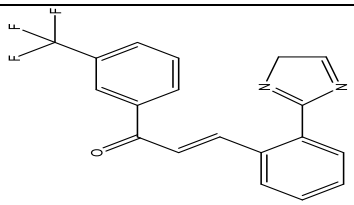
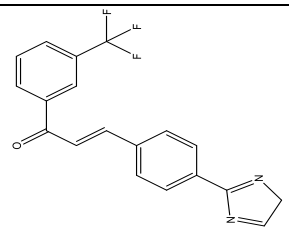
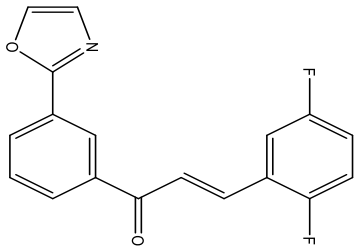
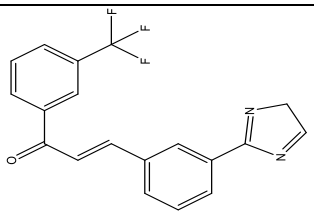


***Figure 11**. Structure-activity relationship (SAR) information are obtained by 3D-QSAR studies*

***Table 4.** Predicted activity and binding affinity score of the newly designed compounds*

No	Chemical structures	CoMSIA	Binding affinity (Kcal/mol)
10b	 (E)-1-(4-methoxyphenyl)-3-(2-(trifluoromethyl) phenyl) prop-2-en-1-one	7.79	-9.2
T1	 (E)-3-(2-(4H-imidazol-2-yl) phenyl)-1-(3-(trifluoromethyl) phenyl) prop-2-en-1-one	8.92	-10.5

T2	 (E)-3-(4-(4H-imidazol-2-yl) phenyl)-1-(3-(trifluoromethyl) phenyl) prop-2-en-1-one	8.26	-9.87
T3	 (E)-3-(2,5-difluorophenyl)-1-(3-(oxazol-2-yl) phenyl) prop-2-en-1-one	8.24	-9.79
T4	 (E)-3-(3-(4H-imidazol-2-yl)phenyl)-1-(3-(trifluoromethyl) phenyl) prop-2-en-1-one	7.95	-9.54

All predicted compounds have pIC_{50} values greater than the most potent inhibitor in the data set (compound **10b**), indicating that our generated model is reliable and can be applied to predict new MAO-B inhibitors. In agreement with the CoMFA and CoMSIA contours map, the introduction of bulky groups at the meta-R1 position (compound **T3**) on ring A and at the ortho-meta-R2 positions on the ring B (compounds **T1** and **T4**) led to excellent inhibitory activities against MAO-B (**T1**, pIC_{50} = 8.92; **T3**, pIC_{50} = 8.24; **T4**, pIC_{50} = 7.95).

In order to understand the binding modes of the newly designed compounds with the MAO-B receptor, molecular docking studies were applied. The compound **T1** was selected and compared with the most active compound in the dataset (compound **10b**). The 2D binding modes of compounds **T1** and **10b** are described in **Figures 12** and **13**, resepectively. The 3D binding modes of compounds **T1** and **10b** are presented in **Figure S1** in supplementary data.

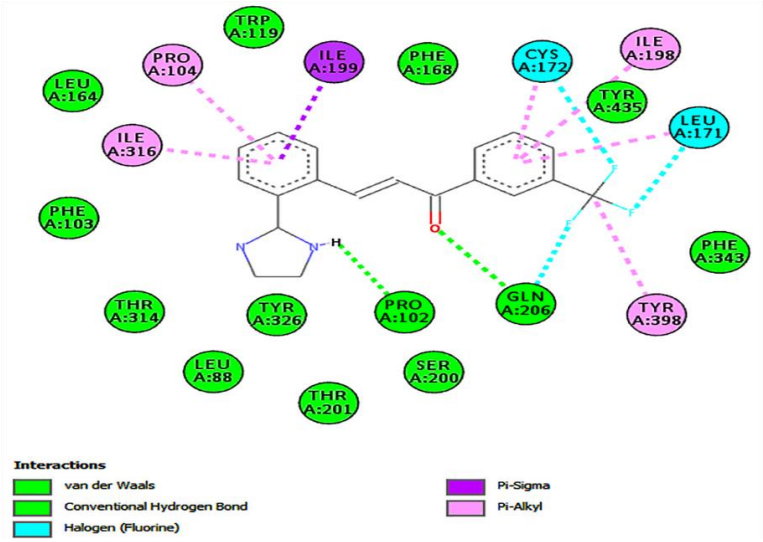


Figure 12. *2D interactions of compound* ***T1*** *with the MAO-B enzyme (PDB ID: 2BK3)*

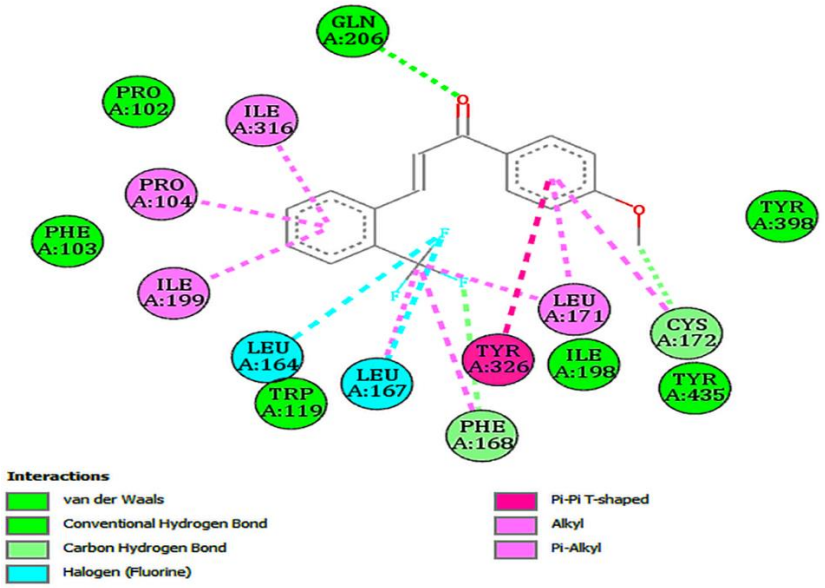


Figure 13. *2D interactions of compound* ***10b*** *with the MAO-B enzyme (PDB ID: 2BK3)*

It can be seen that compound **T1** formed several types of interactions (**Figure 12**) and fitted well in the binding pocket of MAO-B receptor. The most important interactions are the conventional hydrogen bond interactions with GLN 206 and PRO 102, the halogen interactions (Fluorine) with LEU 171, CYS 172, and GLN 206 and the hydrophobic interactions (Pi-sigma and Pi-Alkyl) with PRO 104, LEU 171, CYS 172, ILE 198, ILE 199, ILE 316 and TYR 398. Compound **10b** also forms the same types of interactions (**Figure 13**) compared to **T1** but differs in the number of interactions and distance between the ligand and a particular amino acid. To take into account the number and distance of interactions, a binding affinity score was determined (**Table 4**). All predicted compounds showed low binding affinity values compared to compound **10b**. From this finding we can cobcule that the excellent activity of compound T1 and other predicted compounds might be releated to their electrostatic and hydrogen bond interctions with the main residues of MAO-B binding pocket (GLN 206, PRO 104, LEU 171,

CYS 172). The 2D binding modes of the other predicted compounds are illustrated in **Figures S2, S3,** and **S4** in the supplementary data. This result is in agreement with that obtained by 3D-QSAR studies (compounds with high potency activity showed low binding affinity in the binding pocket of MAO-B receptor).

4. Drug-likeness and molecular property analysis

The drug-likeness and molecular properties analysis of the newly designed compounds as well as the most effective inhibitor in the collection are listed in **Table 5**. All compounds showed a suitable MW value (MW < 500 Daltons) necessary for successful penetration into the central nervous system (CNS). The number of hydrogen bond acceptors and hydrogen bond donors for all predicted compounds were in accordance with Lipinski's rule of five (Lipinski et al., 1997). Also, all the compounds are in accordance with Ghos, Veber, Egan, and Muegge rules. The topological polar surface area (TPSA) and the lipophilicity (Log P <5) of all the compounds were observed to fall within the range. Moreover, the calculated synthetic feasibility suggests that the proposed compounds are easily synthesizable. In summary, it can be concluded that the proposed MAO-B inhibitors presented satisfactory drug-like characteristics.

Table 5. *In silico prediction of molecular properties and synthetic accessibility of proposed compounds*

No	Physicochemical Properties					Druglikeness						Synthetic accessibility
	MW	Log P	HBA	HBD	TPSA	Lipinski	Ghose	Veber	Egan	Muegge	Bioavailability Score	
10b	306.28	4.61	2	0	26.30	Yes, 0 violation	Yes	Yes	Yes	Yes	0.55	2.70
T1	342.32	4.43	3	0	41.79	Yes, 0 violation	Yes	Yes	Yes	Yes	0.55	3.55
T2	342.32	4.23	3	0	41.79	Yes, 0 violation	Yes	Yes	Yes	Yes	0.55	3.56
T3	342.32	4.23	3	0	41.79	Yes, 0 violation	Yes	Yes	Yes	Yes	0.55	3.51
T4	311.28	4.51	3	0	43.10	Yes, 0 violation	Yes	Yes	Yes	Yes	0.55	2.93

MW = molecular weight, LogP = octanol–water partition coefficient, HBA = number of hydrogen bond acceptor, HBD = number of hydrogen bond donor, $N_{violations}$ = violations from Lipinski's rule, TPSA = topological polar surface area.

5. ADMET prediction

The completion of a drug's journey through the body is measured in terms of adsorption, distribution, metabolism, elimination, and toxicity (ADMET). The results include human intestinal absorption (HIA), blood-brain barrier (BBB), Caco-2 permeability, water solubility (log mol/L), CYP450 substrate, and its inhibitor (2D6, 3A4, 1A2, 2C19, and 3A4) and toxicity were summarized in **Table 6**. All compounds showed significant ADMET parameters. The Caco-2 cell permeability in vitro is an important parameter to assess the intestinal absorption of the drug. A compound is considered to have a high Caco-2 permeability if it has a Papp> 8 *10-6 cm/s. For the pkCSM predictive model, high Caco-2 permeability would translate into predicted values > 0.90. Our results indicated that all designed compounds showed a Caco-2 value higher than 0.90. A common measure of the degree of BBB penetration is the ratio of the steady-state concentrations of the drug molecule in the brain and the blood, usually expressed as log (Cbrain/Cblood) or, more simply, log BB. Compounds with log BB > 0.3 cross the BBB readily, while compounds with log BB < -1.0 are only poorly distributed to the brain (M. H. Abraham et al., 1997). The moderate values of log BB (**Table 6**), indicate that these compounds may be effective in treating Parkinson's disease. In addition, the metabolism of the predicted compounds was evaluated by the inhibitory or substrate behavior of the cytochrome P450 enzymes (CYPs). All the predicted compounds were found to be substrates of CYP3A4, whereas no compounds were found to be substrates of CYP2D6. Moreover, the results of inhibition studies indicate that all the compounds might inhibit CYP1A2, CYP2C19, and CYP2C9. There were no compounds found to inhibit CYP2D6 or CYP3A. Also, the Ames toxicity test indicates that the predicted compounds are not dangerous. However, the maximum tolerated dose in humans was 0.24 mg/kg/day for compounds **T1**, **T2**, and **T3** and 0.60 mg/kg/day for compound **T4**.

***Table 6**. In silico ADMET prediction of newly predicted compounds and the most potent inhibitor in the dataset*

Compounds	Absorption			Distribution	Metabolism							Excretion	Toxicity	
					CYP									
	Water solubility	Caco2 permeability	Intestinal absorption human (HIA)	Blood-Brain Barrier Permeability	2D6	3A4	1A2	2C19	2C9	2D6	3A4	Total Clearance	AMES toxicity	Max. tolerated dose (human)
					Substrate			Inhibitor						
	Numeric (Log mol/L)	Numeric (log Papp in 10^{-6} cm/s)	Numeric (% Absorbed)	Numeric (log BB)	Categorical (Yes/No)							Numeric (log ml/min/kg)	Categorical (Yes/No)	Numeric (log mg/kg/day)
10b	-5.678	1.55	93.92	0.27	NO	YES	YES	YES	YES	NO	NO	0.023	NO	0.67
T1	-5.74	1.49	94.46	0.055	NO	YES	YES	YES	YES	NO	NO	0.39	NO	0.24
T2	-5.74	1.49	94.46	0.055	NO	YES	YES	YES	YES	NO	NO	0.38	NO	0.24
T3	-5.74	1.49	94.46	0.055	NO	YES	YES	YES	YES	NO	NO	0.37	NO	0.24
T4	-4.82	1.47	96.95	0.008	NO	YES	YES	YES	YES	NO	NO	0.30	NO	0.60

6. Prediction of activity spectra of substances (PASS)

The biological activity spectrum of the newly designed compounds was predicted using the PASS program. For each compound, the probability to be active 'Pa', to be inactive 'Pi', and the type of activity are listed in **Table S3**. The unsaturated ketone derivatives exhibited a variant of activities reflected by high values of Pa. The probability of compound **10b** being active against Caspase 3, feruloyl esterase and aspulvinone dimethylallyltransferase is 87%, 86.6%, and 83.1%, respectively. Compound **T2** showed a high probability of being active against phobic disorder treatment (pa = 0.838), aspulvinone dimethylallyltransferase (pa = 0.829) and mucomembranous protector (pa = 0.812).

7. Molecular dynamics simulation

In order to validate the pose issued from the docking analysis and to confirm the stability of docked compounds in the binding pocket of the MAO-B enzyme, compounds **b10** and **T1** were subjected to 100 ns of MD simulation. The results are summarized in **Figures 14 and 15.**

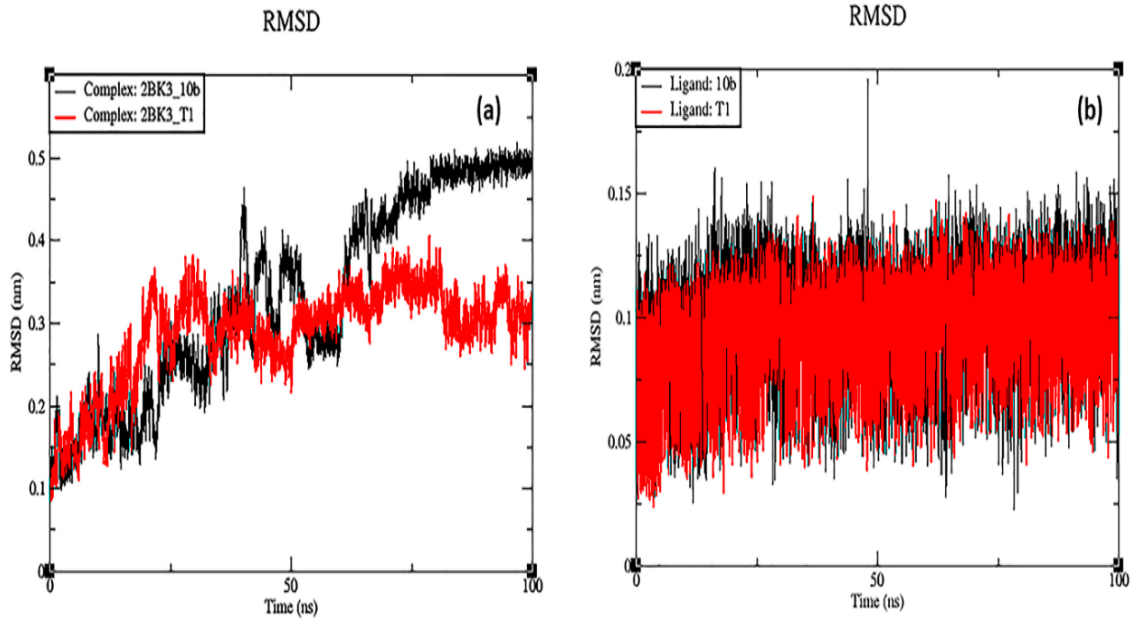


Figure 14. *RMSD plots of 100 ns of MD simulation. (a) for backbone and (b) for ligand. In all figures, the red color represents the 2BK3_T1 complex and the black color represents the 2BK3_b10 complex*

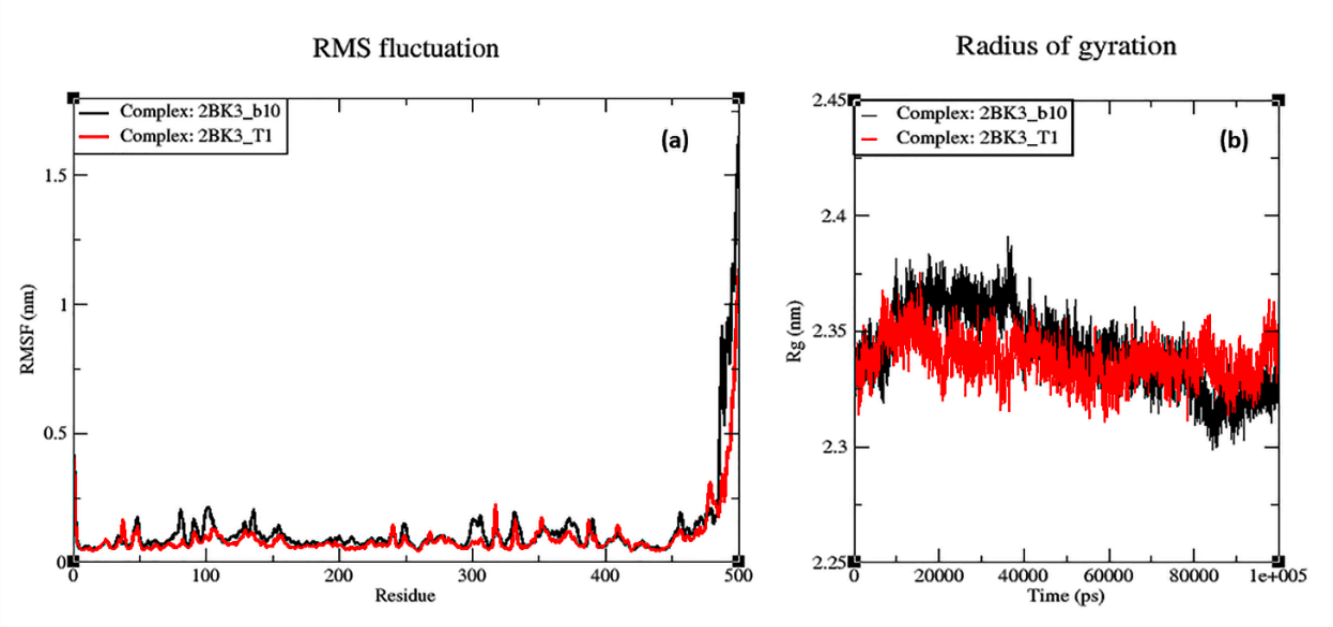


Figure 15*. MD simulation of ligand-protein complex. (a) and (b) represent, respectively, the plots of RMSF (C-alpha) and Rg over the 100 ns of MD silulation at 300 K. In all figures, the red and black colors represent the 2BK3_T1 and 2BK3_b10 complexes*

The root mean square deviations (RMSDs) of the backbone and ligand were calculated against the initial structure of protein-ligand complexes. The results are presented in **Figures 14 (a)** and **14 (b).** As described in **Figure 14 (a)**, the backbone of the 2BK3_T1 complex (red) showed a significant RMSD value compared to the 2BK3_b10 complex, with an average value of 0.24 and 0.32 nm, respectively. Simultaneously, the average RMSD values of T1 and b10 ligands were 0.07 and 0.11 nm, respectively (**Figure 14 (b)**). These results demonstrated that the T1 compound was more stable during the 100 ns of MD simulation compared to compound b10. The movement of amino acid residues during the 100 ns of MD simulation was analyzed using the root mean square fluctuation (RMSF) (**Figure 15 (a))**. This parameter is an indicator of residual flexibility. From **Figure 15 (a),** we can see that the **T1** compound shows fewer conformational changes during the MD simulation compared to the **b10** compound. The stability of systems was also verified by plotting the radius of gyration (Rg) (**Figure 15 (b)**). The results suggest that the 2BK3_T1 complex is stable and compact during the 100 ns of MD simulation.

8. MM-GBSA binding free energy

The binding free energy of the simulated complexes was computed to revalidate the inhibitor affinity predicted by docking studies. The binding free energies were calculated using the last 10 ns of MD trajectories. The total of the nonpolar, polar, and non-bonded interaction energies for all complexes was determined using the g_mmpbsa tools implemented in GROMACS and the results are presented in **Table 7**.

Table 7. *Binding free energy of 2BK3_b10 and 2BK3_T1 complexes using MM-GBSA method*

Complexes	ΔE_{vdw} (KJ/mol)	ΔE_{ele} (KJ/mol)	ΔG_{polar} (KJ/mol)	SASA (KJ/mol)	$\Delta G_{binding}$ (KJ/mol)
2BK3_T1	-146.472	-334.669	87.851	-16.216	-409.506
2BK3_b10	0.00	-0.032	-35.938	4.087	-31.883

As shown in Table 7, the binding free energy of 2BK3_b10 and 2BK3_T1 was -31.883 KJ/mol and -409.506 KJ/mol, respectively. The increased potency of 2BK3_T1 over 2BK3_b10 is attributable to receptor-ligand electrostatic interactions, which account for the majority of the binding affinity contributions (-334.669 KJ/mol). These results showed that T1 compound is more stable than b10 compound in the binding pocket of MAO-B receptor.

IV. Conclusion

In this study, docking based 3D-QSAR, ADMET prediction and molecular dynamics simulations were performed on a series of MAO-B inhibitors. First, molecular docking analysis was applied to select the optimal conformation of each ligand for further 3D-QSAR analysis. Then, docking-based alignment was used to generate CoMFA and CoMSIA models. The best CoMFA and CoMSIA models yielded r^2 values of 0.98 and 0.99, q^2 values of 0.74 and 0.74,

and SEE values of 0.15 and 0.12, respectively. The results revealed the importance of steric fields, electrostatic fields, and H-bond acceptors in controlling the activity of the compounds. Based on the results of the CoMFA and CoMSIA contour maps, a new series of MAO-B inhibitors were designed and their activity was predicted using the most appropriate 3D-QSAR model. All predicted compounds showed good bioavailability based on the drug-likeness and ADMET prediction. Finally, the MD simulation and MM_GBSA methods were used to prove the stability of the predicted compounds against the MAO-B enzyme. The results of this studies demonstrate the relevance of unsaturated ketone derivatives in inhibiting MAO-B enzymatic activity, and they might pave the way for the creation of further inhibitory derivatives as potential MAO-B inhibitors.

Conflict of interest

All authors declare that they have no conflict of interest in this work.

References

Abraham, M. H., Takács-Novák, K., & Mitchell, R. C. (1997). On the partition of ampholytes: Application to blood-brain distribution. *Journal of Pharmaceutical Sciences*, *86*(3), 310–315. https://doi.org/10.1021/js960328j

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindah, E. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001

Academy, S., & Sciences, M. (1991). The discovery of antidepressants: A winding path. *A. Pletscher*, *47*. https://doi.org/https://doi.org/10.1007/BF02041242

Arasappan, A., Padilla, A. I., Jao, E., Bennett, F., Bogen, S. L., Chen, K. X., Pike, R. E., Sannigrahi, M., Soares, J., Venkatraman, S., Vibulbhan, B., Saksena, A. K., Girijavallabhan, V., Tong, X., Cheng, K., & Njoroge, F. G. (2009). Toward Second Generation Hepatitis C Virus NS3 Serine Protease Inhibitors : Discovery of Novel P4 Modified Analogues with Improved Potency and Pharmacokinetic Profile. *J. Med. Chem*, *52*, 2806–2817. https://doi.org/https://doi.org/10.1021/jm801590u

Bach, A. W. J., Lant, N. C., Johnsont, D. L., Abell, C. W., Bembenek, M. E., Kwan, S., Seeburg, P. H., & Shiht, J. C. (1988). cDNA cloning of human liver monoamine oxidase A and B : Molecular basis of differences in enzymatic properties. *Proc. Nat!. Acad. Sci.*, *85*(July), 4934–4938. https://doi.org/https://doi.org/10.1073/pnas.85.13.4934

Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., Dinola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(8), 3684–3690. https://doi.org/10.1063/1.448118

Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & Mackerell, A. D. (2012). Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ 1 and χ 2 Dihedral Angles. *J. Chem. Theory Comput.*, *8*(9), 3257–3273. https://doi.org/https://doi.org/10.1021/ct300400x

Binda, C., Khalil, A., Li, M., Mattevi, A., Castagnoli, N., & Edmondson, D. E. (2005). Demonstration of Isoleucine 199 as a Structural Determinant for the Selective Inhibition of Human Monoamine Oxidase B by Specific Reversible Inhibitors *. *Journal of Biological Chemistry*, *280*(16), 15761–15766. https://doi.org/10.1074/jbc.M500949200

Bush, B. L., & Nachbar, R. B. (1993). Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *Journal of Computer-Aided Molecular Design*, *7*(5), 587–619. https://doi.org/10.1007/BF00124364

Caballero, J., Saavedra, M., Fernández, M., & González-Nilo, F. D. (2007). Quantitative structure-activity relationship of rubiscolin analogues as δ opioid peptides using comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). *Journal of Agricultural and Food Chemistry*, *55*(20), 8101–8104. https://doi.org/10.1021/jf071031h

Cesura, B. A. M., & Pletscher, A. (1992). The new generation of monoamine oxidase inhibitors. *Progress in Drug Research / Fortschritte Der Arzneimittelforschung / Progrès Des Recherches Pharmaceutiques*, *38*, 224–225. https://doi.org/https://doi.org/10.1007/978-3-0348-7141-9_3

Choi, J. W., Jang, B. K., Cho, N., Park, J. H., Yeon, S. K., Ju, J., Lee, Y. S., Han, G., Pae, A. N., Kim, D. J., & Park, K. D. (2015). Synthesis of a Series of Unsaturated Ketone Derivatives as Selective and Reversible Monoamine Oxidase Inhibitors. *BIOORGANIC & MEDICINAL CHEMISTRY*, *23*(19), 6486–6496. https://doi.org/10.1016/j.bmc.2015.08.012

Chu, H., He, Q. xiu, Wang, J. wei, Deng, Y. ting, Wang, J., Hu, Y., Wang, Y. qiang, & Lin, Z. hua. (2020). 3D-QSAR, molecular docking, and molecular dynamics simulation of a novel thieno[3,4-d]pyrimidine inhibitor targeting human immunodeficiency virus type 1 reverse transcriptase. *Journal of Biomolecular Structure and Dynamics*, *38*(15), 4567–4578. https://doi.org/10.1080/07391102.2019.1697366

Clark, M., Cramer, R. D., & Van Opdenbosch, N. (1989). Validation of the general purpose tripos 5.2 force field. *Journal of Computational Chemistry*, *10*(8), 982–1012. https://doi.org/10.1002/jcc.540100804

Clarke, A., Brewer, F., Johnson, E. S., Mallard, N., Hartig, F., Taylor, S., & Corn, T. H. (2003). A new formulation of selegiline: Improved bioavailability and selectivity for MAO-B inhibition. *Journal of Neural Transmission*, *110*(11), 1241–1255. https://doi.org/10.1007/s00702-003-0036-4

Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *Journal of the American Chemical Society*, *110*(18), 5959–5967. https://doi.org/10.1021/ja00226a005

D, S. Z. M. (1985). A clinical overview of monoamine oxidase inhibitors. *Psychosomatics*, *26*(3), 240-246,251. https://doi.org/10.1016/S0033-3182(85)72877-0

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, *7*(October 2016), 1–13. https://doi.org/10.1038/srep42717

El Aissouq, A., Chedadi, O., Bouachrine, M., & Ouammou, A. (2021). Identification of Novel SARS-CoV-2 Inhibitors: A Structure-Based Virtual Screening Approach. *Journal of Chemistry*, *2021*. https://doi.org/10.1155/2021/1901484

El Aissouq, A., Chedadi, O., Kasmi, R., Elmchichi, L., En-nahli, F., Goudzal, A., Bouachrine, M., Ouammou, A., & Khalil, F. (2021). Molecular Modeling Studies of C-Glycosylfavone Derivatives as GSK-3β Inhibitors Based on QSAR and Docking Analysis. *Journal of Solution Chemistry*, *50*(5), 808–822. https://doi.org/10.1007/s10953-021-01083-6

El Aissouq, A., Toufik, H., Stitou, M., Ouammou, A., & Lamchouri, F. (2020). In Silico Design of Novel Tetra-Substituted Pyridinylimidazoles Derivatives as c-Jun N-Terminal Kinase-3 Inhibitors, Using 2D/3D-QSAR Studies, Molecular Docking and ADMET Prediction. *International Journal of Peptide Research and Therapeutics*, *26*, 1335–1351. https://doi.org/10.1007/s10989-019-09939-8

El Mchichi, L., El Aissouq, A., Kasmi, R., Belhassan, A., El-Mernissi, R., Ouammou, A., Lakhlifi, T., & Bouachrine, M. (2021). In silico design of novel Pyrazole derivatives containing thiourea skeleton as anti-cancer agents using: 3D QSAR, Drug-Likeness studies, ADMET prediction and molecular docking. *Materials Today: Proceedings*, *xxxx*. https://doi.org/10.1016/j.matpr.2021.03.152

Fabbri, M., Rosa, M. M., Abreu, D., & Ferreira, J. J. (2015). Clinical pharmacology review of safinamide for the treatment of Parkinson ' s disease. *Future Medicine*, *5*. https://doi.org/https://doi.org/10.2217/nmt.15.46

Goudzal, A., El Aissouq, A., El Hamdani, H., & Ouammou, A. (2020). QSAR modeling, molecular docking studies and ADMET prediction on a series of phenylaminopyrimidine-(thio) urea derivatives as CK2 inhibitors. *Materials Today: Proceedings*, *xxxx*. https://doi.org/10.1016/j.matpr.2020.08.044

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide : A New Approach for Rapid , Accurate Docking and Scoring . 2 . Enrichment Factors in Database Screening. *J. Med. Chem*, *47*(7), 1750–1759. https://doi.org/https://doi.org/10.1021/jm0310885

Israf, D. A., Khaizurin, T. A., Syahida, A., Lajis, N. H., & Khozirah, S. (2007). Cardamonin inhibits COX and iNOS expression via inhibition of p65NF- □ B nuclear translocation and I □ -B phosphorylation in RAW 264 . 7 macrophage cells. *Molecular Immunology*, *44*, 673–679. https://doi.org/10.1016/j.molimm.2006.04.025

Jegham, S., George, P., Recherche, S., Snc, D. D. R., & Carrières, R. (1998). Monoamine oxidase A and B inhibitors. *Expert Opinion on Therapeutic Patents Jegham*, *8*(9), 1143–1150. https://doi.org/https://doi.org/10.1517/13543776.8.9.1143

Jo, S., Yarishkin, O., Hwang, Y. J., Chun, Y. E., Park, M., Woo, D. H., Bae, J. Y., Kim, T., Lee, J., Chun, H., Park, H. J., Lee, D. Y., Hong, J., Kim, H. Y., Oh, S., Park, S. J., Lee, H., Yoon, B., Kim, Y., … Lee, C. J. (2014). ARTICLES GABA from reactive astrocytes impairs memory in mouse models of Alzheimer ' s disease. *Nature Publishing Group*, *20*, 886–896. https://doi.org/10.1038/nm.3639

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L., Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, *926*(79). https://doi.org/10.1063/1.445869

K. VANOMMESLAEGHE, E. HATCHER, C. ACHARYA, S. KUNDU, S. ZHONG, J. SHIM, E. DARIAN, O. GUVENCH, P. LOPES, I. VOROBYOV, A. D. M. J. (2010). CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem*, *31*(4), 671–690. https://doi.org/https://doi.org/10.1002/jcc.21367

Klebe, G, Abraham, U., & Mietzner, T. (1994). {M}olecular {S}imilarity {I}ndices in a {C}omparative {A}nalysis ({CoMSIA}) of {D}rug {M}olecules to {C}orrelate and {P}redict their {B}iological {A}ctivity. *J. Med. Chem.*, *37*, 4130–4146.

Klebe, Gerhard, & Abraham, U. (1999). Comparative Molecular Similarity Index Analysis (CoMSIA) to study hydrogen-bonding properties and to score combinatorial libraries. *Journal of Computer-Aided Molecular Design*, *13*(1), 1–10. https://doi.org/10.1023/A:1008047919606

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. *23*, 3–25.

Mahapatra, D. K., Asati, V., & Bharti, S. K. (2015). European Journal of Medicinal Chemistry Chalcones and their therapeutic targets for the management of diabetes : Structural and pharmacological perspectives. *European Journal of Medicinal Chemistry*, *92*, 839–865. https://doi.org/10.1016/j.ejmech.2015.01.051

Mallikharjuna Rao Lambu, Suresh Kumar, Syed Khalid Yousuf, Deepak K Sharma, Altaf hussain, Ajay Kumar, Fayaz Malik, and D. M. (2013). Medicinal chemistry of dihydropyran based medium ring macrolides related to aspergillides : selective inhibition of PI3K α ! *J. Med. Chem*, *56*(15), 6136–6145. https://doi.org/10.1021/jm400515c

Marin, D. B., Bierera, L. M., Lawlorb, B. A., Ryana, T. M., Jacobsonc, R., Schmeidlerd, J., Mohsa, R. C., & Davisd, K. L. (1995). L-Deprenyl and physostigmine for the treatment of Alzheimer ' s disease. *Psychiatry Research*, *58*(3). https://doi.org/https://doi.org/10.1016/0165-1781(95)02714-8

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Software News and Updates AutoDock4 and AutoDockTools4 : Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, *30*(16), 2785–2791. https://doi.org/10.1002/jcc

Nosé, S. (1991). Molecular Dynamics Simulations at Constant Temperature and Pressure. *Computer Simulation in Materials Science*, *2384*(1980), 21–41. https://doi.org/10.1007/978-94-011-3546-7_2

Oussama, C. (2022). In silico Prediction of Novel SARS-CoV 3CL pro Inhibitors : a Combination of 3D-QSAR , Molecular Docking , ADMET Prediction , and Molecular Dynamics Simulation. *Biointerface Research in Applied Chemistry*, *12*(4), 5100–5115.

Peretz, C., Segev, H., Rozani, V., Gurevich, T., El-Ad, B., Tsamir, J., & Giladi, N. (2016). Comparison of selegiline and rasagiline therapies in Parkinson disease: A real-life study. *Clinical Neuropharmacology*, *39*(5), 227–231. https://doi.org/10.1097/WNF.0000000000000167

Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, *58*(9), 4066–4072. https://doi.org/10.1021/acs.jmedchem.5b00104

Ramadan, M., Sayed, E., Hamadah, H., El, A., & Fodah, R. (2014). European Journal of Medicinal Chemistry Antiobesity , antioxidant and cytotoxicity activities of newly synthesized chalcone derivatives and their metal complexes. *European Journal of Medicinal Chemistry*, *76*, 517–530. https://doi.org/10.1016/j.ejmech.2014.02.021

Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. In *Journal of Molecular Biology* (Vol. 234, Issue 3, pp. 779–815). https://doi.org/10.1006/jmbi.1993.1626

Salum, L. B., Altei, W. F., Chiaradia, L. D., Cordeiro, M. N. S., Canevarolo, R. R., Melo, C. P. S., Winter, E., Mattei, B., Daghestani, H. N., Santos-silva, M. C., Creczynski-pasa, T. B., Yunes, R. A., Yunes, J. A., Andricopulo, A. D., Day, B. W., Nunes, R. J., & Vogt, A. (2013). European Journal of Medicinal Chemistry migration inhibitors. *European Journal of Medicinal Chemistry*, *63*, 501–510. https://doi.org/10.1016/j.ejmech.2013.02.037

Saura, J., Bleuel, Z., Ulrich, J., Mendelowitsch, A., Chen, K., Shih, J. C., Malherbe, P., Da Prada, M., & Richards, J. G. (1996). Molecular neuroanatomy of human monoamine oxidases A and B revealed by quantitative enzyme radioautography and in situ hybridization histochemistry. *Neuroscience*, *70*(3), 755–774. https://doi.org/10.1016/S0306-4522(96)83013-2

Scrutton, N. S. (2004). Chemical aspects of amine oxidation by flavoprotein enzymes. *Natural Product Reports*, *21*(6), 722–730. https://doi.org/10.1039/b306788m

Sharma, H., Patil, S., Sanchez, T. W., Neamati, N., Schinazi, R. F., & Buolamwini, J. K. (2011). Bioorganic & Medicinal Chemistry Synthesis , biological evaluation and 3D-QSAR studies of 3-keto salicylic acid chalcones and related amides as novel HIV-1 integrase inhibitors. *Bioorganic & Medicinal Chemistry*, *19*(6), 2030–2045. https://doi.org/10.1016/j.bmc.2011.01.047

Shih, J. C., Chen, K., & Ridd, M. J. (1999). MONOAMINE OXIDASE : From Genes to Behavior. *Annual Review of Neuroscience*, *22*, 197–217. https://doi.org/https://doi.org/10.1146/annurev.neuro.22.1.197

Sinha, S., Medhi, B., & Sehgal, R. (2013). Chalcones as an Emerging Lead Molecule for Antimalarial Therapy : A Review. *Journal of Modern Medicinal Chemistry*, *1*, 64–77.

Stitou, M., Toufik, H., Bouachrine, M., & Lamchouri, F. (2020). Quantitative structure–activity relationships analysis, homology modeling, docking and molecular dynamics studies of triterpenoid saponins as Kirsten rat sarcoma inhibitors. *Journal of Biomolecular Structure and Dynamics*, *1102*(39), 1. https://doi.org/10.1080/07391102.2019.1707122

Tipton, K. F. (2018). 90 Years of Monoamine Oxidase: Some Progress and Some Confusion. In *Journal of Neural Transmission* (Vol. 125, Issue 11). Springer Vienna. https://doi.org/10.1007/s00702-018-1881-5

Toledo Warshaviak, D., Golan, G., Borrelli, K. W., Zhu, K., & Kalid, O. (2014). Structure-based virtual screening approach for discovery of covalently bound ligands. *Journal of Chemical Information and Modeling*, *54*(7), 1941–1950. https://doi.org/10.1021/ci500175r

Tong, J., Zhan, P., Wang, X. S., & Wu, Y. (2017). Quionolone carboxylic acid derivatives as HIV-1 integrase inhibitors: Docking-based HQSAR and topomer CoMFA analyses. *Journal of Chemometrics*, *31*(12), 1–13. https://doi.org/10.1002/cem.2934

Trott, O., & Olson, A. J. (2009). Software News and Update AutoDock Vina : Improving the Speed and Accuracy of Docking with a New Scoring Function , Efficient Optimization , and Multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc

Waibel, S., Reuter, A., & Ludolph, A. C. (2004). Rasagiline alone and in combination with riluzole prolongs survival in an ALS mouse model. *Journal of Neurology*, *251*, 1080–1084. https://doi.org/10.1007/s00415-004-0481-5

Waller, C. L., Oprea, T. I., Giolitti, A., & Marshall, G. R. (1993). Three-Dimensional QSAR of Human Immunodeficiency Virus (I) Protease Inhibitors. 1. A CoMFA Study Employing Experimentally-Determined Alignment Rules. *Journal of Medicinal Chemistry*, *36*(26), 4152–4160. https://doi.org/10.1021/jm00078a003

Worachartcheewan, A., Nantasenamat, C., Isarankura-Na-Ayudhya, C., & Prachayasittikul, V. (2013). Quantitative population-health relationship (QPHR) for assessing metabolic syndrome. *EXCLI Journal*, *12*(11), 569–583. https://doi.org/10.2478/s11696-013-0398-5

Wu, S., Qi, W., Su, R., Li, T., Lu, D., & He, Z. (2014). CoMFA and CoMSIA analysis of ACE-inhibitory, antimicrobial and bitter-tasting peptides. *European Journal of Medicinal Chemistry*, *84*, 100–106. https://doi.org/10.1016/j.ejmech.2014.07.015

www.3dsbiovia.com. (n.d.). *BIOVIA - Scientific Enterprise Software for Chemical Research, Material Science R&D*. https://www.3dsbiovia.com/

www.pymol.org. (n.d.). *PyMOL | pymol.org*. Retrieved January 3, 2020, from https://pymol.org/2/

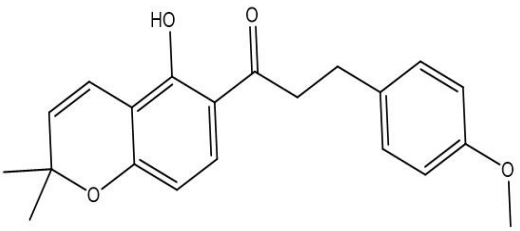
Chapitre III : Identification des agents anti-Parkinsoniens ciblant la protéine MAO-B en utilisant la modélisation par homologie, le docking moléculaire, la simulation par dynamique moléculaire (MD) et les approches ADMET

Résumé :

La monoamine oxydase-B (MAO-B) est une cible médicamenteuse potentielle pour la maladie d'Alzheimer et la maladie de Parkinson. L'objectif de l'article 3 consiste à trouver des inhibiteurs de la MAO-B, on se basant sur des composés à base de produits naturels. Pour atteindre cet objectif, une base de données de 142 molécules a été selectionnée pour réaliser le criblage virtuel par les méthodes de la modélisation moléculaire telles que, la modélisation par homologie, le docking moléculaire, la dynamique moléculaire et l'analyse ADMET. Les molécules sélectionnées sont des métabolites secondaires, dérivées de la chalcone.

Premièrement, les composés collectés ont été dessinés et optimisés par le champ de force MM2 en utilisant le logiciel ChemBioOffice. Ensuite, les structures 3D ont été générées par la méthode DFT en utilisant Becke-three-Lee-Yang-Parr (B3LYP) avec l'ensemble de base 6-31G (d, p) aidé par le logiciel Gaussian 09. La méthode de la modélisation par homologie a été utilisée pour construire la structure 3D du récepteur monoamine oxydase B humaine (hMAO-B). Après l'optimisation de cette structure par le progiciel GROMACS 5.1.4, un criblage virtuel a été réalisé sur l'ensemble des molécules. Les calculs ont été effectués à l'aide du logiciel PyRx en utilisant l'option Vina. Après élimination des molécules ayant une faible affinité, le reste a été soumis à l'analyse ADMET pour évaluer les propriétés pharmacocinétiques de ces composés. Les meilleurs inhibiteurs de la MAO-B ont été validés par la méthode de docking moléculaire en utilisant le programme AutoDock 4.2.

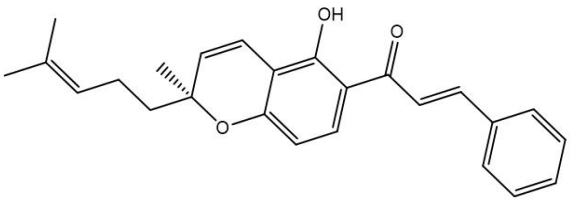
À partir de l'ensemble des 142 molécules, trois molécules (P34, P60 et P81) (Figure a1) ont été considérées comme inhibiteurs potentiels de la MAO-B par rapport aux médicaments de référence (sélégiline, rasagiline et safinamid). Le composé P34 est une crotaramine extraite de la plante « Crotolaria rosmosissima », le composé P60 est un Spinochalcone B extrait de la plante « Tephrosia spinosa » et le composé P81 est un Artoindonesianin J extrait de la plante « Artocarpus bracteata ». L'étude de docking moléculaire a révélé que ces composés étaient bien orientés dans le site actif de la MAO-B. La faible affinité de liaison des composés P60 (-10,77 Kcal/mol), P34 (-10.0 kcal/mol) et P81 (-10.64 kcal/mol) est due à la formation des liaisons hydrogènes, hydrophobiques (Alkyl et Pi-Alkyl) et électrostatiques (Pi-Sigma) avec les principaux résidus de la cavité du substrat. Enfin, leur stabilité a été déduite par la méthode de dynamique moléculaire en utilisant le logiciel GROMACS 5.1.4.



P34



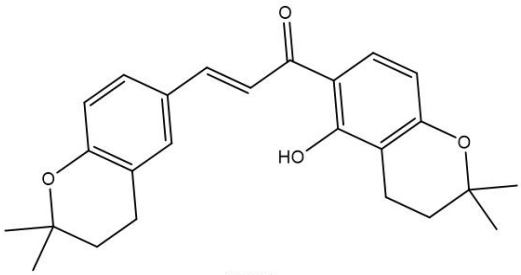
Crotolaria rosmosissima



P60



Tephrosia spinosa



P81



Artocarpus bracteata

Figure a1. *Potentiel inhibiteurs de la MAO-B : P34 (Crotolaria rosmosissima), P60 (Tephrosia spinosa) et P81(Artocarpus bracteata)*

.

Contents lists available ScienceDirec
Neuroscience Letters
journal homepage: www.elsevier.com/locate/neulet

Homology modeling, virtual screening, molecular docking, molecular dynamic (MD) simulation, and ADMET approaches for identification of natural anti-Parkinson agents targeting MAO-B protein

Abdellah El Aissouq [a,*], Mohammed Bouachrine [c], Abdelkrim Ouammou [b], Fouad Khalil [a]

[a] Laboratory of Processes, Materials, and Environment (LPME), Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[b] LIMOME Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco
[c] MCNS Laboratory, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco

ARTICLEINFO

Keywords:
hMAO-B

Neurodegenerative disorders
Natural products
Structure-based virtual screening
ADMET analysis

ABSTRACT

Monoamine oxidase-B (MAO-B) is a flavin-dependent enzyme involved in various neurodegenerative disorders. Here, a dataset of 142 chalcone derivatives, collected from various natural plants, was screened by combining structure-based virtual screening and ADMET approaches. The goal is to discover novel natural chalcones as potential MAO-B inhibitors. With the help of the Gaussian 09.5 software, the 3D chemical structures of compounds were optimized using the DFT method. The 3D structure of the hMAO-B enzyme was built using the Modeller software. The optimized structures were subjected to virtual screening by Autodock Vina, implicated in PyRx software. Among the 142 natural substances, 43 were selected based on their binding affinity. Then, the pharmacokinetic proprieties and toxicity of these compounds were evaluated using ADMET analysis. Ten compounds were predicted to have ADMET characteristics with no side effects. The binding modes and interactions of the top selected compounds were then evaluated using AutoDock 4.2. Compounds P60 and P81 were found to be potential inhibitors of MAO-B compared to rasagiline, safinamid, and selegiline, the reference drugs. The stability of the selected compounds was confirmed by MD simulation. Based on this finding, compounds P60 and P81 could be considered potential hMAO-B inhibitors.

1. Introduction

Monoamine oxidases (MAOs) are mitochondrial-associated enzymes that catalyze the oxidative deamination of monoamine neurotransmitters. MAOs have been extensively studied and are composed of two isoforms, MAO-A and MAO-B, which are both linked with the mitochondrial outer membrane and display different substrate specificities, inhibitor sensitivities, and tissue localizations. MAO-B predominates in the human brain, and its activity is frequently increased in individuals with neurodegenerative disorders including Alzheimer's and Parkinson's [1, 2]. Currently, three MAO-B drugs are frequently used in the treatment of Parkinson's disease: selegiline **(1)**, rasagiline **(2)**, and safinamid **(3)** (Fig. 1). Selegiline and rasagiline are selective and irreversible MAO-B inhibitors, whereas, safinamid is a compound with a more complex mechanism of action associated with dopaminergic and non-dopaminergic properties [3]. To date, no natural substance has developed to be a potential drug against MAO-B. Thus, development of new natural substances to inhibit the enzymatic activity of MAO-B is required.

Chalcones are a large class of organic compounds that may be obtained from natural sources or by synthetic procedures [4]. Due to their basic structure **(4)** and a broad spectrum of pharmacological activities, such as anti-inflammatory [5], antifungal [6, 7], anti-oxidants [8], anti- cancer [9, 10], anti-Alzheimer [11, 5], these compounds have received great attention. Among the various pharmacologically intriguing applications of chalcones, several research groups have lately evaluated their potential as MAO-B inhibitors [12, 13].

The use of computational methods in medicinal chemistry has become a

Abbreviations: MAO-B, Monoamine oxidase-B; DFT, Density Functional Theory; B3LYP, Becke-three-Lee-Yang-Parr; CADD, Computer-aided drug design; ADMET, Absorption, distribution, metabolism, excretion, and toxicity; HIA, Human intestinal absorption; BBB, Blood brain barrier; PDB, Protein data bank; MD, Molecular dynamic; RMSD, Root mean square deviation; RMSF, Root mean square deviation fluctuation; Rg, Radius of gyration.

* Corresponding author.
E-mail address: abdellah.elaissouq@usmba.ac.ma (A. El Aissouq). https://doi.org/10.1016/j.neulet.2022.136803

Received 16 May 2022; Received in revised form 9 July 2022; Accepted 11 July 2022 Available online 14 July 2022

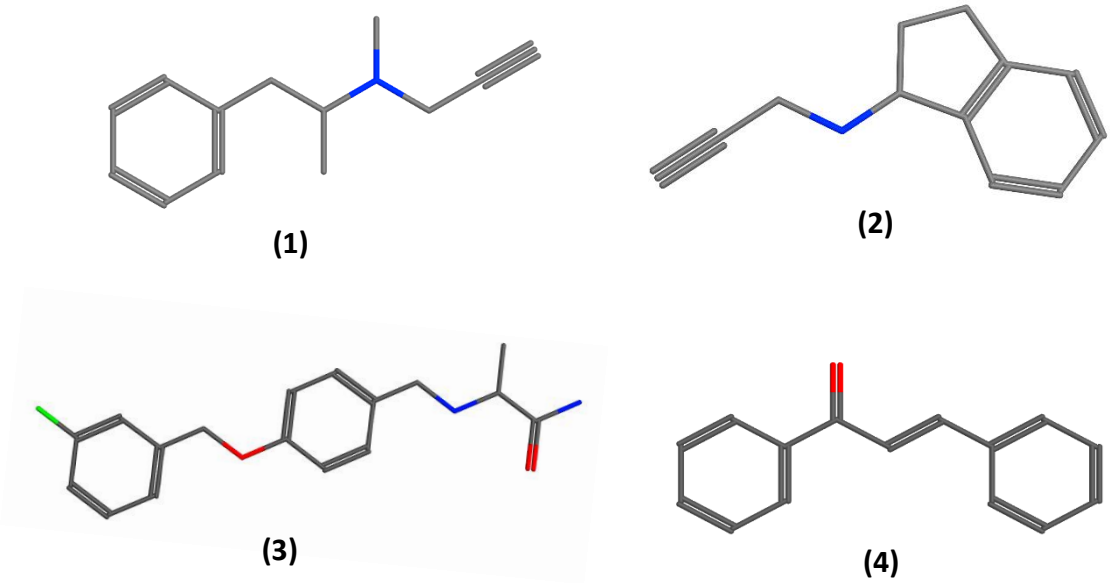


Figure 1. *Chemical structures of MAO-B drugs and chalcone :* ***(1)*** *selegiline,* ***(2)*** *rasagiline,* ***(3)*** *safinamide and* ***(4)*** *chalcone*

very useful tool for screening and evaluating new therapeutic molecular compounds capable of interacting with particular molecular receptors [14]. Currently, CADD (computer-aided drug design) methods such as structure-based virtual screening (SBVS) are widely utilized in drug discovery and development [15][16]. These techniques proved to be effective in various stages of the drug development process, reducing both the cost and time required to develop a drug [17]. Homology modeling, virtual screening, molecular docking, and molecular dynamic (MD) simulation are among the most frequently used SBDD strategies due to their wide range of applications in the analysis of molecular recognition events such as binding modes, energetic interactions, and induced conformational changes [18]. In addition to these methods, the pharmacokinetics proprieties of a compound, as well as its toxicity were evaluated through absorption, distribution, metabolism, excretion, and toxicity (ADMET) prediction [19]. In this study, a dataset of 142 chalcone derivatives obtained from various natural plants was collected via an advanced literature search. The objective of this work is to identify new natural chalcones as potential MAO-B inhibitors.

I. Materials and methods

1. Data set collection and optimization

A data set of 142 bioactive secondary metabolites obtained from various plants was collected via an advanced literature search (Table S1 in supplementary data). The collected compounds are known as chalcone derivatives and display a wide range of activities. All collected compounds were initially sketched and minimized by the MM2 force field using ChemBioOffice software. Then, the 3D structures were generated by the DFT approach using Becke-three-Lee-Yang-Parr (B3LYP) with the 6-31G (d, p) basis set aided by Gaussian 09 program package.

2. Homology modeling

The homology modeling was used to build the 3D structure of human monoamine oxidase B (hMAO-B). First, the amino acid sequence of hMAO-B (520 amino acids) was collected from the UniProt database (http://www.uniprot.org/) (sequence ID: P27338). The crystal structure of hMAO receptor chain A (PDB-ID: 1GOS) was relieved from the protein data bank (PDB) and used as a template for homology modeling, based on the result of the basic local alignment search tool (BLAST). The alignment of the template and sequence was done using ClustalW [20] and then the structure of MAO-B was modeled in Modeller V9.23. Among the ten models predicted, the best model based on the lowest root mean square deviation (RMSD), GA341 and optimized protein energy (DOPE) score was selected. The selected model was then energy minimized in GROMACS 5.1.4 package using the CHARMM36 all-atom force field [21]. Then, the structure was taken for evaluation using Ramachandran plot [22], Procheck [23] and Errat [24]. The active site of the modeled structure was generated by MOE software and finally, the generated model was then used for further studies.

3. Virtual screening

Initial virtual screening of 142 natural compounds was done using the AutoDock Vina module implemented in PyRx 0.8 employing the empirical free energy force field and Lamarckian genetic algorithm conformational search with the default settings. The modeled structure (target) of MAO-B and ligands were converted to pdbqt format. Then, a grid box was set to cover the binding pocket of the target with the following dimension in Å: center (X, Y, Z) = (56.50, 138.70, 16.13) and dimensions (X, Y, Z) = (31, 41, 30) with an exhaustiveness of 8. The top-ranked compounds were then evaluated using ADMET analysis, molecular docking and MD simulations.

4. ADMET analysis

ADMET is a pharmacokinetic short form for absorption, distribution, metabolism, excretion and toxicity and it is important in drug discovery [25]. In this study, it is used to assure the safety and efficacy level of the hit ranked compounds. These include Caco2 permeability, human intestinal absorption (HIA), blood brain barrier (BBB), total tolearance, AMES toxicity, hepatotoxicity, and human max. tolerated dose. The interpretation of each parameter is shown in **Table 1**. All calculation was performed using the pKCM web server [26], which accepts compounds in the SMILES format.

Table 1. *Interpretation of some ADMET parameters*

Parameter	Unit	Interpretation
Caco2 permeability	Numeric (log Papp in 10-6 cm/s)	High Caco2 permeability translate in predicted values > 0.90.
Human intestinal absorption (HIA)	Numeric (% Absorbed)	Compounds with HIA less than 30% are considered poorly absorbed in human intestine.
Blood brain barrier (BBB) permeability	Numeric (log BB)	A compound is considered to penetrate the BBB easily if the value of Log BB is greater than 0.3 while compounds with LogBB less than -1 are poorly distributed to the brain.
Total tolearance	Numeric (log ml/min/kg)	The predicted total tolearance log (CLtot) of a given compound is given in log (ml/min/kg).
AMES toxicity	Categorical (Yes/No)	Compounds with Ames positive are considered mutagenic.
Hepatotoxicity	Categorical (Yes/No)	It predicts whether a given compound is likely to be associated with disrupted normal function of the liver.
Human max. tolerated dose	Numeric (log mg/kg/day)	Compounds with a maximum recommended tolerated dose of less or equal to 0.477 log mg/kg/day are considered low, and high if greater than 0.477.

5. Molecular docking

Molecular docking is an essential part of drug design process, and it is used in this work to validate the binding affinity of the top 10 screened compounds. AutoDock 4.2 [27] was used to perform the docking study which requires both the receptor and ligands to be in pdbqt format. All compounds were docked in the same previously binding site using 50 iterations. The most reasonable binding configurations were then subjected to MD simulation, taking into account both ligand-receptor interactions and binding affinity. The results were analyzed and visualized by PyMol and Discovery studio software's.

6. Molecular dynamic (MD) simulation

Dynamic behavior and stability of ligand-protein complexes were evaluated using MD simulation on the GROMACS 5.1.4 package [28]. The CHARMM36 all-atom force field [21] was used to generate protein topology. The ligand topology and parameters were produced using the General Force Field (CGenFF) server [29]. The TIP3P water model was used to solve each system [30], followed by neutralization with the appropriate concentrations of Na^+ and Cl^- [31]. The energy of each system was then minimized using the steepest descent algorithm with a maximum of 50,000 steps and a maximum force of 10.0 KJ/mol. The temperature and pressure were maintained at 300 K and 1.01325 bar, respectively. Each system was equilibrated with NVT and NPT ensembles. After that, the ligand-receptor complexes were subjected to 100 ns of MD simulation with a time step of 2 femtoseconds (fs). Based on the dynamics trajectory, various parameters such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), number of hydrogen bonds (HB) between protein and ligand were calculated using in-built functions of CROMACS 5.1.4 software.

II. Results and discussion

1. Homology model

3D structure of hMAO-B was built using Modeller V9.23. The initial BLAST search results showed that the sequence of hMAO-B (520 aa) shares 100% sequence identity with the Chain A of the selected template (PDB ID: 1GOS). The sequence alignment between the template and the hMAO-B modeled structure is shown in **Figure S1**. Among the generated models, the best one was selected based on the lowest DOPE score (-58176.859). Additionally, the GA341 score, which indicates a model's reliability, was determined to be 1. The 3D-structure of the modeled structure is shown in **Figure 2**.



Figure 2. *3D structure of the modeled hMAO-B protein*

1.1. Energy minimization

Energy minimization was performed in GRROMACS 5.1.4 package using the CHARMM36 all-atom force field. First, the modeled protein was immersed in a triclinic box of TIP3P water molecules with a distance of 1.0 nm [30]. Sodium ions (Na+) were added to neutralize the charge of the system. Then, the system were subjected to 50,000 steps of energy minimization using the steepest descent algorithm. The minimized structure was used in further studies.

1.2. Validation of the retained homology hMAO-B model

The quality of the chosen model was then evaluated by calculating the RMSD of the superimposed target and template structures, Ramachandran plots, Verify-3D and ProSA (z-score). As shown in **Figure 3 (a)**, the superimposition of the modeled structure and its template resulted in a backbone RMSD value of 0.144 Å. The Ramachandran plot is shown in **Figure 3 (b)**, which showed the distribution of psi and phi angles for the amino acid residues. The result showed a good distribution of the amino acid residues of the hMAO-B model. About 94.6% (423 residues) of the residues in favored regions, 4.5% residues (20 residues) in additional allowed regions, 0.2% residues (1 residues) in generously allowed regions and only 0.7% residues (3 residues) in disallowed regions. The plot of ProSA z-score (**Figure 3 (c)**) indicates that the overall quality is within the range of scores. The results of the above validation methods revealed that the homology model of hMAO-B is satisfactory and reliable for further studies

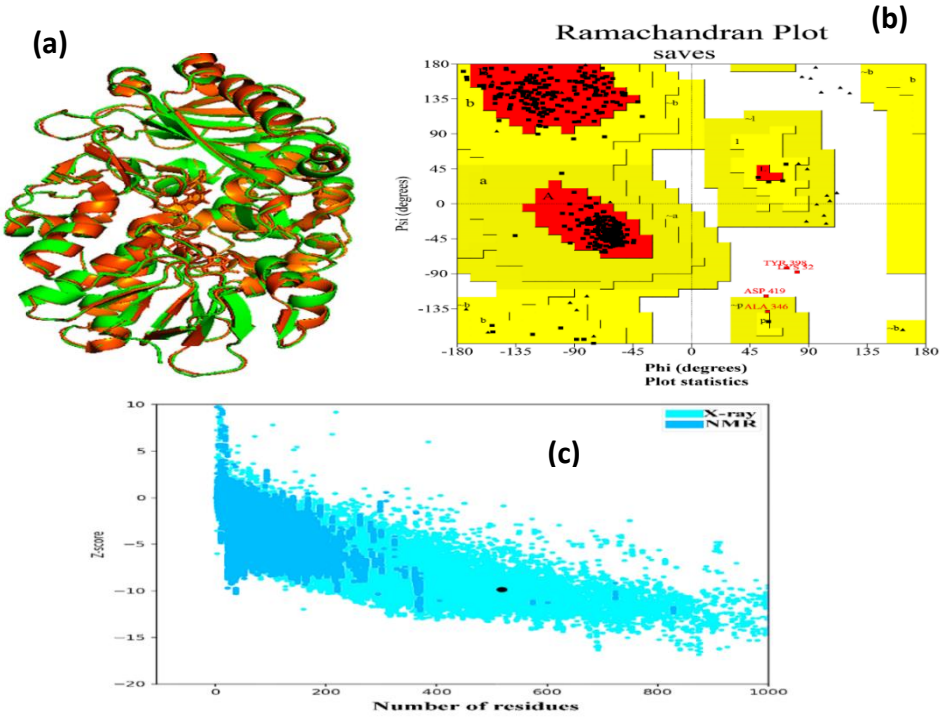


Figure 3. *Homology modeling of hMAO-B protein. (a), superposition of 3D structures between the modeled structure and the 1GOS (Chain A) template, (b), Ramachandran plot of best hMAO-B model, (c), ProSA z-score plot of the modeled structure*

1.3. Binding pocket

Before modeling, the binding pocket of the modeled structure must be determined. MOE software [32] was used to identify the binding pocket of the modeled structure (**Figure 4**). The site with the highest site score was selected out of 26 different sites offered by MOE software. 48 residues were used to line the active site of hMAO-B receptor. All residues found within 3Å of the active site were examined and utilized to generate the docking grid.

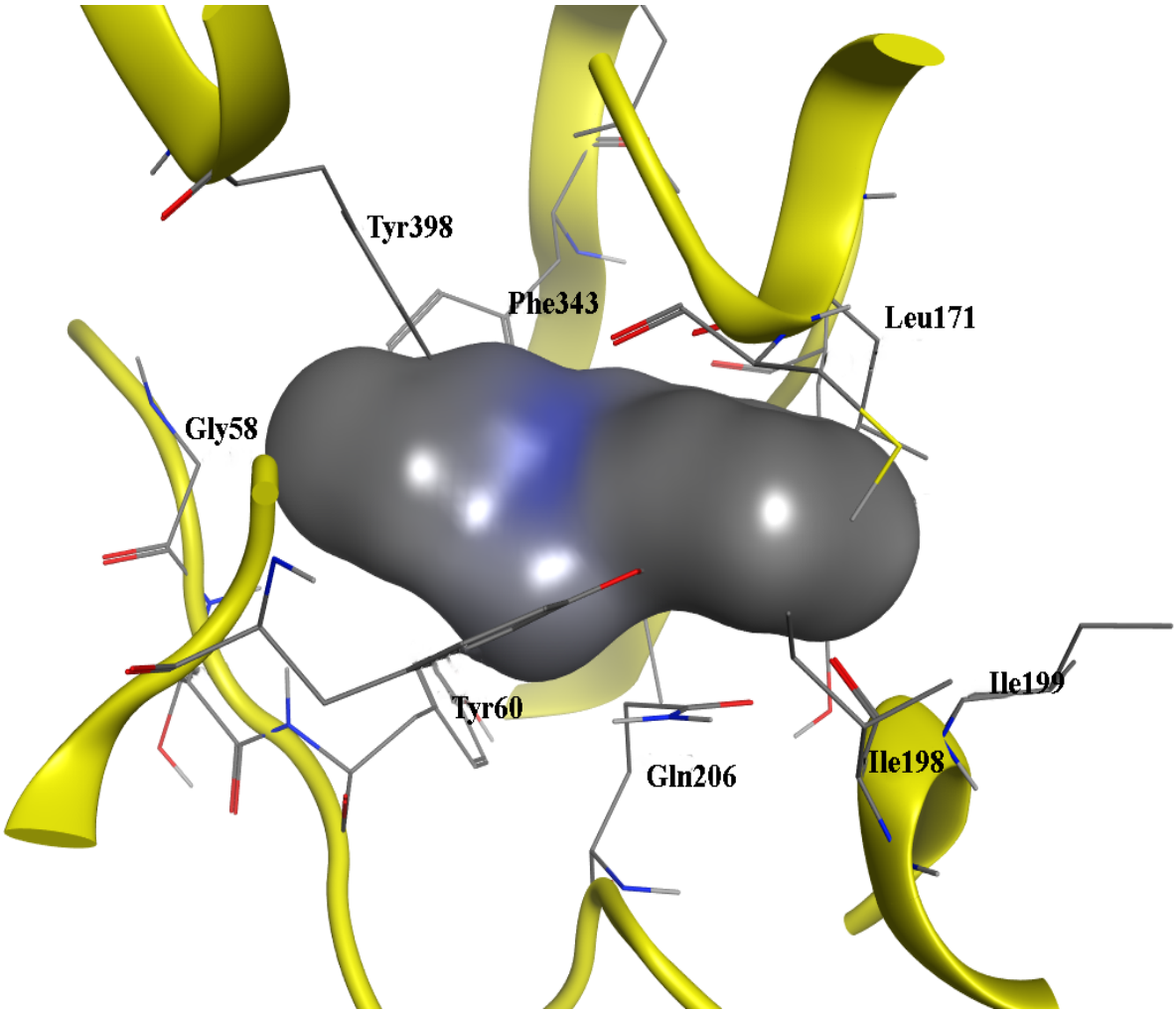


Figure 4. *Binding pocket of the modeled hMAO-B protein*

2. Virtual screening

Initial virtual screening of 142 natural compounds was carried out using autodock vina, implemented in PyRx software. The lowest binding scores between all compounds and target varied between -2.8 and -11.1 kcal/mol, with an average of -8.27 kcal/mol (see Table S2 in supplementary data). Compounds with a binding affinity of less than -9.0 Kcal/mol were considered best for hMAO-B inhibition. The top 43 ranked compounds were selected for further analysis (**Table 2**). We also performed docking studies on rasagiline, safinamid and selegiline

as comparative standard drugs. All selected compounds have shown the potential to inhibit the enzymatic activity of hMAO-B, compared to standard drugs (**Table 2**).

Table 2. *Binding affinity of the best ranked compounds*

No	Binding Affinity (Kcal/mol)	No	Binding Affinity (Kcal/mol)	No	Binding Affinity (Kcal/mol)
P1	-9.3	P45	-9.1	P107	-9.5
P4	-9.7	P53	-9.2	P110	-10.5
P6	-10.4	P56	-9.9	P112	-9.8
P7	-10.1	P60	-9.6	P113	-9.2
P10	-9.2	P61	-9.4	P114	-10.2
P11	-9.9	P62	-9.3	P115	-10.0
P13	-9.8	P70	-10.1	P116	-10.9
P15	-9.2	P72	-11.1	P117	-9.1
P16	-9.9	P81	-9.1	P121	-9.2
P18	-10.4	P92	-9.0	P122	-9.0
P22	-10.2	P96	-9.4	P129	-9.1
P26	-9.6	P97	-9.0	Rasagiline	-6.8
P31	-9.3	P100	-10.7	Safinamid	-8.9
P33	-10.2	P101	-10.1	Selegiline	-6.2
P34	-9.8	P104	-9.2		
P35	-9.8	P106	-9.8		

3. ADMET analysis

Pharmacokinetic proprieties and toxicity of the hit-ranked compounds by AutoDock Vina were then submitted to ADMET analysis. Within the 43 screened compounds, 28 were found to have a high Caco2 permeability, 42 were found to have a high absorption level in the human intestine (HIA), and 29 were predicted to be able to cross the BBB. In addition, 35 were found to be non-

toxic in the AMES test, and 41 were considered to have no hepatotoxicity. The results are presented in **Table S3** (see supplementary data). Based on the ADMET analysis, 10 compounds were predicted to have ADME characteristics with no side effects (**Table 4**).

Table 4. *ADMET characteristics of the top ranked compounds*

No	LogBB	Caco2 permeability	HIA	Total tolearance	AMES toxicity	Hepatotoxicity	Max. tolerated dose (human) (log mg/kg/day)
P34	0.04	1.34	94.02	0.24	No	No	0.42
P56	0.19	1.18	92.45	0.12	No	No	0.06
P60	-0.14	1.32	94.61	0.13	No	No	046
P61	-0.92	1.03	93.03	0.14	No	No	0.76
P62	-0.94	1.01	93.22	0.24	No	No	0.77
P81	0.02	1.35	94.72	0.21	No	No	0.19
P96	0.06	0.93	93.16	0.25	No	No	0.31
P104	-0.014	1.35	94.31	0.38	No	No	-0.03
P113	-0.8	1.06	90.45	0.24	No	No	0.18
P117	0.07	1.39	93.82	0.14	No	No	0.40
Rasagiline	0.62	1.52	94.67	1.20	No	Yes	0.31
Safinamid	-0.26	1.015	94.08	1.02	Yes	No	0.53
Selegiline	1.02	1.52	92.57	1.01	No	Yes	0.19

4. Molecular docking analysis

The hit-ranked compounds were then submitted to molecular docking using AutoDock 4.2. The results are summarized in **Table 5**. As described in this table, all docked compounds showed higher binding affinity compared to the standard hMAO-B drugs. Additionally, we can see that

compounds P34, P60, and P81 had the highest binding affinity for the hMAO-B protein's binding pocket. The constant of inhibition (KI) values for P34 (Ki= 0.047 µM), P60 (Ki= 0.013 µM), and P81 (Ki= 0.016 µM) indicate that these compounds are very likely to be active against the hMAO-B enzyme.

Table 5. *Binding affinity of the top ranked compounds by ADMET analysis*

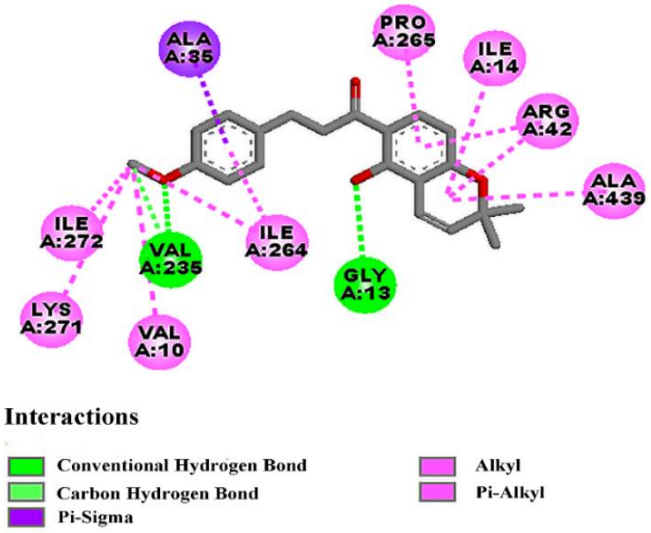
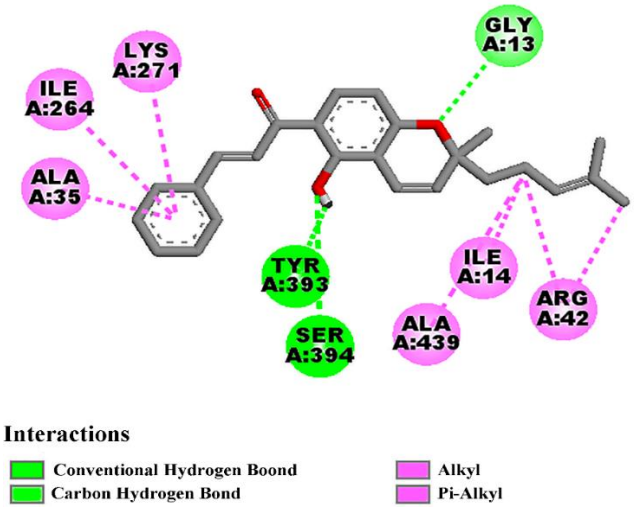
No	Binding affinity Autodock	Ki (µM)
P34	**-10.0**	**0.047**
P56	-8.81	0.350
P60	**-10.77**	**0.013**
P61	-8.64	0.468
P62	8.01	1.34
P81	**-10.64**	**0.016**
P96	-7.88	1.68
P104	-8.32	0.803
P113	-5.6	78.2
P117	-8.88	0.309
Rasagiline	-6.28	25
Safinamid	-7.91	1.6
Selegiline	-5.55	85.61

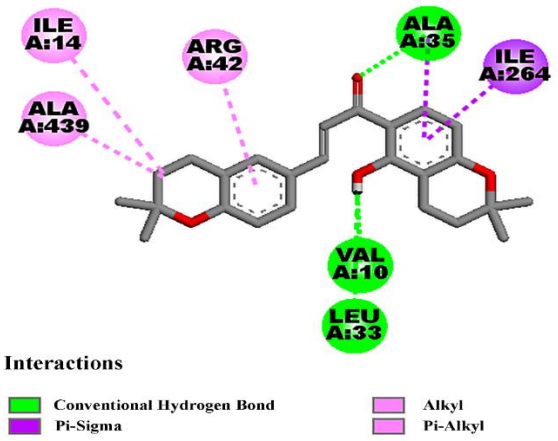
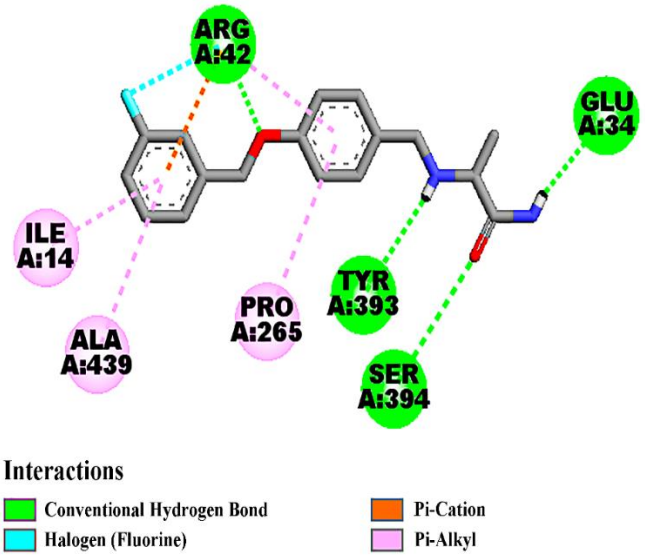
In order to predict the binding modes and interactions between the ligand and receptor, compounds P34, P60 and P81 were selected. The results are presented in **Table 6**. The interaction analysis using discovery studio showed that the high binding affinity of compound P60 (-10.77 Kcal/mol) is due to the formation of three hydrogen bonds with TYR 393, SER 394, and GLY 13. The hydrogen bond interactions of compound P34 were observed in VAL 235 and GLY 13, and P81 hydrogen bond interactions were ALA 35, VAL 10 and LEU 33. Moreover, safinamid, which has a high binding affinity (-7.91 Kcal/mol) compared with the other reference drugs (Rasagiline and Selegiline), formed four hydrogen bond interactions with ARG 42, TYR 393, SER 394, and LEU 34. In addition to hydrogen bonds, several types of interactions were observed in the inhibitor-enzyme complexes, including hydrophobic (pi-Alkyl, alkyl, and pi-Sigma) and electrostatic (pi-cation) interactions. **Figure 5** illustrates the 3D-binding site between the target protein and the top-ranked compounds.

Similarly, the results of inhibitor constant (Ki) values suggested that substances with high binding affinity had low Ki. From these results, we can conclude that the excellent activity of compound P60 could be related to the interaction of hydrogen bonds with residues TYR393 and SER 394. This result is also confirmed by Abell et al., who demonstrate the role of TYR

393, GLU 34 and ARG 42 in the cataliytic activity of hMAO-B [33]. Also, a detailed analysis of the other ranked compounds (see **Table S4** and **Figure S2** in the supplementary data) showed the importance of hydrogen bond interactions with the residues TYR 393, SER 394, ARG 42, and GLU 34, and hydrophobic interactions with the residues ILE 14, ARG 42, ALA 439, LYS 271, and ILE 264 in the stability of chalcone derivatives in the binding pocket of the hMAO-B receptor.

Table 6. *2D-binding site of P34, P60, P81, and safinamid in the binding pocket of hMAO-B receptor*

NO	Hydrogen Bonds	Others Interactions	2D-Binding site of interactions
P34	VAL 235, GLY 13	ALA 35, ILE 272, LYS 271, VAL 10, ILE 264, PRO 265, ILE 14, ARG 42, ALA 439	
P60	TYR 393, SER 394, GLY 13	ALA 35, ILE 264, LYS 271, ILE 14, ALA 439, ARG 42	

P81	ALA 35, VAL 10, LEU 33	ILE 14, ARG 42, ALA 439, ILE 264	
Safinamid	ARG 42, TYR 393, SER 394, LEU 34	ARG 42, ILE 14, ALA 439, PRO 265,	

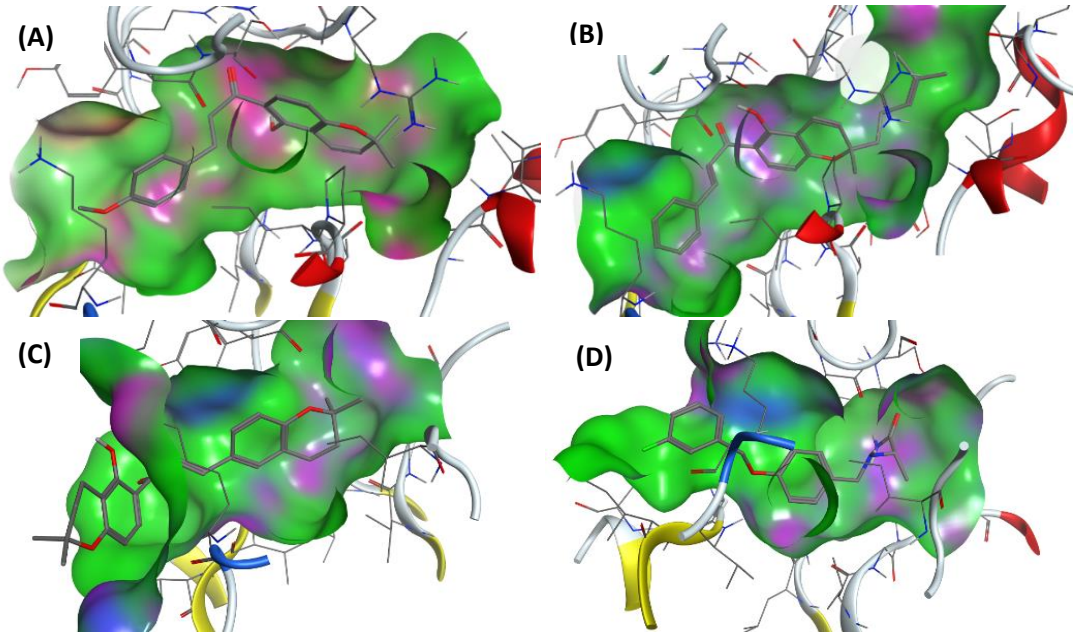


Figure 5. *3D-binding site between the target protein and the top-ranked compounds.* ***(A)*** *P34 (-10.0 Kcal/mol),* ***(B)*** *P60 (-10.77 Kcal/mol),* ***(C)*** *P81 (-10.64 Kcal/mol) and* ***(D)*** *Safinamid (-7.91 Kcal/mol)*

5. MD simulation

To study the stability of a ligand inside a protein's binding site, molecular dynamics (MD) simulations give a deeper understanding of the ligand through a variety of statistical parameters. In a similar manner, the top-ranked compounds (P60 and P81) by Autodock 4.2 were subjected to MD simulations. Each system was simulated for 100 ns using the GROMACS 5.1.4 package. Throughout the procedure, every system was monitored and analyzed in detail, including root mean square deviation (RMSD) and fluctuation (RMSF), and radius of gyration (Rg).

Calculating the root mean square deviation (RMSD) is a good way for examining conformational changes and estimating the stability of proteins and complexes. The RMSD plots of backbone atoms and ligands were presented in **Figure 6 (a)** and **7 (b)**. As seen in **Figure 6 (a)**, at the start of the simulations, specifically in the first 50 ns, all systems exhibited a fast increase in RMSD values. After the 50 ns delay, the systems remained relatively stable with average values of 0.3, 0.43, and 0.45 for hMAO-B_P60 (black), hMAO-B_P81(red), and hMAO-B_Safinamid (green) complexes, respectively. However, the average RMSD values of ligands P60, P80, and safinamide were 0.11, 0.08, and 0.19 nm, respectively (**Figure 6 (b)**). As a result, it can be concluded that the binding of ligands to proteins results in structure stability of complexes.

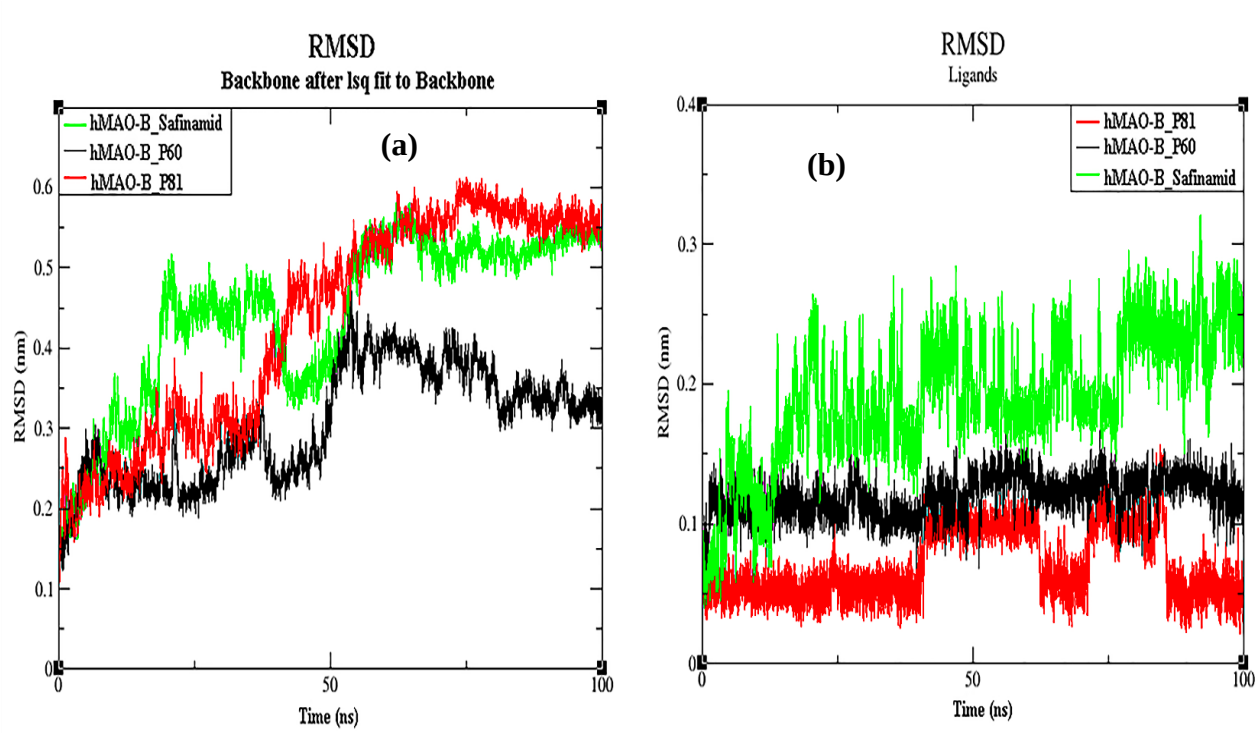


Figure 6. *RMSD plots of 100 ns of MD simulat ion.* ***(a)*** *For backbone and* ***(b)*** *for ligand. In all figures, the red black represents the hMAO-B_P60 complex, the red represents hMAO-B_P81 complex, and the green color represents the hMAO-B_Safinamid complex*

The root-mean-square fluctuation (RMSF) of hMAO-B in three simulated systems (hMAO-B_P60 (black), hMAO-B_P81 (red), and hMAO-B_Safinamid (green)) was calculated and graphically plotted in **Figure 7 (a)**. This parameter is an indicator of residual flexibility. The RMSF graphs demonstrate that only a few residues fluctuated with an intensity greater than 3Å. Additionally, the active site residues of the simulated systems showed fluctuations < 3 Å, which indicated that active site regions of hMAO-b were stabilized by ligands. For example, the most stable complex (hMAO-B_P60) showed low RMSF values at TRY 393, and TYR398 residues.

The radius of gyration (Rg) of all simulated systems was calculated and graphically plotted in **Figure 7 (b)**. Throughout the 100 ns MD simulation, all systems remained stable and compact, with an average Rg values of 2.32 nm for hMAO-B P60 (black), 2.36 nm for hMAO-B P81 (red), and 2.37 nm for hMAO-B Safinamid. The lower Rg value of the hMAO-B P60 system may be explained by the high stability of P60 in the binding pocket of the hMAO-B receptor.

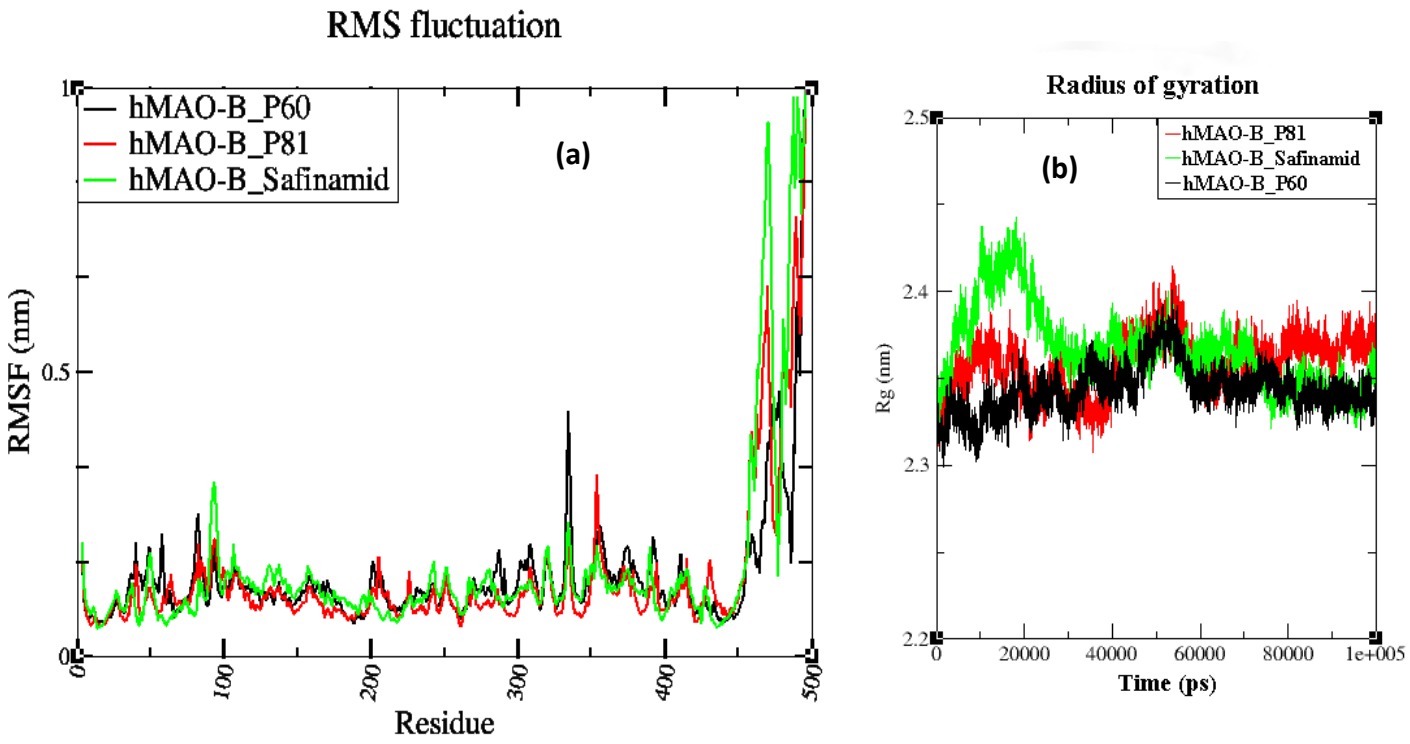


Figure 7. *MD simulation of ligand-protein complex.* ***(a)*** *and* ***(b)*** *represent, respectively, the plots of RMSF (C-alpha) and Rg over the 100 ns of MD silulation at 300 K. In all figures, the black, red, and green colors represent the hMAO-B_P60, hMAO-B_P81, and hMAO-B_Safinamid complexes, respectively*

III. Conclusion

In this work, a dataset of 142 chalcone derivatives, collected from various natural plants, was screened by combining homology modeling, virtual screening, ADMET analysis, molecular docking, and molecular dynamic (MD) simulation. The objective is to discover novel natural chalcones as potential MAO-B inhibitors. First, the 3D structure of human MAO-B (hMAO-B) was built using homology modeling. The dataset of compounds was sketched and initially minimized by the MM2 force field, then the 3D structures were generated using the DFT approach. All compounds were screened using AutoDock Vina, implicated in PyRx. Among the 142 natural compounds, 43 were selected for ADMET analysis. Based on the results of ADMET analysis, 10 compounds were evaluated by using AutoDock 4.2. Finally, MD simulation was performed to evaluate the stability of the top ranked compounds (P60 and P81) in the binding pocket of the modeled hMAO-B receptor. This research could lead to the development of new natural compounds that act as hMAO-B inhibitors.

Acknowledgment

The authors want to thank the Moroccan Association of Theoretical Chemistry (MATC) for its relevant help regarding to the Software.

Conflicting interests

The authors declare that they have no conflicting interests in regard to the publication of this article.

References

[1] D.B. Marin, L.M. Bierera, B.A. Lawlorb, T.M. Ryana, R. Jacobsonc, J. Schmeidlerd, R.C. Mohsa, K.L. Davisd, L-Deprenyl and physostigmine for the treatment of Alzheimer ' s disease, Psychiatry Res. 58 (1995). https://doi.org/https://doi.org/10.1016/0165-1781(95)02714-8.

[2] S. Jo, O. Yarishkin, Y.J. Hwang, Y.E. Chun, M. Park, D.H. Woo, J.Y. Bae, T. Kim, J. Lee, H. Chun, H.J. Park, D.Y. Lee, J. Hong, H.Y. Kim, S. Oh, S.J. Park, H. Lee, B. Yoon, Y. Kim, Y. Jeong, I. Shim, Y.C. Bae, J. Cho, N.W. Kowall, H. Ryu, E. Hwang, D. Kim, C.J. Lee, ARTICLES GABA from reactive astrocytes impairs memory in mouse models of Alzheimer ' s disease, Nat. Publ. Gr. 20 (2014) 886–896. https://doi.org/10.1038/nm.3639.

[3] M. Fabbri, M.M. Rosa, D. Abreu, J.J. Ferreira, Clinical pharmacology review of safinamide for the treatment of Parkinson ' s disease, Futur. Med. 5 (2015). https://doi.org/https://doi.org/10.2217/nmt.15.46.

[4] M.J. Matos, S. Vazquez-Rodriguez, E. Uriarte, L. Santana, Potential pharmacological uses of chalcones: A patent review (from June 2011-2014), Expert Opin. Ther. Pat. 25 (2015) 351–366. https://doi.org/10.1517/13543776.2014.995627.

[5] Z. Nowakowska, A review of anti-infective and anti-inflammatory chalcones, Eur. J. Med. Chem. 42 (2007) 125–137. https://doi.org/10.1016/j.ejmech.2006.09.019.

[6] P. Boeck, P.C. Leal, R.A. Yunes, V. Cechinel Filho, S. López, M. Sortino, A. Escalante, R.L.E. Furlán, S. Zacchino, Antifungal activity and studies on mode of action of novel xanthoxyline-derived chalcones, Arch. Pharm. (Weinheim). 338 (2005) 87–95. https://doi.org/10.1002/ardp.200400929.

[7] A.M. Clark, Maclura, Planta Med. 67 (2001) 87–89. https://doi.org/10.1055/s-2001-10621.

[8] N. Yayli, O. Üçüncü, E. Aydin, Y. Gök, A. Yaşar, C. Baltaei, N. Yildirim, M. Küçük, Stereoselective photochemistry of heteroaryl chalcones in solution and the antioxidant activities, J. Photochem. Photobiol. A Chem. 169 (2005) 229–234. https://doi.org/10.1016/j.jphotochem.2004.06.015.

[9] C. Nakamura, N. Kawasaki, H. Miyataka, E. Jayachandran, I.H. Kim, K.L. Kirk, T. Taguchi, Y. Takeuchi, H. Hori, T. Satoh, Synthesis and biological activities of fluorinated chalcone derivatives, Bioorganic Med. Chem. 10 (2002) 699–706. https://doi.org/10.1016/S0968-0896(01)00319-4.

[10] A. Modzelewska, C. Pettit, G. Achanta, N.E. Davidson, P. Huang, S.R. Khan, Anticancer activities of novel chalcone and bis-chalcone derivatives, Bioorganic Med. Chem. 14 (2006) 3491–3495. https://doi.org/10.1016/j.bmc.2006.01.003.

[11] P. Singh, A. Anand, V. Kumar, Recent developments in biological activities of chalcones: A mini review, Eur. J. Med. Chem. 85 (2014) 758–777. https://doi.org/10.1016/j.ejmech.2014.08.033.

[12] B. Lakshminarayanan, S.C. Baek, J.P. Lee, N. Kannappan, G.F. Mangiatordi, O. Nicolotti, T. Subburaju, H. Kim, B. Mathew, Ethoxylated Head of Chalcones as a New Class of Multi-Targeted MAO Inhibitors, ChemistrySelect. 4 (2019) 6614–6619. https://doi.org/10.1002/slct.201901093.

[13] H.U.M. Valdecilla, e on is ly us fo on rs Pe ot r d tri e on is us fo rs Pe r d al n, Lett. Org. Chem. 12 (2015) 405–410.

[14] B.L. Larvol, L.J. Wilkerson, In silico drug discovery: tools for bridging the NCE gap., Nat. Biotechnol. 16 (1998) 33–34. https://doi.org/10.1038/5412.

[15] A. El Aissouq, O. Chedadi, R. Kasmi, L. Elmchichi, F. En-nahli, A. Goudzal, M. Bouachrine, A. Ouammou, F. Khalil, Molecular Modeling Studies of C-Glycosylfavone Derivatives as GSK-3β Inhibitors Based on QSAR and Docking Analysis, J. Solution Chem. 50 (2021) 808–822. https://doi.org/10.1007/s10953-021-01083-6.

[16] O. Chedadi, A. El Aissouq, Y. El Ouardi, M. Bouachrine, A. Ouammou, 3D-QSAR and molecular docking studies of 4-methyl quinazoline derivatives as PI3Kα inhibitors, J. Indian Chem. Soc. 98 (2021) 100183. https://doi.org/10.1016/j.jics.2021.100183.

[17] A. EL AISSOUQ, M. BOUACHRINE, A. OUAMMOU, F. KHALIL, Computational investigation of unsaturated ketone derivatives as MAO-B inhibitors by using QSAR, ADME/Tox, molecular docking, and molecular dynamics simulations, Turkish J. Chem. 46 (2022) 687–703. https://doi.org/10.55730/1300-0527.3360.

[18] A. El Aissouq, O. Chedadi, M. Bouachrine, A. Ouammou, Identification of Novel SARS-CoV-2 Inhibitors: A Structure-Based Virtual Screening Approach, J. Chem. 2021 (2021). https://doi.org/10.1155/2021/1901484.

[19] L. El Mchichi, A. El Aissouq, R. Kasmi, A. Belhassan, R. El-Mernissi, A. Ouammou, T. Lakhlifi, M. Bouachrine, In silico design of novel Pyrazole derivatives containing thiourea skeleton as anti-cancer agents using: 3D QSAR, Drug-Likeness studies, ADMET prediction and molecular docking, Mater. Today Proc. (2021). https://doi.org/10.1016/j.matpr.2021.03.152.

[20] W. Clustal, improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice Thompson, Julie D.; Higgins, Desmond G.; Gibson, Toby J, Nucleic Acids Res. 22 (1994) 4673–4680. https://doi.org/https://doi.org/10.1093/nar/22.22.4673.

[21] R.B. Best, X. Zhu, J. Shim, P.E.M. Lopes, J. Mittal, M. Feig, A.D. Mackerell, Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ , ψ and Side-Chain χ 1 and χ 2 Dihedral Angles, J. Chem. Theory Comput. 8 (2012) 3257–3273. https://doi.org/https://doi.org/10.1021/ct300400x.

[22] S.S. Sheik, P. Sundararajan, A.S.Z. Hussain, K. Sekar, Ramachandran plot on the web, Bioinformatics. 18 (2002) 1548–1549. https://doi.org/10.1093/bioinformatics/18.11.1548.

[23] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291. https://doi.org/10.1107/s0021889892009944.

[24] C. Colovos, T.O. Yeates, Verification of protein structures: Patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519. https://doi.org/10.1002/pro.5560020916.

[25] L.G. Ferreira, R.N. Dos Santos, G. Oliva, A.D. Andricopulo, Molecular docking and structure-based drug design strategies, Molecules. 20 (2015) 13384–13421. https://doi.org/10.3390/molecules200713384.

[26] H.S. Roth, R.C. Botham, S.C. Schmid, T.M. Fan, L. Dirikolu, P.J. Hergenrother, Removal of Metabolic Liabilities Enables Development of Derivatives of Procaspase-Activating Compound 1 (PAC-1) with Improved Pharmacokinetics, 1 (2015). https://doi.org/10.1021/acs.jmedchem.5b00413.

[27] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, Software News and Updates AutoDock4 and AutoDockTools4 : Automated Docking with Selective Receptor Flexibility, J. Comput. Chem. 30 (2009) 2785–2791. https://doi.org/10.1002/jcc.

[28] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1–2 (2015) 19–25. https://doi.org/10.1016/j.softx.2015.06.001.

[29] A.D.M.J. K. VANOMMESLAEGHE, E. HATCHER, C. ACHARYA, S. KUNDU, S. ZHONG, J. SHIM, E. DARIAN, O. GUVENCH, P. LOPES, I. VOROBYOV, CHARMM General Force Field: A

Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields, J. Comput. Chem. 31 (2010) 671–690. https://doi.org/https://doi.org/10.1002/jcc.21367.

[30] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 926 (1983). https://doi.org/10.1063/1.445869.

[31] M.B. and A.O. O. Chedadia, A. El Aissouq, Y. El Ouardi, CHEMISTRY In Silico Prediction of Novel (TRIM24) Bromodomain Inhibitors : A Combination of 3D-QSAR , Molecular Docking , ADMET Prediction , and Molecular Dynamics, 24 (2022). https://doi.org/10.22036/PCR.2022.331866.2040.

[32] http://www.chemcomp.com., Molecular Operating Environment (MOE), (n.d.). https://doi.org/10.1056/NEJMcps1402621.

[33] C.W. Abell, S.W. Kwan, Molecular characterization of monoamine oxidases A and B, Prog. Nucleic Acid Res. Mol. Biol. 65 (2000) 129–132. https://doi.org/10.1016/s0079-6603(00)65004-3.

Partie II : Travaux Réalisés

(Applications sur la Maladie d'Alzheimer)

Contexte:

Les deux chapitres de la partie II visent à étudier l'activité anti-Alzheimer de certaines molécules organiques, biologiquement actives. En effet, plusieurs recherches ont montré le rôle de dérivés d'imidazole, de flavonoïde et autres dans l'inhibition des enzymes responsables de la maladie d'Alzheimer. Dans ces deux derniers chapitres, nous allons étudier l'activité inhibitrice de dérivés de la C-Glycosylfavone et de pyridinyl imidazole en tant qu'inhibiteurs de l'enzyme glycogène synthase kinase-3β (GSK-3β) et de l'enzyme p38α kinase, respectivement.

Chapitre IV : Études de modélisation moléculaire de dérivés de la C-Glycosylfavone en tant qu'inhibiteurs de GSK-3β, basées sur l'étude QSAR et l'analyse de docking

Résumé :

La glycogène synthase kinase-3 (GSK-3) est une sérine/thréonine kinase initialement découverte pour son rôle majeur dans la signalisation intracellulaire induite par l'insuline. Chez les mammifères, cette enzyme existe en deux isoformes étroitement liées : GSK-3α et GSK-3β. Les deux gènes sont presque identiques dans leur domaine catalytique (98% d'identité) mais divergent au niveau des régions N- et C-terminales. La GSK-3β, fortement exprimée dans les neurones, a la particularité d'être active ; elle est régulée négativement par la phosphorylation sur son résidu sérine 9. Son activité est inhibée par plusieurs stimuli externes, notamment l'insuline, les ligands Wnt et la voie Akt. La GSK-3β est également phosphorylée sur la tyrosine 216, probablement par autophosphorylation, ce qui augmente son activité.

Dans ce chapitre, une série de molécules dérivées de la C-Glycosylfavone a été sectionnée pour réaliser des études de relation quantitative structure-activité (QSAR) et de docking moléculaire. Tout d'abord, des études 2D-QSAR ont été appliquées pour relier la variation de l'activité inhibitrice aux descripteurs moléculaires. Le logiciel MOE (Molecular Operating Environment) a été utilisé pour calculer les descripteurs moléculaires. Un ensemble de 354 descripteurs moléculaires a été calculé pour chaque composé de l'ensemble d'apprentissage. Après élimination des descripteurs qui sont corrélés entre eux et des descripteurs qui n'ont aucune influence sur l'activité étudié, la méthode de régression des moindres carrés partiels (PLS) a été utilisée pour générer le modèle 2D-QSAR. Le meilleur modèle sélectionné a donné un coefficient de détermination (r^2) de 0.85 avec une erreur quadratique moyenne (RMSE) de 0.04. La capacité prédictive et la stabilité du modèle élaboré ont été vérifiées par la validation interne et externe. Les résultats montrent que le modèle développé a une meilleure capacité prédictive. Par la suite, une analyse de docking moléculaire a été réalisée pour expliquer les modes d'interactions et de liaisons de ces molécules dans le site actif du récepteur GSK-3β. Sur la base du modèle 2D-QSAR et des résultats du docking moléculaire, une nouvelle série de molécules dérivées de C-Glycosylfavone a été prédite. Toutes les molécules montrent une activité inhibitrice élevée par rapport à la molécule la plus active de la série.

Journal of Solution Chemistry (2021) 50:808–822 https://doi.org/10.1007/s10953-021-01083-6

Molecular Modeling Studies of C- Glycosylfavone Derivatives as GSK- 3β Inhibitors Based on QSAR and Docking Analysis

Abdellah El Aissouq[1] · Oussama Chedadi[2] · Rania Kasmi[2] · Larbi Elmchichi[3] · Fatima En- nahli[3] · Amina Goudzal[2] · Mohammed Bouachrine[3,4] · Abdelkrim Ouammou[2] · Fouad Khalil[1]

Received: 18 September 2020 / Accepted: 16 April 2021 / Published online: 14 May 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Glycogen synthase kinase-3 beta (GSK-3β) is implicated in abnormal hyperphosphorylation of the tau protein and its inhibitors may be a promising therapeutic approach for treating Alzheimer's disease. Here, a series of C-glycosylfavone derivatives as GSK-3β inhibitors was selected to perform two-dimensional quantitative structure activity relationship (2D-QSAR) method and docking analysis. The 2D-QSAR model was generated and validated using a dataset of 23 compounds and a test set of 5 compounds, respectively. The best model selected by the partial-least-squares (PLS) regression method revealed a regression coefficient (r^2) value of 0.85 and the mean-square-error (*MSE*) value of 0.04. The predictive ability and stability of the generated model was verified by external and internal validations, and gave the regression coefficient values of 0.93 and 0.72, respectively. Molecular docking analysis using AutoDock vina was carried out to explain the binding modes of C-glycosylfavone ligands with the GSK-3β receptor. Based on the obtained results, a novel series of C-glycosylfavone derivative was designed and their activity and binding affinity were predicted. The generated work could be helpful for the design and development of novel GSK-3β inhibitors.

Keywords GSK-3β · Alzheimer's disease · C-Glycosylfavone derivatives · 2D-QSAR · Docking analysis

✉ Abdelkrim Ouammou abdelkrim.ouammou@usmba.ac.ma

[1] Laboratory of Processes, Materials and Environment (LPME), Faculty of Science and Technology, Sidi Mohamed Ben Abdellah University, Fez, Morocco

[2] Engineering Laboratory of Organometallic, Molecular Materials and Environment (LIMOME), Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco

[3] MCNS Laboratory, Faculty of Sciences, Moulay Ismail University, Meknes, Morocco

[4] EST Khenifra, Sultan Moulay Sliman University, Beni-Mellal, Morocco

I. Introduction

Glycogen synthase kinase-3 (GSK-3), is a Serine/Threonine kinase that implicated in various human diseases including type 2 diabetes [1], Alzheimer's disease [2] and inflammation [3]. In mammals, two GSK-3 genes have been identified [4], GSK-3α and GSK-3β with β gene being highly expressed in neuronal tissue. Both genes are almost identical within their catalytic domain (98% identity) but diverge at the N- and C-terminal regions [4, 5]. Contrary to their similar structure, GSK-3 genes are known to act on different substrates [6]. GSK-3β has been proposed as the link between the two neuropathological hallmarks of AD [5]. Indeed, it has been suggested that Aβ may favor tau phosphorylation by activation of GSK-3 [7]. Furthermore, inhibition of GSK-3β reduces the production of β-amyloid [8, [9]. GSK-3β is also known to phosphorylate the microtubule-associated protein tau in mammalian cell [10] and reduces acetylcholine synthesis, which is in accordance with the cholinergic deficit present in AD [11]. In addition, several studies have demonstrated that GSK-3β activity is increased in AD brains [12 - 14]. These results suggest that AD in a GSK-3β-dependent manner and GSK-3β inhibition is expected to be a promising therapeutic approach for AD.

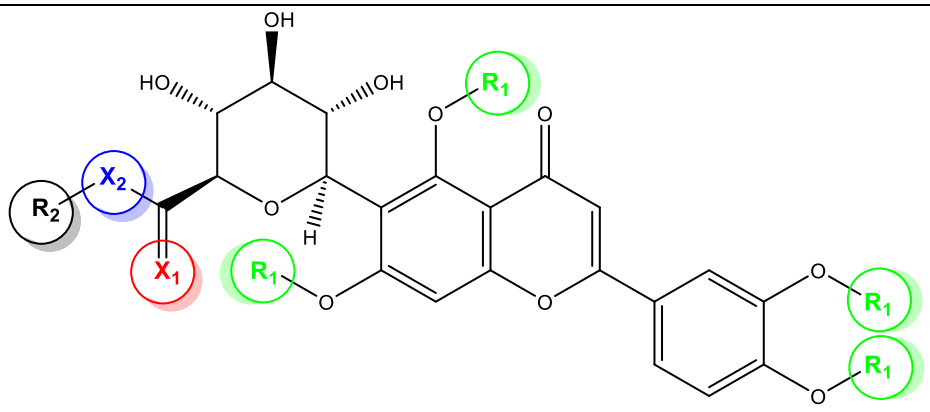
In this paper, a series of C-Glycosylfavone derivatives as GSK-3β inhibitors was used to perform molecular modeling by two-dimensional quantitative structure-activity relationship (2D-QSAR) and molecular docking analysis. The selected molecules are known as GSK-3β inhibitors and have IC50 values ranging from 0.59 µM to 239.2 µM [15]. The QSAR model was built using the values of pIC50 as dependent variables and the molecular operating environment (MOE) descriptors as independent variables. Partial least squares (PLS) regression method was applied in order to explore the mathematical relationships between the both variables [16, 17]. The robustness of the generated QSAR model and its predictive ability was verified by internal and external validations. Y-randomization test [18] was also used to ensure the robustness of the generated QSAR model. In order to validate the generated QSAR model and to study the binding site interaction and energy between the C-Glycosylfavone ligands and GSK-3β receptor molecular docking analysis was used. The aims of this study are to find a relationship between structure-activity and to identify new C-Glycosylfavone derivatives with high potence activity.

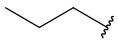
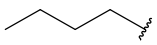
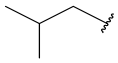
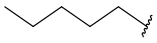
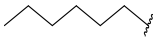
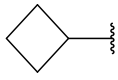
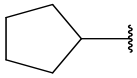
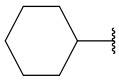
II. Materials and Methods

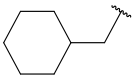
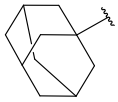
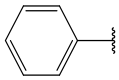
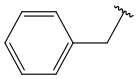
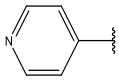
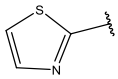
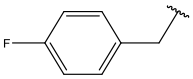
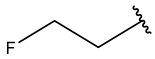
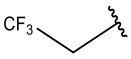
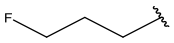
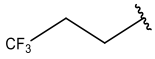
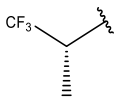
1. Dataset collection

A set of 28 C-Glycosylfavone derivative as Glycogen Synthase Kinase-3 beta (GSK-3β) inhibitors was collected from literature [15]. The IC_{50} values (concentration required for inhibition 50% of GSK-3β activity) of 28 compounds were converted into the corresponding pIC_{50} [$pIC_{50} = \log (1/IC_{50})$] and subsequently used as a dependent variable in the 2D-QSAR analysis. The chemical structures of all compounds were designed and optimized using ChemDraw and and Chem3D software's. Then, the total of compounds was randomly divided into a training set (23 compounds) for generating the 2D-QSAR model and a test set (5 compounds) for validating the predictive ability of the generated model. The structures and their biological activities are presented in Table 1.

Table 1. *Chemical Structures and their experimental biological activity of C-Glycosylfavones*



cmpd	Functional group				GSK-3β inhibition	
No.	R1	X1	X2	R2	IC_{50} (µM)	pIC_{50}
L1	H	CH_2	O	H	184.9	3.7331
L2	CH3	CH_2	O	H	239.2	3.6212
L3	CH3	O	O	H	237.3	3.6247
L4	CH3	O	O	CH_3	135	3.8697
L5	CH_3	O	NH		29.2	4.5346
L6^T	CH_3	O	NH		5.4	5.2676
L7	CH_3	O	NH		33.3	4.4776
L8	CH_3	O	NH		31.1	4.5072
L9	CH_3	O	NH		28	4.5528
L10	CH_3	O	NH		9.4	5.0269
L11^T	CH_3	O	NH		25.8	4.5884
L12	CH_3	O	NH		61.9	4.2083
L13^T	CH_3	O	NH		13.1	4.8827
L14	CH_3	O	NH		9	5.0458

L15	CH_3	O	NH		26	4.5850
L16	CH_3	O	NH		33.5	4.4750
L17	CH_3	O	NH		19.7	4.7055
L18	CH_3	O	NH		15.8	4.8013
L19	CH_3	O	NH		35.6	4.4486
L20[T]	CH_3	O	NH		35.1	4.4547
L21	CH_3	O	NH		56.2	4.2503
L22	CH_3	O	NH		21.7	4.6635
L23	CH_3	O	NH		34.8	4.4584
L24	CH_3	O	NH		19.3	4.7144
L25	CH_3	O	NH		49.5	4.3054
L26	CH_3	O	NH		17.2	4.7645
L27	CH_3	O	NH		0.59	6.2291
L28[T]	CH_3	O	NH		2.3	5.6383

[T] represented the compounds of the test set

2. 2D-QSAR model generation

2.1 Molecular descriptors

The codification of the chemical structure by a diversity of molecular descriptors is a necessary step in QSAR modeling. In this study, a total of 354 descriptors were calculated for each compound in the dataset using MOE 2015 software (http://www.chemcomp.com). The descriptors having the same value or highly inter-correlated were deleted using correlation matrix, option in MOE package, the others were selected. The number of descriptors was then optimized by removing the low correlated with the value of pIC_{50}. Then, the partial least squares

(PLS) analysis [19] implicated in MOE and XLSTAT software's was applied to link the best selected descriptor with the values of pIC_{50}.

2.2 Partial least squares (PLS) analysis

Partial least squares (PLS) regression is a linear regression method [20, 21], especially useful in the cases where the number of descriptors is greater than the number of observations (dataset compounds), and in the cases where the dataset compounds contains highly inter-correlated descriptors. Here, the PLS method was applied to generate the linear 2D-QSAR model, using the values of pIC_{50} as dependent variables (Y variables) and the significant selected descriptors as independent variables (X variables). The goodness of the model is evaluated by how well it performs mapping between the X and Y variables in the training set. This mapping is determined by the squared correlation coefficient (r^2), the root mean square error (RMSE), and the Fischer ratio value (F value).

$$R^2 = 1 - \left[\frac{\Sigma_i (Y_{i\,OBS} - Y_{i\,PRED})^2}{\Sigma_i (Y_{i\,OBS} - \overline{Y_{Mean}})^2}\right] \qquad \text{Eq.1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Y_{i\,OBS} - Y_{i\,PRED})}{N}} \qquad \text{Eq.2}$$

In the above equations, $Y_{i\,OBS}$ and $Y_{i\,PRED}$ are the observed and the predicted activities of compound i, respectively, while $\overline{Y_{Mean}}$ is the mean of observed activity over N compounds in the training set.

2.3 2D-QSAR model validation

To ensure the performance and robustness of the generated QSAR model, internal and external validations were used. The internal and external validations were determined using the leave-one-out (q^2) cross validation method and a test set (r^2_{test}) of 5 compounds. A high value of q^2 and r^2_{test} is considered as a good indicator of high predictive QSAR model [22, 23]. The robustness of the established QSAR model was also evaluated by Y-randomization test [18, 24].

2.4 Applicability Domain

The applicability domain is defined as a theoretical region in chemical space including both the model variables and modeled response [25]. The applicability domain is based on the calculation of the leverage value (h_i) of each compound i, for which QSAR model is used to predict its activity:

$$h_i = x_i^T (X^T X)^{-1} x_i \qquad (i = 1, \ldots, n), \qquad \text{Eq.3}$$

Where x_i is the descriptor row-vector of a query compound, and X is the n*(k-1) matrix of k descriptor values for n data set compounds. A compound i was considered outside the

applicability domain, when the leverage value (h_i) of this compound is greater than the critical value (h*) (h* = 3(d+1)/n, were d is the number of variables and n is a number of compounds in training set). Conversely, a compound i was considered inside the applicability domain, when the h_i is lower than the h*.

III. Molecular docking

Molecular docking study was carried out to understand the binding interactions between ligand and GSK-3β receptor. Firstly, the x-ray diffraction structure of GSK-3β protein was selected from the RCSB Protein Data Bank (PDB ID: 1PYX) [26]. The original ligand was extracted from the protein, the water molecules were removed, and the polar hydrogen atoms were added to the protein. AutoDock tool [27] was used to define the docking area by selecting a cubic grid box (X= 30, Y= 22, and Z= 40 at 0.375 Å) centered at the catalytic site of the enzyme. Finally, the most potent inhibitors in the prepared dataset were docked into the active site of GSK-3β using AutoDock vina [28], and the results were analyzed and visualized using PyMol (https://www. pymol.org) and Discovery Studio 2017 R2 (https://www.3dsbiovia.com) programs.

IV. Results and discussion

1. 2D-QSAR analysis

2D-QSAR analysis on C-Glycosylfavones was performed using MOE software. The MOE descriptors were used as independent variables (X variables) and the pIC_{50} values as dependent variables (Y variables). The correlation between the selected descriptors and pIC_{50} values was established by the PLS method. The final PLS model with its statistical parameters is given below:

$$\mathbf{pIC_{50}} = 0.175 - 10.10\ \mathbf{GCUT_SMR_2} + 0.175\ \mathbf{Kier1} + 1.074\ \mathbf{Vsurf_wp7} \quad \text{Eq.4}$$

$N = 23;\ r^2 = 0.85;\ r^2_{adjust} = 0.83;\ q^2 = 0.74;\ r^2_{test} = 0.94;\ MSE_{model} = 0.04;\ MSE_{LOO} = 0.15;\ F_{value} = 36.56;\ \alpha = 5\%$

where N is the training set compounds, r^2_{adjust} is the adjusted square correlation coefficient, MSE is the mean square error and F is the Fisher ratio.

As described in Eq.4, the best QSAR model was obtained using 3 descriptors. The collinearity of these descriptors was verified by the correlation matrix (Table 2) and the variance inflation +factor (VIF) values (Table 3), which can be calculated as $VIF = 1/(1-r^2)$, where r^2 is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. Generally, if the VIF value is greater than 5, multicollinearity is high. The values of the VIF and the matrix of correlation indicate that the selected descriptors are very significant at the threshold α = 5%.

Table 2. *The correlation matrix of the selected descriptors*

Variables	GCUT_SMR_2	Kier1	Vsurf_Wp7	pIC_{50}
GCUT_SMR_2	1,0000			
Kier1	0,0355	1,0000		
Vsurf_Wp7	-0,5033	0,2390	1,0000	
pIC_{50}	-0,4135	0,7172	0,7105	1,0000

Table 3. *The VIF value of the selected descriptors*

Statistique	GCUT_SMR_2	Kier1	Vsurf_Wp7
Tolerance	0,7210	0,9104	0,6806
VIF	1,3870	1,0984	1,4692

The generated QSAR model explains 85% (r^2 = 0.85) (Fig. 1a) of the total variance in the training set with a small value of error (MSE = 0.04) and an overall Fischer value of 36.56 at the threshold α = 5%. The predictive ability of the evaluated model was verified by the internal and external validations, which gave the correlation coefficient values of 0.74 and 0.94, respectively. The generated QSAR model was also validated by applying Y-randomization test. 100 random shuffles of the pIC50 values were performed. The average coefficients of R^2_m (R^2_m = 0.14) and Q^2_m (Q^2_m = -3.58) indicate that the results in the original model were not due to a chance correlation. Figure 2b, shows the plot of predicted versus observed activities for internal and external validations. The observed and predicted activity with residual of model is shown in Table 4.

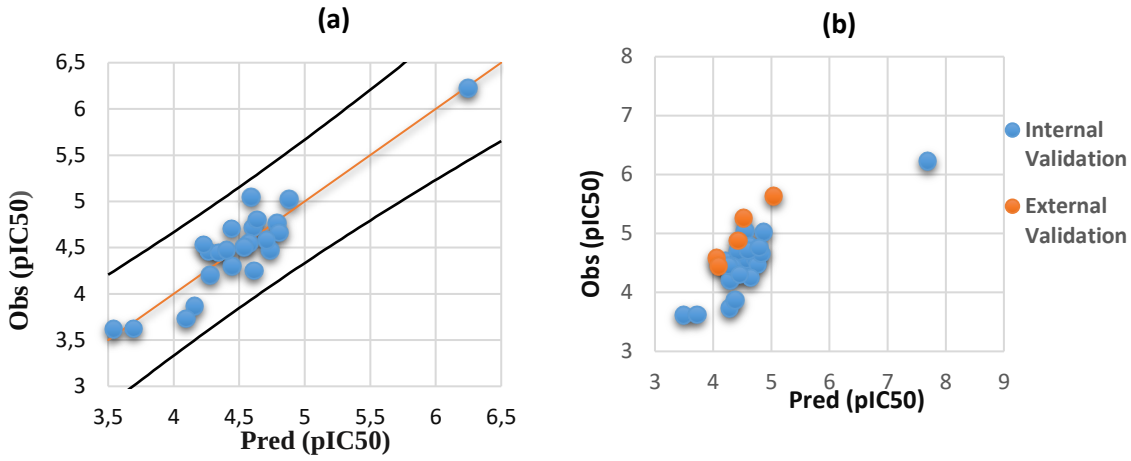


Fig. 1. *Plot of predicted versus observed activities.* ***(a)*** *the generated QSAR model,* ***(b)*** *the internal and external validations*

Table 4. *Predicted and observed with residual values for QSAR model and leave-one-out (LOO) cross validation*

cmpd No.	Obs_pIC$_{50}$	PLS		LOO	
		Pred_pIC$_{50}$	Residual	Pred_pIC$_{50}$	Residual
L1	3.7331	4.0940	-0.3610	4.2832	-0.5501
L2	3.6212	3.5392	0.0820	3.4885	0.1327
L3	3.6247	3.6925	-0.0678	3.7183	-0.0936
L4	3.8697	4.1604	-0.2907	4.3734	-0.5038
L5	4.5346	4.2292	0.3054	4.2091	0.3255
L7	4.4776	4.4012	0.0763	4.3974	0.0802
L8	4.5072	4.5369	-0.0297	4.5408	-0.0336
L9	4.5528	4.5734	-0.0205	4.5747	-0.0218
L10	5.0269	4.8798	0.1470	4.8647	0.1621
L12	4.2083	4.2771	-0.0688	4.2874	-0.0791
L14	5.0458	4.5927	0.4531	4.5436	0.5022
L15	4.5850	4.7059	-0.1209	4.7174	-0.1324
L16	4.4750	4.7348	-0.2599	4.7630	-0.2880
L17	4.7055	4.4406	0.2650	4.4224	0.2831
L18	4.8013	4.6341	0.1673	4.6202	0.1811
L19	4.4486	4.3428	0.1057	4.3215	0.1271
L21	4.2503	4.6113	-0.3611	4.6446	-0.3944
L22	4.6635	4.8047	-0.1412	4.8248	-0.1613
L23	4.4584	4.2641	0.1943	4.2521	0.2063
L24	4.7144	4.6080	0.1064	4.6009	0.1135
L25	4.3054	4.4451	-0.1398	4.4524	-0.1470
L26	4.7645	4.7895	-0.0251	4.7926	-0.0281
L27	6.2291	6.2455	-0.0163	7.6884	-1.4593

1.1.Interpretation of the selected descriptors

The model derived from PLS method indicates that the descriptors such as GCUT_SMR_2, Kier1 and Vsurf_wp7 have a significant influence on the activity of the studied compounds. From Eq.4 we can conclude that Kier1 and Vsurf_wp7 descriptors which represent the molecular shape and the polar surface volume respectively, showed positive effect impact on the value of pIC_{50} whereas GCUT_SMR_2 descriptor which represents the molar refractivity, showed the negative impact. In order to improve the activity of new predicted compounds, it is necessary to understand the meaning of each descriptor in Eq.4 above.

GCUT_SMR_2: The GCUT_SMR_2 is the descriptor that calculated from the eigenvalues of a modified graph distance adjacency [29]. The diagonal of the matrix takes the value of the atomic contribution to molar refractivity (using the Wildman and Crippen SMR method) [30]. The negative sign of the coefficient in the descriptor indicating that compounds with low values of GCUT_SMR_2 may increase the value of pIC_{50}.

Vsurf_wp7: The Vsurf_wp7 descriptor is a 3D descriptor describing the polar volume on the surface area. This descriptor has been shown to be useful in pharmacokinetic property prediction [31]. The positive sign of the coefficient in the descriptor suggesting that the molecule should have polar groups in the surface for interaction with the receptor or water molecules in the receptor.

Kier1: The Kier1 is the first kappa shape index defined by $(n-1)^2/m^2$ where n is the number of atoms in the hydrogen suppressed graph and m is the number of bonds in the hydrogen suppressed graph [32].

1.2.Applicability Domain

The applicability domain of the selected 2D-QSAR model was generated using Williams plot (Fig. 2), taken by MINITA.17 software (http://www.minitab.com), which shows leverage values versus standardized residuals. As describing in this Figure, all compounds are in the applicability domain, except compounds L27 and L4, which have a standardized residual and a lever value greater than $\pm 2\sigma$ and h* = 0.52, respectively.

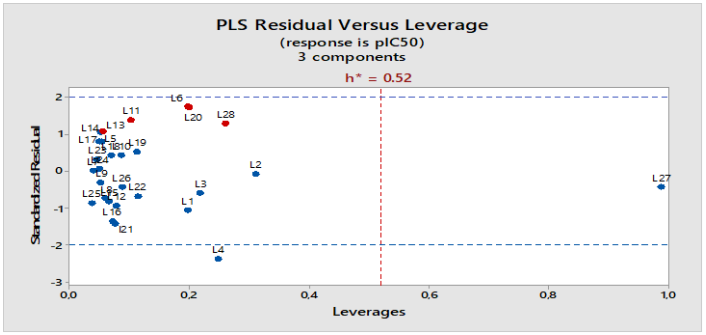


Fig. 2.*Williams plot for the developed 2D-QSAR model. Blue values represent training set compounds and red values represent test set compounds*

2. Molecular docking analysis

Molecular docking study play an important role in drug design approach, is used to predict the binding site of ligands to their active site and also to predict the binding affinity of ligands [33]. In this study molecular docking analysis was applied on the studied compounds. First, the reference ligand taken from the X-ray crystal structure was redocked into the active site of GSK-3β (PDB ID: 1PYX. The best obtained pose gave the root mean square deviation (RMSD) value of 1.29 Å using 50 iterations. The superposition between co-crystallized and redocked ligands as well as the binding site interactions of the native ligand in the active site of GSK-3β is shown in Figure 3.

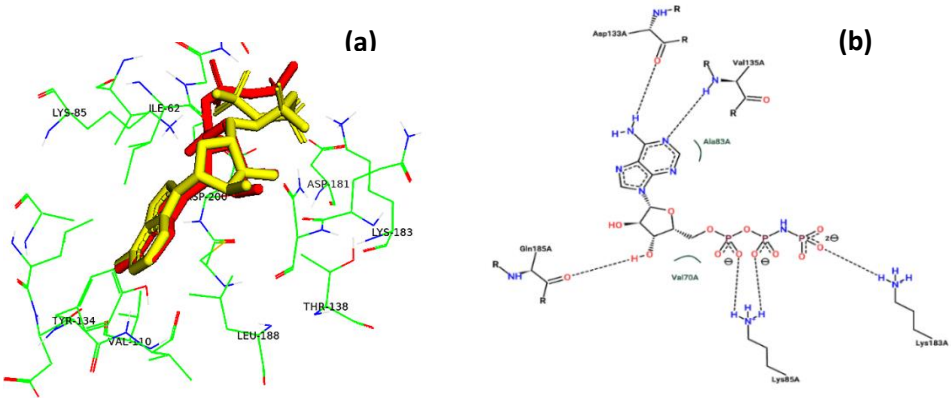


Fig. 3 (a). *Superposition of the co-crystalized ligand on the GSK-3β pocket (the yellow stick represents the re-docked ligand; the red stick represents the reference ligand). (b) PoseView 2D interaction diagram of the reference ligand*

After validation of docking results by re-docking the native ligand into the active site of GSK-3β, compounds L27 and L6 were used to generate the molecular docking studies. The result is shown in Figure 4.

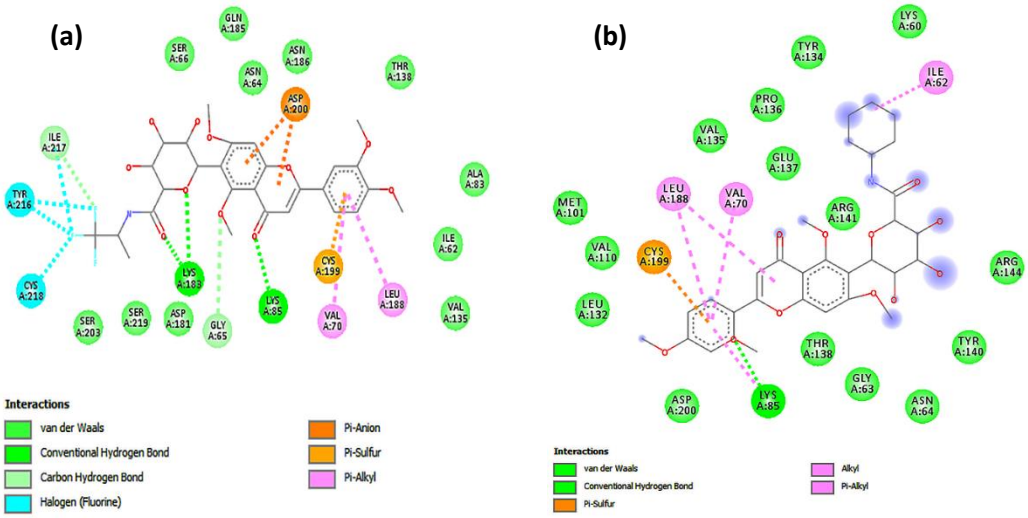


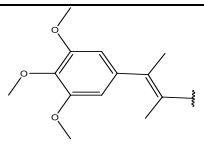
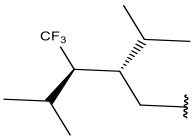
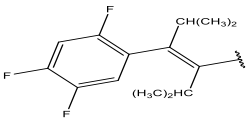
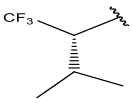
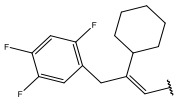
Fig. 4*. Docking results of compounds L27* ***(a)*** *and L6* ***(b)*** *with GSK-3β protein (PDB code: 1PYX)*

As described in Figure 4, compounds L27 and L6 show the same type of interactions. The most important are conventional hydrogen bond interactions, hydrophobic interactions (Alkyl and Pi-Alkyl), and Halogen interactions (Fluorine). Compound L27, the most potent inhibitor in dataset, was fixed via three hydrogen bonds with LYS 85 and LYS 183, four Halogen bonds with CYS 218, TYR 216 and ILE 217, two carbon hydrogen bonds with ILE 217 and GLY 65, two hydrophobic interactions with LEU 188 and VAL 70, two electrostatic interactions with ASP 200 and one pi-Sulfur interaction WITH CYS 199. These types of interactions may explain its low binding energy (-8.5 Kcal/mol) and consequently high pIC_{50} value (pIC_{50} = 6.2291).

3. Structural Design of new C-Glycosylfavone derivatives

QSAR and molecular docking studies not only provide information regarding structural features responsible for biological activity but it can also utilize for designing and predicting new potent compounds. Here, we based on the information obtained by 2D-QSAR model and molecular docking analysis, a novel series of compounds were proposed by modifying some structural features. These compounds were designed and optimized in the same way as that of the dataset compounds. The chemical structures and their predicted pIC_{50} values as well as the values of the binding free energy are presented in Table 5.

***Table** 5. Chemical structures and their predicted pIC_{50} and binding free energy values*

cmpd No.	Functional group				GSK-3β inhibition	
	R1	X1	X2	R2	pIC_{50}	Binding free Energy (Kcal/Mol)
T1	CH3	NH	O		6.617	-9.4
T2	CH3	NH	O		6.462	-8.5
T3	CH3	NH	O		6.417	-9.7
T4	CH3	CH-CH-$(CH_3)_2$	O		6.251	-9.0
T5	CH3	NH	O		6.267	-9.3

As describing in Table 5, all predicted compounds show a pIC_{50} value greater than the most potent inhibitor in the dataset (pIC_{50} (L27) = 6.229), imply our model is reliable and can be applied for designing new potent GSK-3β inhibitors. In order to confirm the high pIC_{50} values of predicted compounds, molecular docking approach using AutoDock Vina was used. The low values of binding free energy (Table 5) reflect the stability of predicted compounds in the binding pocket of GSK-3β protein. The 2D and 3D ligand-receptor interaction diagrams of compounds T1 and T3 are shown in Figures 5 and 6, respectively.

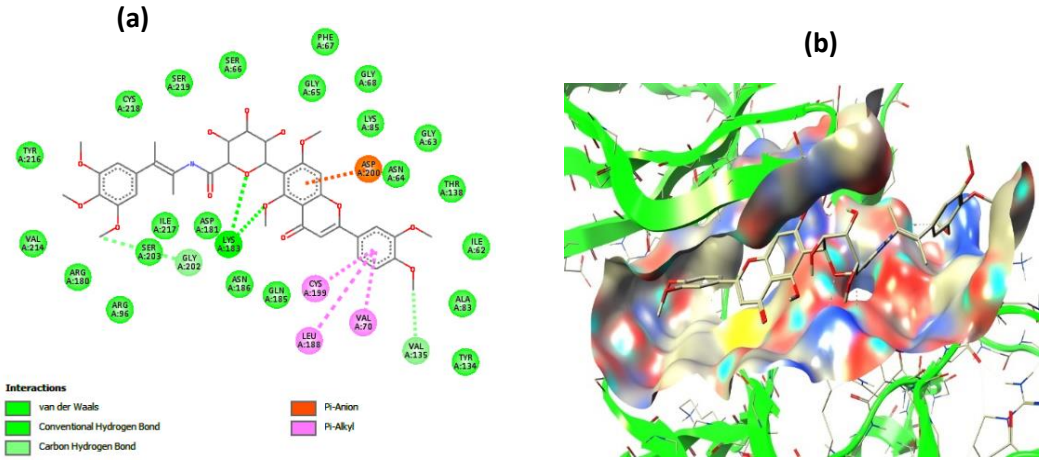


Fig. 5. *Molecular docking of compound T1: (a) 2D view of binding site interactions, (b) 3D view of the binding conformation*

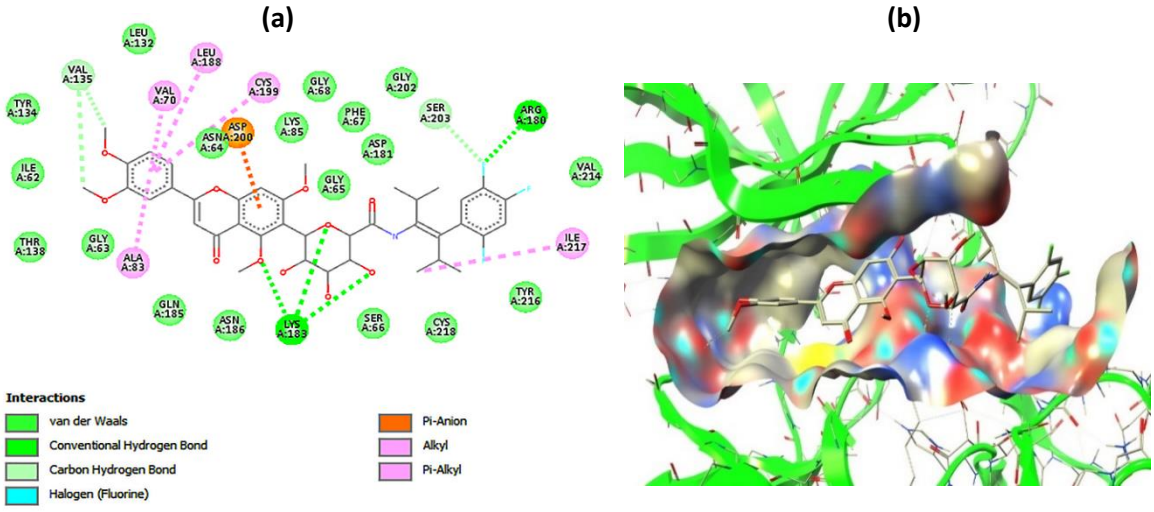


Fig. 6. *Molecular docking of compound T3:* ***(a)*** *2D view of binding site interactions,* ***(b)*** *3D view of the binding conformation*

From the result above we can conclude that, hydrogen bond interaction with Lys 85, Arg 180, Lys 183, Ser 203 and Ile 217 amino acids and hydrophobic interaction with Ile 62, Val 70, Ala 83, Val110, Leu132, Tyr 134, Leu 188 and Cys 199 amino acids play a significant role in the stability of ligand-receptor complexes.

V. Conclusion

In this paper, a theoretical study on a series of 2-(4-aryl-2-methylpiperazin-1-yl)-pyrimidin-4-one derivatives as GSK-3β inhibitors was carried out using 2D-QSAR study and docking analysis. The 2D-QSAR model was generated using MOE descriptors. PLS method was used to correlate the values of pIC_{50} with the values of calculated descriptors. The selected model revealed correlation coefficient (r^2) value of 0.85, the cross validated correlation coefficient (q^2) value of 0.74 and the predictive correlation coefficient (r^2_{test}) value of 0.94. Docking analysis using AutoDock Vina indicate that the most potent inhibitors in dataset exhibited strong binding affinity with the GSK-3β receptor. Based on the result of QSAR and docking analysis a novel series of C-Glycosylfavone derivatives were designed and their pIC_{50} values were predicted using 2D-QSAR model. These obtained results may guide synthetic medicinal chemists to design and synthesize new compounds with an increased biological activity.

Data Availability

The data of this work is available at https://github.com/elai-ssouq/GSK-3-beta

Acknowledgment

The authors want to thank the Moroccan Association of Theoretical Chemistry (MATC) for its relevant help regarding to the Software.

Conflict of Interest

The authors declared no conflict of interest

Compliance with Ethics Requirement

This article does not contain any study with human or animal subjects

References

1. Dokken, B.B., Henriksen, E.J.: Chronic selective glycogen synthase kinase-3 inhibition enhances glucose disposal and muscle insulin action in prediabetic obese Zucker rats. Am. J. Physiol. - Endocrinol. Metab. 291, (2006). https://doi.org/10.1152/ajpendo.00628.2005
2. Llorens-Martín, M., Jurado, J., Hernández, F., Ávila, J.: GSK-3β, a pivotal kinase in Alzheimer disease. Front. Mol. Neurosci. 7, 1–11 (2014). https://doi.org/10.3389/fnmol.2014.00046
3. Dugo, L., Collin, M., Thiemermann, C.: Glycogen synthase kinase 3β as a target for the therapy of shock and inflammation. Shock. 27, 113–123 (2007). https://doi.org/10.1097/01.shk.0000238059.23837.68
4. Woodgett, J.R.: Molecular cloning and expression of glycogen synthase kinase-3/factor A. EMBO J. 9, 2431–2438 (1990). https://doi.org/10.1002/j.1460-2075.1990.tb07419.x
5. Geometry, R., Analysis, G.: GLYCOGEN SYNTHASE KINASE 3 (GSK-3) AND ITS INHIBITORS.

John Wiley & Sons, Inc, Hoboken (2006)

6. Soutar, M.P.M., Kim, W.Y., Williamson, R., Peggie, M., Hastie, C.J., McLauchlan, H., Snider, W.D., Gordon-Weeks, P.R., Sutherland, C.: Evidence that glycogen synthase kinase-3 isoforms have distinct substrate preference in the brain. J. Neurochem. 115, 974–983 (2010). https://doi.org/10.1111/j.1471-4159.2010.06988.x

7. Busciglio, J., Lorenzo, A., Yeh, J., Yankner, B.A.: β-Amyloid fibrils induce tau phosphorylation and loss of microtubule binding. Neuron. 14, 879–888 (1995). https://doi.org/10.1016/0896-6273(95)90232-5

8. Acker, B.A., Jacobsen, E.J., Rogers, B.N., Wishka, D.G., Reitz, S.C., Piotrowski, D.W., Myers, J.K., Wolfe, M.L., Groppi, V.E., Thornburgh, B.A., Tinholt, P.M., Walters, R.R., Olson, B.A., Fitzgerald, L., Staton, B.A., Raub, T.J., Krause, M., Li, K.S., Hoffmann, W.E., Hajos, M., Hurst, R.S., Walker, D.P.: [2 , 3- c] pyridine-5-carboxamide as an agonist of the a 7 nicotinic acetylcholine receptor : In vitro and in vivo activity. 18, 3611–3615 (2008). https://doi.org/10.1016/j.bmcl.2008.04.070

9. Ryder, J., Su, Y., Liu, F., Li, B., Zhou, Y., Ni, B.: Divergent roles of GSK3 and CDK5 in APP processing. Biochem. Biophys. Res. Commun. 312, 922–929 (2003). https://doi.org/10.1016/j.bbrc.2003.11.014

10. Lovestone, S., Reynolds, C.H., Latimer, D., Davis, D.R., Anderton, B.H., Gallo, J., Hanger, D., Mulot, S., Marquardt, B., Stabel, S., Woodgett, J.R., Miller, C.C.J.: Alzheimer's disease-like phosphorylation of the microtubule-associated protein tau by glycogen synthase kinase-3 in transfected mammalian cells. . Curr Biol. 4, 1077–1086 (1994). https://doi.org/10.1016/S0960-9822(00)00246-3

11. Imahori, K.: PDH is inactivated by betaA-induced TPKI-GSK-3beta. Proc. Natl. Acad. 93, 2719–2723 (1996). https://doi.org/10.1073/pnas.93.7.2719

12. Pei JJ, Tanaka T, Tung YC, Braak E, Iqbal K, G.-I.I.: Distribution, levels, and Activity of Glycogen Synthase Kinase-3 in the Alzheimer Disease Brain. J. Neuropathol. Exp. Neurol. 56, 70–78 (1997). https://doi.org/10.1097/00005072-199701000-00007

13. Hardy, J.: A Hundred Years of Alzheimer's Disease Research. Neuron. 52, 3–13 (2006). https://doi.org/10.1016/j.neuron.2006.09.016

14. Leroy, K., Yilmaz, Z., Brion, J.: Increased level of active GSK-3ß in Alzheimer's disease and accumulation in argyrophilic grains and in neurons at different stages of neurofibrillary degeneration. Neuropathol. Appl. Neurobiol. 33, 43–55 (2007). https://doi.org/10.1111/j.1365-2990.2006.00795.x

15. Liang, Z., Li, Q.X.: Discovery of Selective, Substrate-Competitive, and Passive Membrane Permeable Glycogen Synthase Kinase-3 β Inhibitors: Synthesis, Biological Evaluation, and Molecular Modeling of New C - Glycosyl fl avones. ACS Chem. Neurosci. 9, 1166–1183 (2018). https://doi.org/10.1021/acschemneuro.8b00010

16. El Aissouq, A., Toufik, H.: QSAR study of isonicotinamides derivatives as Alzheimr ' s disease inhibitors using PLS-R and ANN methods. 2019 Int. Conf. Intell. Syst. Adv. Comput. Sci. 1–7 (2019). https://doi.org/10.1109/ISACS48493.2019.9068919

17. Stitou, M., Toufik, H., Bouachrine, M., Lamchouri, F.: Quantitative structure–activity relationships analysis, homology modeling, docking and molecular dynamics studies of triterpenoid saponins as Kirsten rat sarcoma inhibitors. J. Biomol. Struct. Dyn. 1102, (2020). https://doi.org/10.1080/07391102.2019.1707122

18. Tropsha, A., Gramatica, P., Gombar, V.K.: The Importance of Being Earnest : Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. Mol. Inform. 22, 69–77 (2003). https://doi.org/https://doi.org/10.1002/qsar.200390007

19. Raghu, T.S., Ramesh, R., Whinston, A.B.: Addressing the homeland security problem: A collaborative decision-making framework. J. Am. Soc. Inf. Sci. Technol. 56, 310–324 (2005). https://doi.org/10.1002/asi.20123

20. El Aissouq, A., Toufik, H., Stitou, M., Ouammou, A., Lamchouri, F.: In Silico Design of Novel Tetra-

Substituted Pyridinylimidazoles Derivatives as c-Jun N-Terminal Kinase-3 Inhibitors, Using 2D/3D-QSAR Studies, Molecular Docking and ADMET Prediction. Int. J. Pept. Res. Ther. 26, 1335–1351 (2020). https://doi.org/10.1007/s10989-019-09939-8

21. Elmchichi, L., Belhassan, A., Lakhlifi, T., Bouachrine, M.: 3D-QSAR Study of the Chalcone Derivatives as Anticancer Agents. J. Chem. 2020, (2020). https://doi.org/10.1155/2020/5268985

22. Kubinyi, H., Hamprecht, F.A., Mietzner, T.: Three-dimensional quantitative similarity-activity relationships (3D QSiAR) from SEAL similarity matrices. J. Med. Chem. 41, 2553–2564 (1998). https://doi.org/10.1021/jm970732a

23. Kasmi, R., Hadaji, E., Chedadi, O., El Aissouq, A., Bouachrine, M., Ouammou, A.: of CDK (anticancer activity) with an application of the molecular docking method. Heliyon. 6, e04514 (2020). https://doi.org/10.1016/j.heliyon.2020.e04514

24. Goudzal, A., Aissouq, A. El, Hamdani, H. El, Ouammou, A.: Materials Today : Proceedings QSAR modeling , molecular docking studies and ADMET prediction on a series of phenylaminopyrimidine- (thio) urea derivatives as CK2 inhibitors. Mater. Today Proc. (2020). https://doi.org/10.1016/j.matpr.2020.08.044

25. Gramatica, P.: Principles of QSAR models validation: Internal and external. QSAR Comb. Sci. 26, 694–701 (2007). https://doi.org/10.1002/qsar.200610151

26. Bertrand, J.A., Thieffine, S., Vulpetti, A., Cristiani, C., Valsasina, B., Knapp, S., Kalisz, H.M., Flocco, M.: Structural Characterization of the GSK-3 b Active Site Using Selective and Non-selective ATP-mimetic Inhibitors. 393–407 (2003). https://doi.org/10.1016/j.jmb.2003.08.031

27. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., Olson, A.J.: Software News and Updates AutoDock4 and AutoDockTools4 : Automated Docking with Selective Receptor Flexibility. J. Comput. Chem. 30, 2785–2791 (2009). https://doi.org/10.1002/jcc

28. Trott, O., Olson, A.J.: Software News and Update AutoDock Vina : Improving the Speed and Accuracy of Docking with a New Scoring Function , Efficient Optimization , and Multithreading. J. Comput. Chem. 31, 455–461 (2009). https://doi.org/10.1002/jcc

29. Petitjean, M.: Applications of the Radius-Diameter Diagram to the Classification of Topological and Geometrical Shapes of Chemical Compounds. J. Chem. Inf. Comput. Sci. 32, 331–337 (1992). https://doi.org/10.1021/ci00008a012

30. Wildman, S.A., Crippen, G.M.: Prediction of physicochemical parameters by atomic contributions. J. Chem. Inf. Comput. Sci. 39, 868–873 (1999). https://doi.org/10.1021/ci9903071

31. Hari, N.S., Moorthy, N., Ramos, M.J., Fernandes, P.A.: Prediction of the relationship between the structural f - eatures of andrographolide derivatives and α -glucosidase inhibitory activity : A quantitative structure-activity - relationship (QSAR) Study. J. Enzyme Inhib. Med. Chem. 26, 78–87 (2011). https://doi.org/10.3109/14756361003724760

32. Hall, L.H., Mohney, B.: The Electrotopological State : An Atom Index for QSAR '. Mol. Inform. 51, 43–51 (1991). https://doi.org/https://doi.org/10.1002/qsar.19910100108

33. Kitchen, D.B., Decornez, H., Furr, J.R., Bajorath, J.: DOCKING AND SCORING IN VIRTUAL SCREENING FOR DRUG DISCOVERY : METHODS AND APPLICATIONS. 3, (2004). https://doi.org/10.1038/nrd1549

Chapitre V: Application des méthodes CADD à la recherche et développement de nouveaux inhibiteurs de la p38a MAP kinase : 2D-QSAR, docking moléculaire, dynamique moléculaire et propriétés ADMET

Résumé :

La protéine p38 kinase est une enzyme appartenant à la famille des kinases MAP (Mitogen-Activated Protein Kinse). Elle joue un rôle important dans la régulation de divers processus biologiques, notamment la réponse cellulaire au stress, l'inflammation est la mort cellulaire. Des anomalies de la régulation de la protéine p38 kinase ont été associées à de nombreuses maladies humaines, notamment le cancer, les maladies inflammatoires chroniques telles que la polyarthrite rhumatoïde et la maladie de Crohn, ainsi que les maladies neurodégénératives telles que la maladie d'Alzheimer et la maladie de Parkinson. Dans ce chapitre, une série de molécules, dérivées du pyridinyl-imidazole, en tant qu'inhibiteurs de la protéine p38α Kinase, a été sélectionnée pour réaliser les études 2D-QSAR, de docking moléculaire et de la dynamique moléculaire. L'objectif est de trouver une corrélation entre la variation de l'activité biologique de cette série de molécules et les descripteurs moléculaires calculés par le logiciel MOE et par conséquent, la prédiction de nouvelles molécules de même famille. Le modèle 2D-QSAR a été élaboré à l'aide de la méthode de régression des moindres carrés partiels (PLS). Les résultats ont révélé l'importance de descripteurs PEOE_RPC+ et SMR_VSA2 dans la variation de l'activité biologique. Le modèle 2D-QSAR développé a révélé une valeur de coefficient de détermination (r^2) de 0.85, une valeur du coefficient de détermination ajusté (r_a^2) de 0.83 et une valeur d'erreur quadratique moyenne (RMSE) de 0.16. Les tests de la validation interne et externe ont montré que le modèle développé a un pouvoir prédictif élevé (r_{cv}^2=0.78 et r_{test}^2=0.92). L'affinité des ligands (dérivés du pyridinyl imidazole) dans le site de liaison du récepteur MAP kinase p38α a également été déterminée par l'étude de docking moléculaire en utilisant le logiciel autoDock. Les résultats ont révélé l'importance des liaisons hydrogène dans la stabilité des ligands dans le site de liaison du récepteur MAPK p38 α. À l'aide de ces résultats, une nouvelle série de même famille de composés a été prédite. L'étude pharmaceutique par l'approche ADMET a montré que ces molécules ont des propriétés pharmaceutiques élevées. L'analyse par dynamique moléculaire montre que ces molécules prédites sont stables dans le site de liaison du récepteur MAPK p38α. Les résultats de l'étude pourraient ouvrir la voie au développement de nouveaux dérivés du pyridinyl-imidazoles ayant la capacité d'inhiber l'activité enzymatique de la MAP kinase p38a.

NJC

PAPER

Check for updates

Cite this: *New J. Chem.*, 2022, 46, 20786

Computer-aided drug design applied to a series of pyridinyl imidazole derivatives targeting p38a MAP kinase: 2D-QSAR, docking, MD simulation, and ADMET investigation

Abdellah El Aissouq*[a], Aissam Lachhab[a], Said El Rhabori [a], Mohammed Bouachrine[c], Abdelkrim Ouammou[b] and Fouad Khalil [a]

Received 25th July 2022,

Accepted 5th October 2022 DOI: 10.1039/d2nj03686j

rsc.li/njc

The p38 mitogen-activated protein kinase (MAPK) is a crucial target for chronic inflammation. Here, 2D-QSAR and molecular docking studies were performed on a new series of pyridinyl imidazole derivatives as p38a MAPK inhibitors. The goal is to find a correlation between the chemical structures and inhibitory activity of the p38a MAPK. The partial least squares (PLS) regression method was used to build the 2D-QSAR model. The results revealed that the generated model yielded 0.85 for the determination coefficient (r^2), 0.83 for the adjusted determination coefficient (r_a^2), and 0.16 for the root-mean-square error (RMSE). Internal and external validations were used to figure out how well the model could predict the future. The r_{cv}^2 and r_{test}^2 values were 0.78 and 0.92, indicating that the developed model is robust and capable of predicting the activity of new compounds. The affinity of the ligands (pyridinyl imidazole derivatives) and the binding pocket of the p38a MAP kinase receptor was also determined by molecular docking using autoDock tools. The results revealed the importance of conventional hydrogen bonds in the binding pocket of the p38a MAPK receptor. Based on these results, a novel series of compounds were predicted and evaluated using molecular dynamics (MD) and ADMET prediction investigations. The study's findings might pave the way for the development of new pyridinyl imidazole derivatives that are able to inhibit the enzymatic activity of the p38a MAP kinase.

1. Introduction

Alzheimer's disease (AD) is a progressive brain disease that causes people to lose their memories. Amyloid deposits and neurofibrillary tangles are two of the pathological features of the disease; the underlying cause has been intensively studied for several decades, but the etiology is poorly understood. The notion of inflammation and oxidative stress is one of the theories suggested to explain the pathology of Alzheimer's disease [1]. However, it is unclear how early and where in the pathogenic cascade these processes occur. Several extracellular stimuli are transformed into specific cellular responses by activating mitogen-activated protein kinase (MAPK) signaling pathways. Among MAPK family members, the p38α MAPK subfamilies are preferentially activated by cellular stresses, such as those induced by proinflammatory cytokines [2, 3]. Recent studies on the underlying biology of AD in cellular and animal models have indicated that p38α MAPK can orchestrate diverse events related to AD, such as neurotoxicity, tau phosphorylation, synaptic dysfunction, and neuroinflammation [4, 5]. Thus, the inhibition of p38α MAPK is considered a promising strategy for treating AD. Among the most widely pursued drug target class, the protein kinases 3, p38α MAPK represents an intensively studied kinase, for which numerous inhibitors have been developed to date, but still no approved drug exists [6].

The discovery of a safe and efficacious drug is a complex, time-consuming, and expensive process. Computational methodologies driven by cheminformatics tools play a central role in the high-throughput lead discovery and optimization process especially when the information about the ligand (ligand-based) or structure of the biological target (structure-based) is available. Here, a novel series of 37 pyridinylimidazole derivatives was used to perform ligand-based drug design (LBDD) by applying the method of two-dimensional quantitative structure-activity relationship (2D-QSAR). The selected compounds are known as p38 MAPK inhibitors, and their IC_{50} values range from 3 µM to 1603 µM [7]. The 2D-QSAR model was developed using the partial least squares (PLS) regression method [8], employing the inhibitory activity of p38α MAPK as a dependent variable and molecular descriptors (calculated on Molecular Operating Environment (MOE) software) as independent variables. The internal and external validations of the developed model were verified by a cross-validation process and a set of 6 compounds, respectively. The methods of structure-based drug design (SBDD) including molecular docking and molecular dynamics (MD) simulation were carried out to derive more

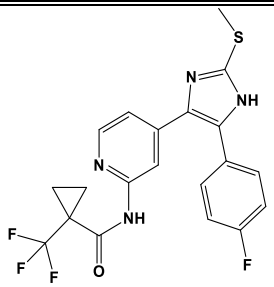
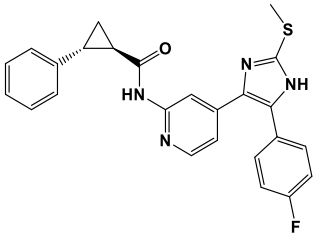
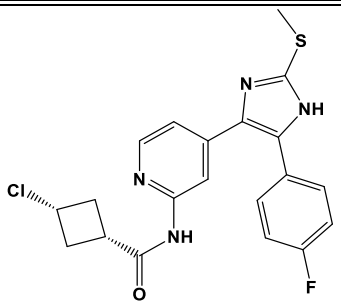
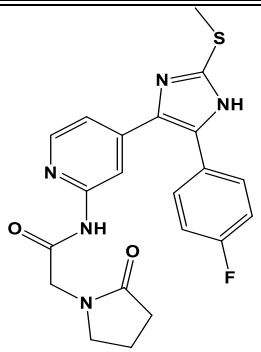
insights on protein-ligand interaction along with stability profile of protein-ligand complexes[9, 10]. In addition, the pharmacokinetic proprieties of our studied compounds were evaluated using absorption, distribution, metabolism, excretion, and toxicity (ADMET) investigation. The objective is to develop new pyridinylimidazole derivatives capable of inhibiting the enzymatic activity of p38α MAP kinase.

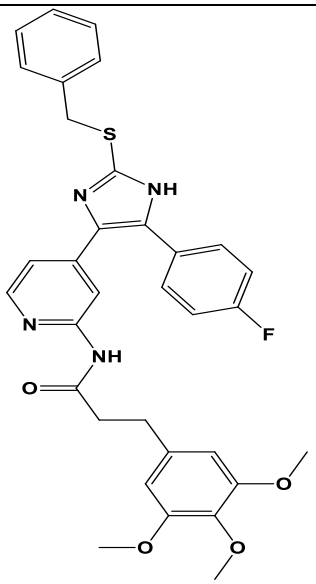
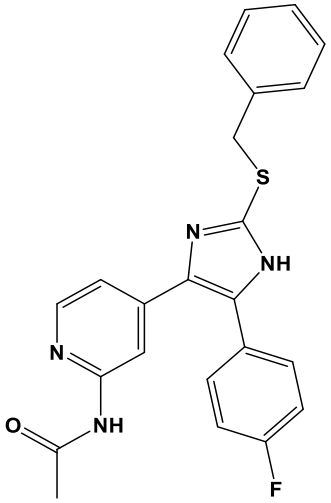
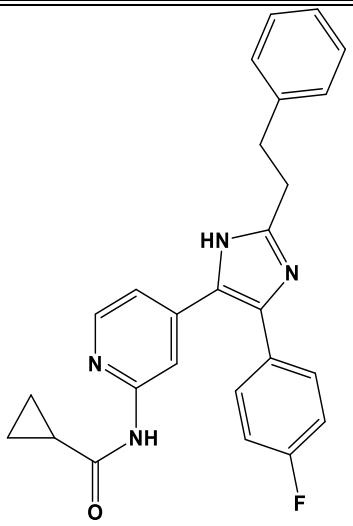
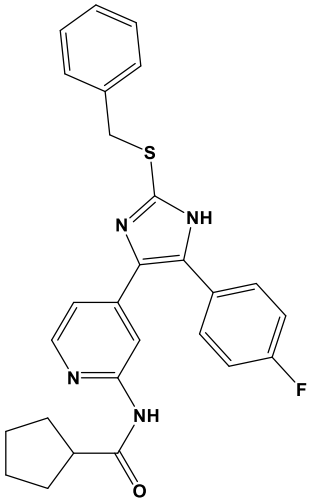
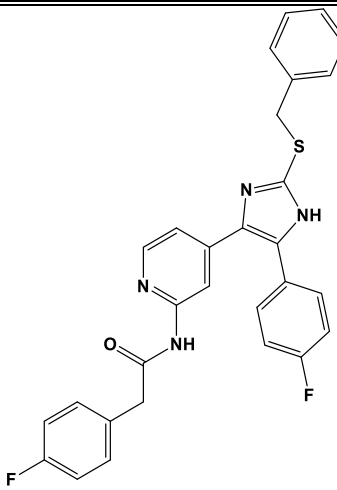
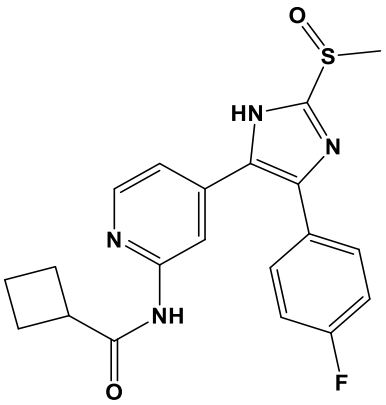
II. Materials and methods

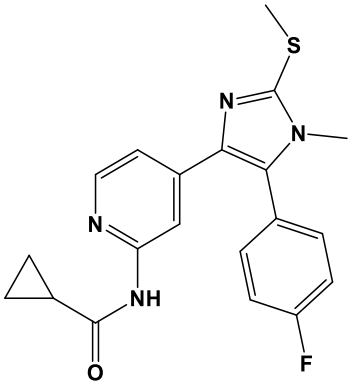
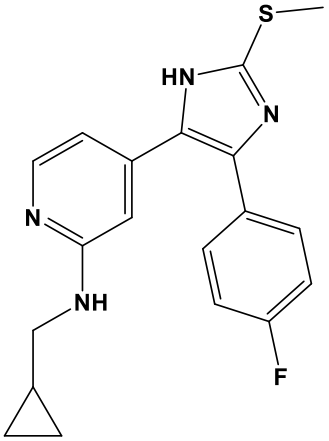
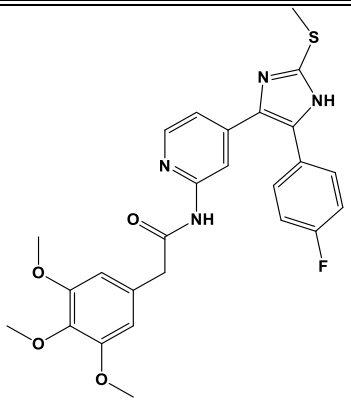
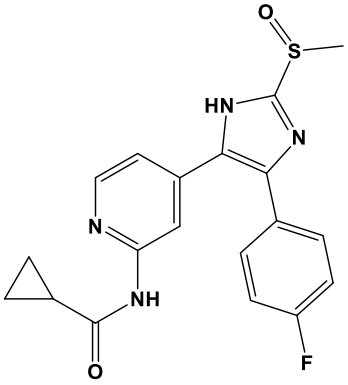
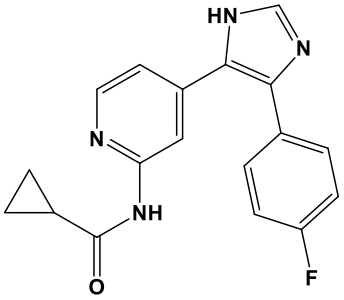
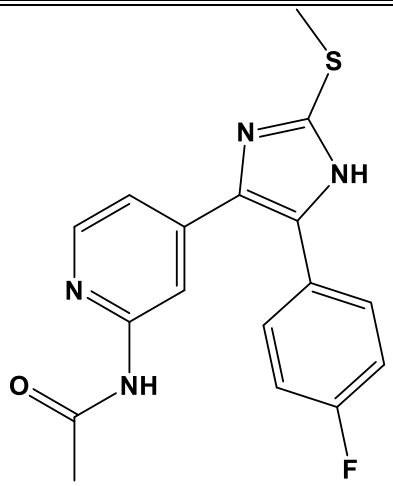
1. Data set collection and optimization

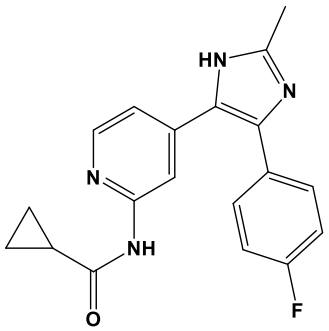
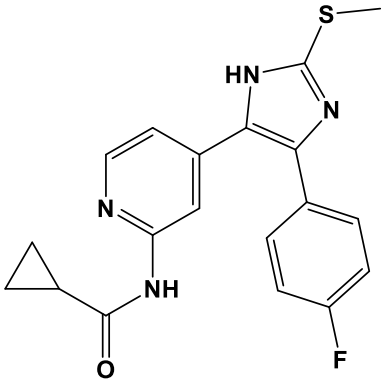
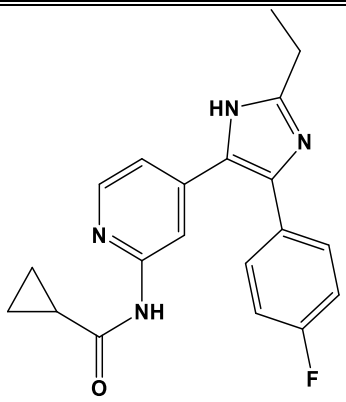
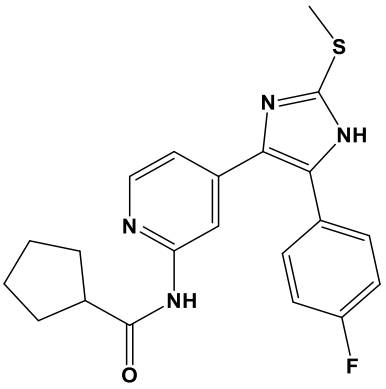
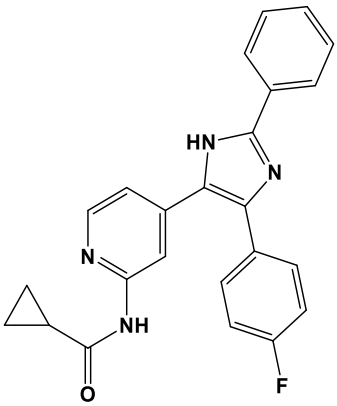
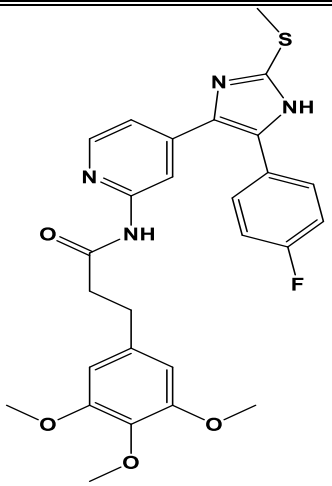
In this work, a dataset of 37 Pyridinylimidazole derivatives (**Table 1**) was collected to perform computer-aided drug design (CADD) approaches [7]. The selected compounds are known as p38α MAPK inhibitors and their IC_{50} values range from 3 µM to 1603 µM. The IC_{50} values in the µM range were transformed to molar (M) values, and their negative logarithm (pIC_{50} = - log (IC_{50})) was used as a dependent variable in future 2D-QSAR study. The dataset was then randomly split into a training set and a test set, with 31 compounds in the training set and 6 in the test set.

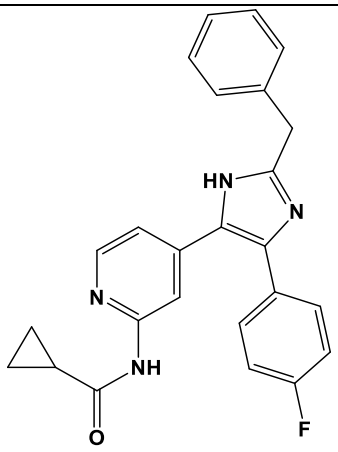
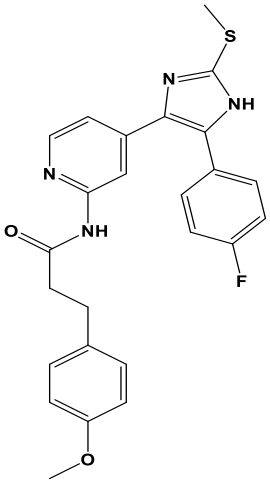
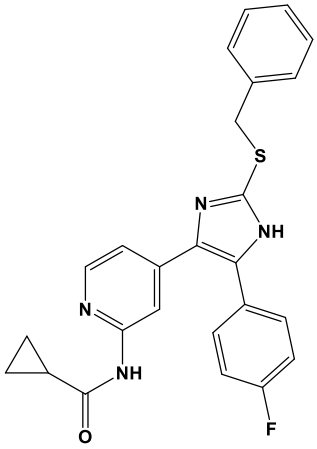
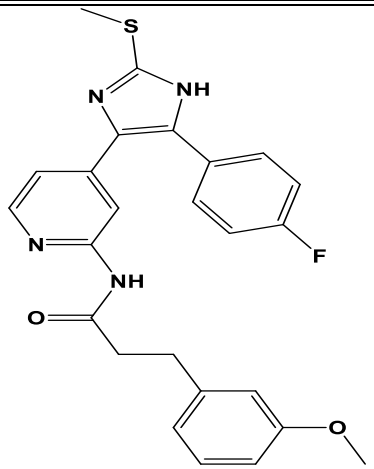
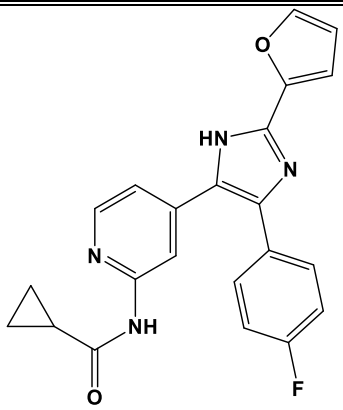
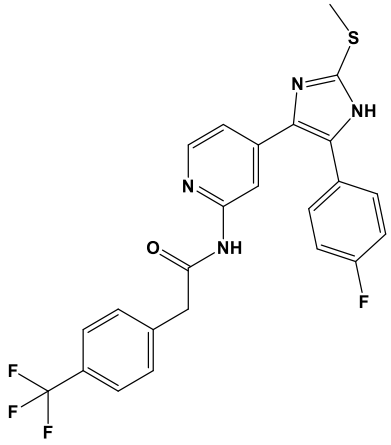
Table 1. *Structures of chemicals and their corresponding activity values (* represents the test set compounds)*

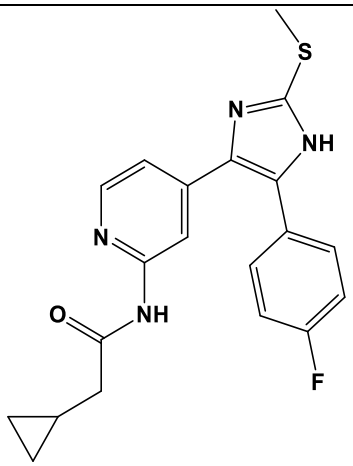
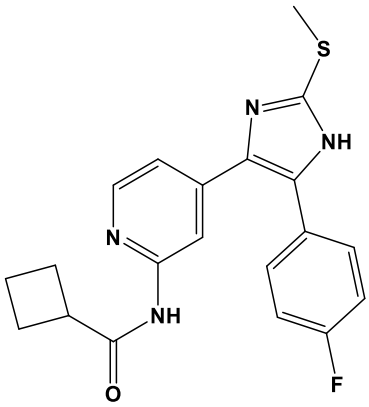
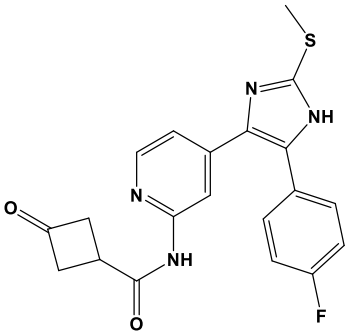
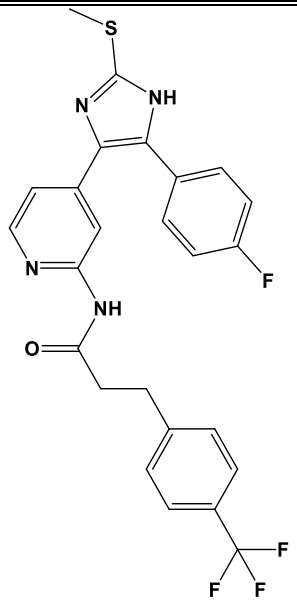
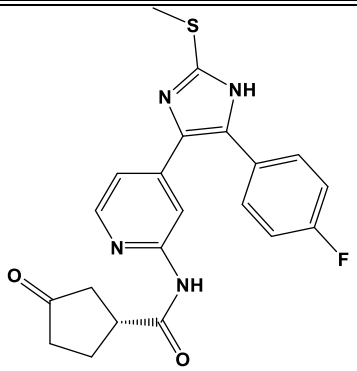
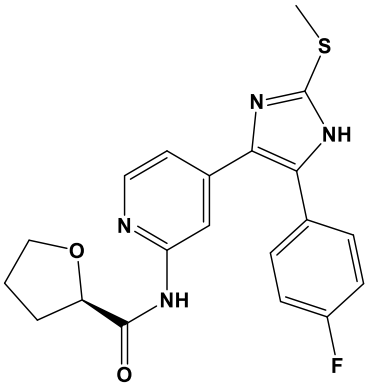
No	Chemical structures	OBS_p IC_{50}	No	Chemical structures	OBS_p IC_{50}
L1*		6.7352	L20		7.8239
L2		7.7447	L21		7.6383

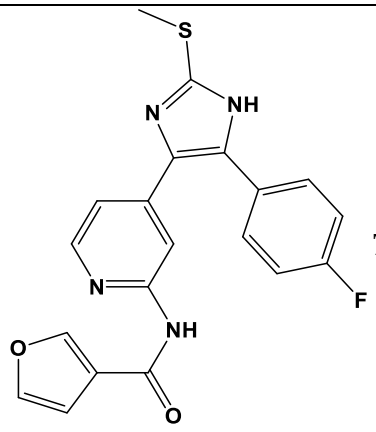
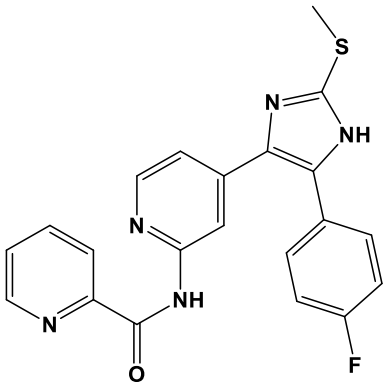
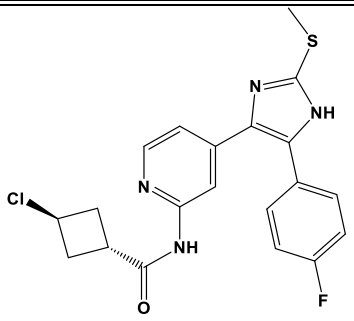
L3*		7.2147	L22		7.2924
L4*		7.3872	L23*		7.6990
L5		8.5229	L24		7.5376

L6		5.7951	L25		7.8539
L7*		7.0555	L26		7.3565
L8		7.7212	L27		7.3468

L9		7.7212	L28		7.6198
L10		7.7959	L29*		7.3979
L11		7.5686	L30		7.8861

L12		7.7447	L31		7.2596
L13		7.6990	L32		7.4437
L14		7.7447	L33		7.5376

L15		7.8239	L34		7.5686
L16		7.6990	L35		7.5686
L17		7.6778	L36		7.5528

L18		7.6576	L37		7.3768
L19		7.4202			

Before modeling, the chemical structures of all compounds were built using ChemBio3D Ultra 14.0 software (https ://chema xon.com). Then all compounds were energy-minimized using molecular mechanics-2 (MM2) force field [11]. The optimized structures were then used in future studies.

2. 2D-QSAR model

2.1.Molecular descriptors

Calculating molecular descriptors is an essential step in QSAR research since they quantitatively represent the chemical information associated with compounds [12]. In this study, a set of 354 descriptors was generated using molecular operating environment (MOE) software (http:// www. chemcomp. com). The descriptors with identical or null values were eliminated, and the remaining ones were chosen. To reduce the number of descriptors, those with low correlation coefficients with activity were omitted. The best-selected descriptors were used as independent variables in the QSAR study.

2.2.2D-QSAR model development

The 2D-QSAR model was constricted using the PLS regression method. The best-selected descriptors were utilized as X variables, while p38 MAP kinase inhibitory activity was used as a Y variable. The equation that relates both variables (X and Y variables) is given below:

$$\mathbf{pIC_{50} = a_0 + a_1 D_1 + a_2 D_2 + \ldots\ldots + a_n D_n} \qquad \textbf{Eq.1}$$

The statistical significance of the QSAR model was determined by calculating the squared correlation coefficient (r^2), the adjusted coefficient of determination (r_a^2), the root-mean-square error (RMSE), and the Fischer ratio value (F value).

$$\mathbf{R^2 = 1 - \left[\frac{\sum_i (Y_i\ obs - Y_i pred)^2}{\sum_i (Y_i\ obs - \bar{Y}_i\ mean)^2}\right] = 1 - \frac{SSE}{SST}} \qquad \textbf{Eq.2}$$

$$\mathbf{R^2_{Adjusted} = \frac{(n-1)\, r^2 - p}{n-1-p}} \qquad \textbf{Eq.3}$$

$$\mathbf{RMSE = \sqrt{\frac{\sum_{I=1}^{n} (yi\ obs - yi\ pred)}{n}}} \qquad \textbf{Eq.4}$$

$$\mathbf{F = \frac{\frac{\sum_i (Y_i\ pred - \bar{Y}_i\ mean)^2}{p}}{\frac{\sum_i (Y_i\ obs - Y_i\ pred)^2}{n-p-1}}} \qquad \textbf{Eq.5}$$

In the equations above, Y_i obs and Y_i pred are the observed and predicted activities for the i^{th} compound in the training set. $\bar{Y}$mean is the average activity of the compounds in the training set, SSE and SST are the residual of squares and the total sum of squares, respectively, n is the total number of compounds in the training set and p is the number of descriptors. The value of r^2 should be close to one, the value of RMSE should be close to zero, and the value of F should be high for a successful model. The model that is developed is used to understand how the activity varies when any of the descriptors (D_n) is changed.

2.3.2D-QSAR Model Validation

The robustness of the developed model was evaluated through internal and external validations [13, 14]. The internal validation was performed using the leave-one-out cross-validation method. In this process, each compound in the training set was eliminated, and the remaining dataset (n-1) was used to develop a model and predict the target value of the compound that had been removed. The quality of the model is evaluated by a high value of r_{cv}^2. The external validation was performed on a set of 6 compounds that were not used in the training set. The predicted ability of the devolved model was verified by a high value of $r^2_{\ test}$.

2.4. Applicability domain

The applicability domain is an area in chemical space that contains both the model variables and the predicted response[15]. The applicability domain of our studied compounds was determined using the leverage value (hi) of each compound i, for which a QSAR model was used to predict its activity:

$$h_i = x_i^T (X^T X)^{-1} x_i \qquad (i = 1. \ldots. n). \qquad \text{Eq.6}$$

Where xi is the descriptor row-vector for a query compound and X is the n*(k-1) matrix with k descriptor values for n compounds in the data collection. When a compound's leverage value (hi) is higher than the critical value (h*) (h* = 3(d+1)/n, where d is the number of variables and n is the number of compounds in the training set), the compound is considered to be outside the applicability domain. On the other hand, a compound i was thought to be in the applicability domain when hi is less than h*.

3. Molecular docking analysis

Molecular docking was carried out using Autodock tools. The human p38α MAPK in complex with imidazolyl pyridine inhibitor 11b was downloaded from the protein data bank (PDB ID : 5ML5) [16]. The ligand was separated from the protein. The water molecules were taken out of the protein, and the polar hydrogen atoms and Kollman charges were added. The final file was minimized by UCSF Chimera software and saved in pdbqt format for further analysis. Then, a grid box was positioned to include the active site of the crystal structure with the following dimensions in: center (X, Y, Z) = (21.423, 2,502, -21.989) and dimensions (X, Y, Z) = (X, Y, Z (40, 42, 40). Then, the AutoDock 4.2 was used to dock the most powerful inhibitors from the prepared dataset into the active site of the p38α MAP kinase receptor. The results were analyzed using Discovery Studio 2017 R2 software (https://www.3dsbiovia.c).

4. MD simulation

MD simulation was performed using the GROMACS 5.1.4 package [17] with the CHARMM36 all-atom force field [18]. The topology parameters of ligands were generated using the General Force Field (CGenFF) web server [19]. The ligand and protein complex were put in a cubic periodic box with a minimum distance of 2 nm from the box's sides and filled with a TIP3P water model [35]. Each system was then neutralized by adding the appropriate concentration of Na^+. The systems were equilibrated using NVT and NPT ensembles for 100 ps [20]. After equilibration, production runs of 50 ns were performed at a constant temperature of 300 K (V-rescale method) and pressure of 1.01325 (Parrinello-Rahman method) using the Leapfrog algorithm [21]. Using the built-in functions of the GROMACS 5.1.4 software, different parameters were calculated based on the dynamics trajectory [22]. These parameters included the

root mean square deviation (RMSD), radius of gyration (Rg), and the number of hydrogen bonds (HB) between the protein and the ligand.

5. Prediction of ADMET properties

The predicted compounds as well as the most active compound in the dataset were subjected to ADMET prediction using the PreADMET web server (https://preadmet.webservice.bmdrc.org/). First, the coordinates of the compound were copied as mol format into the server's input page and submitted for ADMET parameters calculation. The predicted parameters of ADMET include blood brain permeability, aqueous solubility, human intestinal absorption, Ames test, and rodent carcinogenicity were used to evaluate the pharmacokinetic proprieties of a given compound [23].

6. Results and discussion

2D-QSAR Study

PLS analysis

In the first step of the study, the linear 2D-QSAR model was constructed using the PLS method. Two factors were used to determine the optimal model. The PLS equation for the data set is given below:

$$\mathbf{pIC_{50}} = 5.854 - 10.506 \times \mathbf{PEOE_RPC^{+}} + 0.099 \times \mathbf{SMR_VSA2} \quad \text{Eq.7}$$

$N = 31$, $RMSE = 0.16$, $r^2 = 0.85$, $r_a^2 = 0.83$, $F_{model} = 77.04$, $p < 0.001$, $\alpha = 5\%$, $r^2_{cv} = 0.77$, $RMSE_{CV} = 0.19$ and $r^2_{test} = 0.92$

N = the number of modeled responses

As shown in Eq. 7, the best 2D-QSAR model was built by combining two descriptors (PEOE RPC^{+} and SMR VSA2) that have a low correlation with each other and a high correlation with pIC_{50} (**Table 2**). The occurrence of collinearity between these descriptors was further confirmed by computing the variance inflation factor (VIF) values (**Table 3**), which are calculated as VIF = 1/ (1 r^2), where r is the multiple correlation coefficient of one descriptor's effect regressed on the remaining molecular descriptors. The VIF value of a descriptor should be less than 5. As seen in **Table 3**, all descriptors have a VIF value less of than 5, suggesting that they are very significant at the threshold $\alpha = 1\%$.

Table 2. *Matrix of Correlation of the selected descriptors*

Variables	PEOE_RPC+	SMR_VSA2	pIC_{50}
PEOE_RPC+	1.0000		
SMR_VSA2	0.2336	1.0000	
pIC_{50}	-0.2425	0.8061	1.0000

Table 3. *VIF values of selected descriptors*

Statistique	PEOE_RPC+	SMR_VSA2
Tolérance	0.9454	0.9454
VIF	1.0577	1.0577

The developed 2D-QSAR model explains 85% (**Figure 1**) of the total variance in the training set with a small value of error (RMSE = 0.16) and a Fisher value of 77.04. Internal validation by the leave-one-out cross-validation (LOOCV) process yielded 0.77 for the determination coefficient (r^2_{cv}) and 0.19 for the root-mean-square error (RMSE). External validation by a set of 6 compounds yielded 0.92 of the determination coefficient (r^2_{test}). **Table 4** shows the values of some of the descriptors in Eq. 7 as well as the predicted activity and residual values.

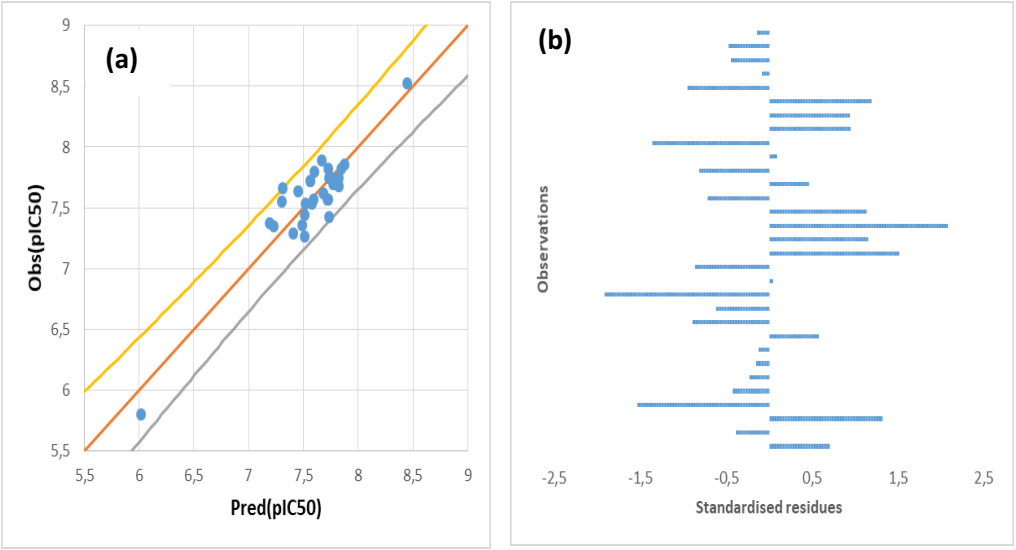


Figure 1*. (a) Plot of predicted activities versus experimental ones, (b) standardisd residues of training set compounds*

Table 4. *The value of the observed activity, descriptors, predicted activity, and residual*

	NO	pIC_{50}	PEOE_RPC+	SMR_VSA2	pIC_{50}			
					pIC_{50}_PRED	Residual	pIC_{50}_PRED_CV	Residual
Training set	L27	7.3468	0.1188	26.5996	7.2300	0.1168	7.2223	0.1245
	L28	7.6198	0.1170	31.0104	7.6835	-0.0637	7.6862	-0.0664
	L30	7.8861	0.0770	26.5996	7.6691	0.2170	7.6264	0.2596
	L31	7.2596	0.0918	26.5996	7.5131	-0.2535	7.5360	-0.2763
	L32	7.4437	0.0918	26.5996	7.5132	-0.0695	7.5194	-0.0757
	L33	7.5376	0.1588	34.3591	7.5753	-0.0377	7.5904	-0.0528
	L35	7.5686	0.1570	34.3591	7.5936	-0.0250	7.6029	-0.0343
	L20	7.8239	0.1018	31.0104	7.8440	-0.0201	7.8452	-0.0213
	L15	7.8239	0.1127	31.0104	7.7289	0.0950	7.7251	0.0988
	L34	7.5686	0.1139	31.0104	7.7159	-0.1473	7.7218	-0.1532
	L16	7.6990	0.1058	31.0104	7.8013	-0.1023	7.8061	-0.1071
	L19	7.4202	0.1119	31.0104	7.7372	-0.3170	7.7499	-0.3297
	L2	7.7447	0.1119	31.0104	7.7372	0.0075	7.7369	0.0078
	L17	7.6778	0.1040	31.0104	7.8206	-0.1428	7.8279	-0.1501
	L36	7.5528	0.1118	26.5996	7.3036	0.2492	7.2900	0.2628
	L21	7.6383	0.0980	26.5996	7.4485	0.1898	7.4356	0.2026
	L18	7.6576	0.1108	26.5996	7.3139	0.3437	7.2953	0.3623
	L37	7.3768	0.1226	26.5996	7.1899	0.1868	7.1752	0.2015
	L22	7.2924	0.1016	26.5996	7.4109	-0.1185	7.4180	-0.1256
	L5	8.5229	0.0886	35.7206	8.4470	0.0758	8.4245	0.0983
	L26	7.3565	0.1353	31.0104	7.4912	-0.1347	7.5050	-0.1484
	L24	7.5376	0.1324	31.0104	7.5220	0.0156	7.5207	0.0169
	L6	5.7951	0.1177	14.2249	6.0204	-0.2253	6.4089	-0.6138
	L8	7.7212	0.1218	30.3173	7.5652	0.1560	7.5581	0.1631
	L9	7.7212	0.1217	30.3173	7.5658	0.1554	7.5588	0.1625
	L10	7.7959	0.1185	30.3173	7.5999	0.1960	7.5921	0.2038
	L11	7.5686	0.1064	30.3173	7.7264	-0.1577	7.7327	-0.1641
	L12	7.7447	0.1034	30.3173	7.7582	-0.0135	7.7588	-0.0141
	L13	7.6990	0.1021	30.3173	7.7723	-0.0733	7.7759	-0.0769
	L14	7.7447	0.0972	30.3173	7.8232	-0.0785	7.8283	-0.0836
	L25	7.8539	0.0985	31.0104	7.8782	-0.0243	7.8799	-0.0260
Test set	L1	6.7352	0.1653	34.3590	7.5072	-0.7720	-	-
	L3	7.2147	0.0677	26.5990	7.7658	-0.5511	-	-
	L4	7.3872	0.1007	31.0104	7.8554	-0.4682	-	-
	L7	7.0555	0.0749	26.5996	7.6908	-0.6353	-	-
	L23	7.6990	0.0962	31.0104	7.9025	-0.2035	-	-
	L29	7.3979	0.1110	31.0104	7.7467	-0.3487	-	-

1.1.Applicability Domain

The applicability domain of the chosen 2D-QSAR model was determined using MINITA.17's Williams plot (**Figure 2**), which displays leverages values vs standardized residuals. All compounds are in the applicability domain, except for compounds L1 and L6, which have a standardized residual and lever value greater than 3 and h* = 0.29, respectively. Although these compounds are considered outside the applicability domain.

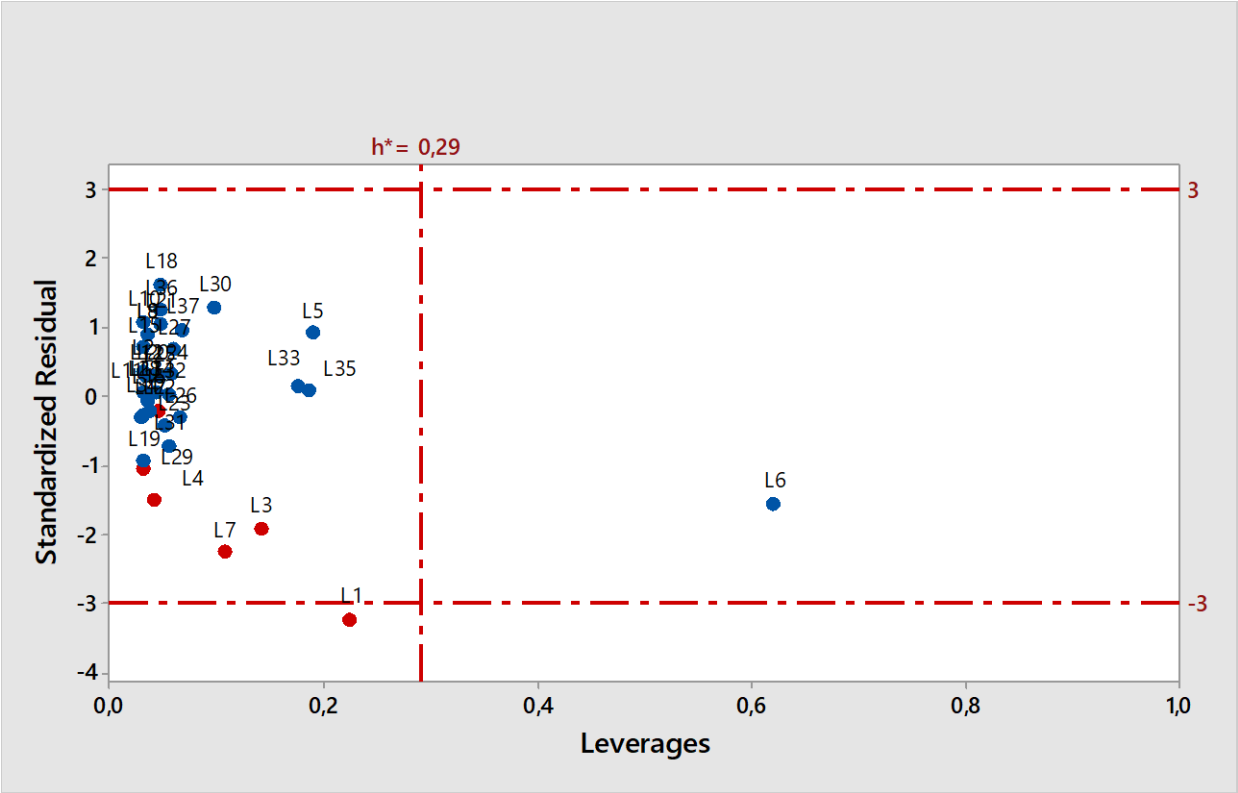


***Figure 2**. Williams plot for the developed 2D-QSAR model. Red and blue dots represent the test and training sets compounds, respectively*

2. Molecular Docking Analysis

Molecular docking was employed in order to determine the binding mode of interactions and affinity between the pyridinylimidazole derivatives and the p38MAPK receptor. First, the docking validation was performed by docking the original ligand into the binding pocket of the p38α MAPK receptor (PDB ID: 5ML5). The obtained best pose gave a root-mean-square deviation (RMSD) value of 1.29 Å. The superposition between the re-docked and the original ligands into the binding pocket of the p38α MAPK receptor is shown in **Figure 3**.

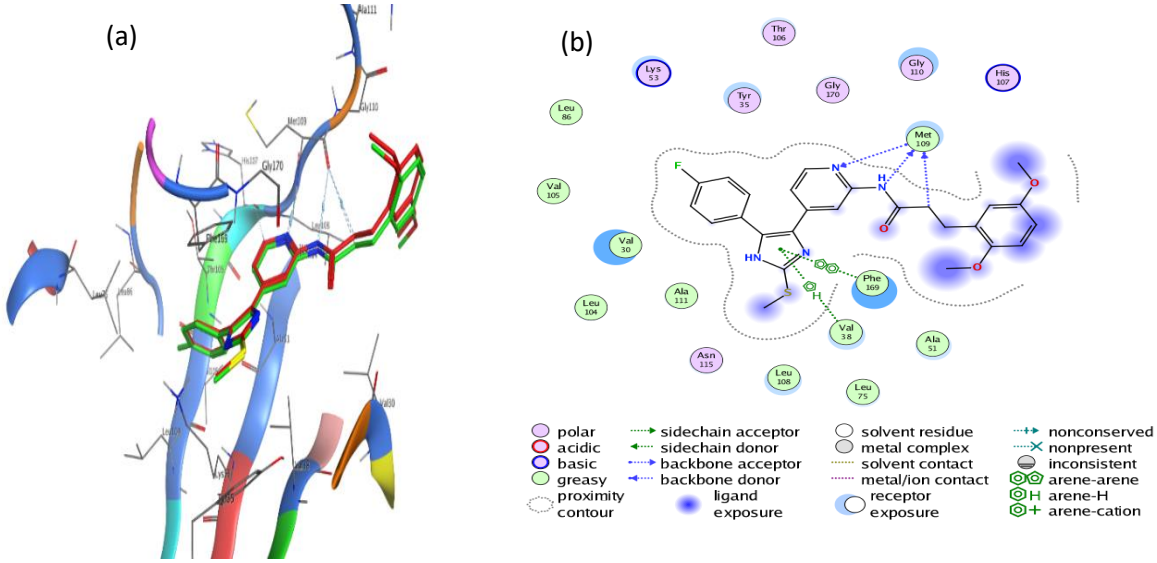


Figure 3. *Docking validation.* ***(a)*** *Superposition between the re-docked and the original ligands into the binding pocket of the p38α MAPK (PDB ID: 5ML5),* ***(b)*** *2D interactions of the native ligand into the binding p38α MAPK receptor*

After docking validation, the active compound (L5) in the collected data was used to perform docking studies. The result is shown in **Figure 4**.

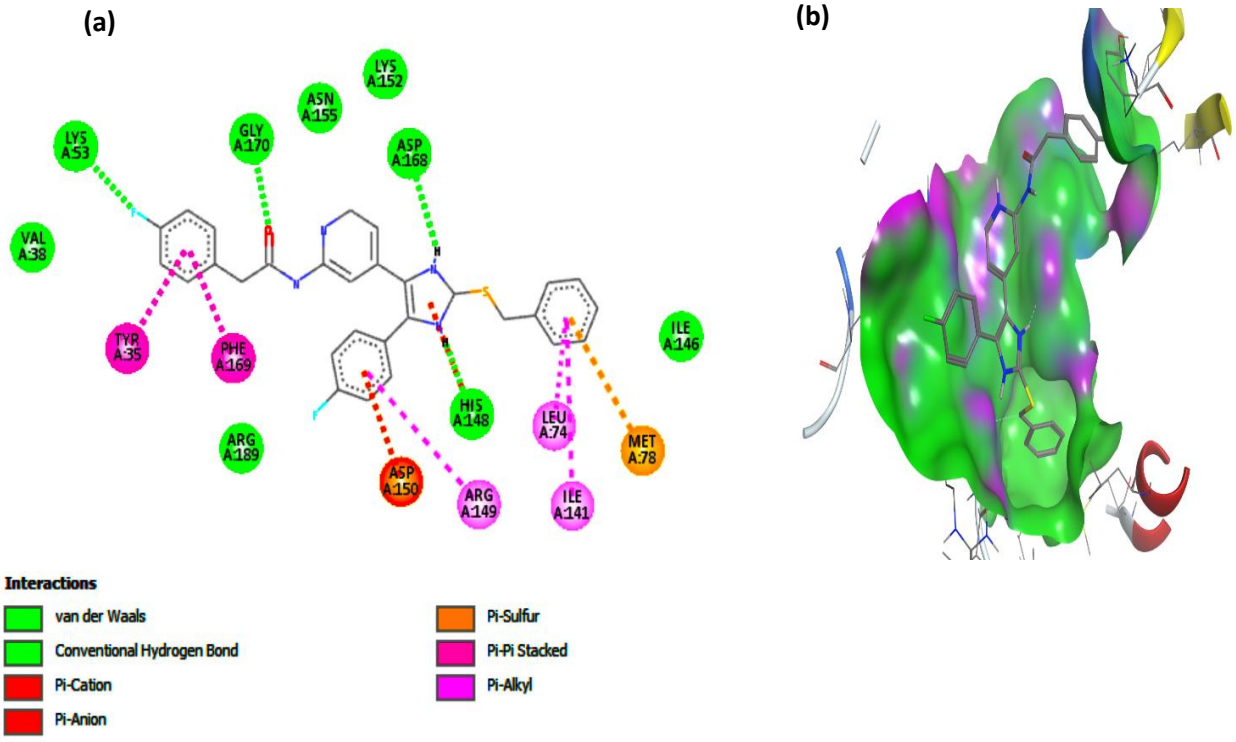


Figure 4. *Docking analysis of compound L5 into the binding pocket of p38α MAPK receptor.* ***(a)*** *2D ligand-receptor complex interactions,* ***(b)*** *3D ligand-receptor complex interactions*

As described in **Figure 4**, compound L5 forms several interactions with the p38α MAPK receptor. The most important interactions are conventional hydrogen bond interactions with LYS 53, GLY 170, ASP 168 and HIS 148. The stability of compound L5 in the receptor's binding pocket is demonstrated by a low binding free energy (-10.6 kcal/mol).

3. Structural Design of New Pyridinylimidazole derivatives

Based on the results of 2D-QSAR and molecular docking studies, a novel series of Pyridinylimidazole derivatives was proposed. The 2D-QSAR analysis revealed the importance of PEOE_RPC^{+} and SMR_VSA2 descriptors to explain the variation of the p38α MAPK inhibitory activity. The PEOE_RPC+ descriptor is associated with the relative positive partial charge (the largest positive qi divided by the sum of the positive qi). The negative influence of PEOE_RPC+ in Eq.7 indicates that the partial positive charge may decrease the inhibitory activity of the p38α MAPK receptor. **Figure 5** shows the calculated partial charge of each atom in compound L5 (the most active compound). The SMR_VSA2 descriptor represents the sum of van der Waals surface area such that Ri (Molar Refractivity for atom i) is in the range (0.26 - 0.35). The positive sign of this descriptor in Eq. (5) indicates the importance of molar refractivity in improving activity. In addition to the 2D-QSAR analysis, molecular docking study revealed the importance of conventional hydrogen bond interactions to improve the inhibitory activity of the p38α MAPK.

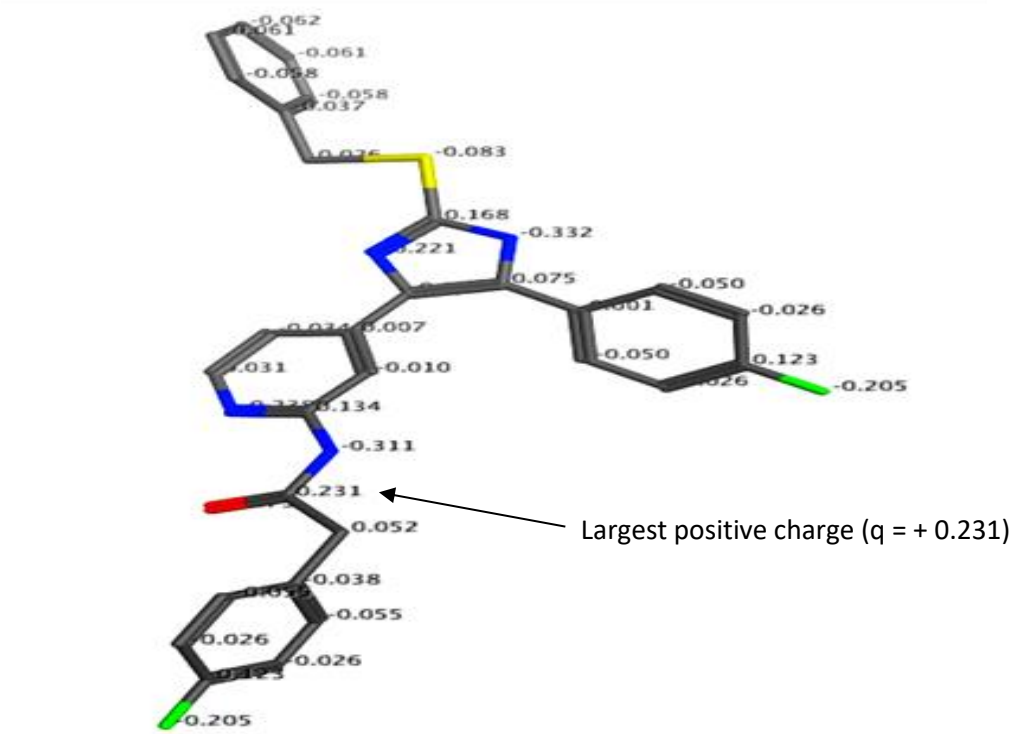
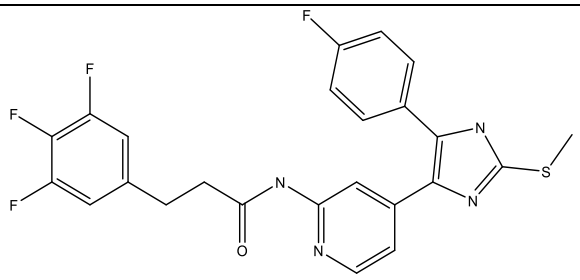
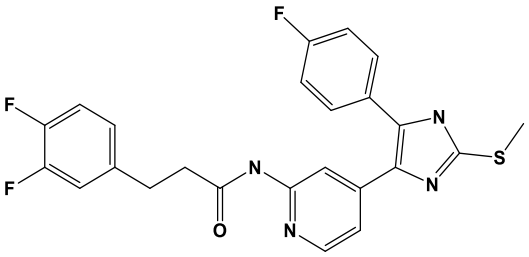
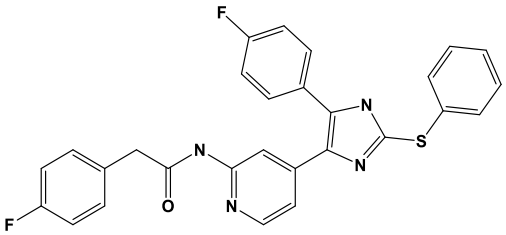


Figure 5. *The calculated partial charge of the most active compound in the dataset (Compound L5)*

Based on the above information, three compounds were proposed, and their activity was predicted using the generated 2D-QSAR model. The compounds were designed and optimized in the same way as those in the dataset. **Table 5** shows the chemical structures, the activity values that go with them, and the value of the binding free energy.

***Table** 5. Chemical structures and their predicted activity as well as the binding free energy of predicted compounds*

Compounds	Structures	pIC_{50}	Binding free energy (kcal/mol)
T1		10.2558	-10.6
T2		9.3042	-10.5
T2		8.5665	-9.4

As shown in Table 5, all predicted compounds have a higher pIC_{50} value than the most potent inhibitor in the dataset (pIC_{50} (L5) = 8.5229), which means that our model is accurate and can be used to design new p38α MAPK inhibitors. The predicted compounds were also evaluated by molecular docking analysis. The low values of binding free energy reflect the stability of these compounds in the binding pocket of the p38α MAPK receptor.

4. Molecular dynamics simulations

Molecular dynamics (MD) simulation is a technique for studying the movement and trajectories of a molecule with other molecule along with various inter-atomic and intermolecular interactions within the system [23, 24]. In this work, MD simulation was performed to evaluate the dynamic behavior and stability of our studied compounds in the p38α MAPK receptor binding pocket. The detailed analysis of the resulting MD trajectories was performed using different modules of GROMACS, which included the "rms module" for root mean square deviation (RMSD), the "gyrate module" for radius of gyration, and the "hbond module" for hydrogen bonding analysis.

RMSD analysis

RMSD is a measure of the average deviation of an atom's position from the reference frame for a specific frame. The smaller the gap, the greater the stability of the protein or protein-ligand complex. The RMSD plot of each studied system is presented in **Figure 6**.

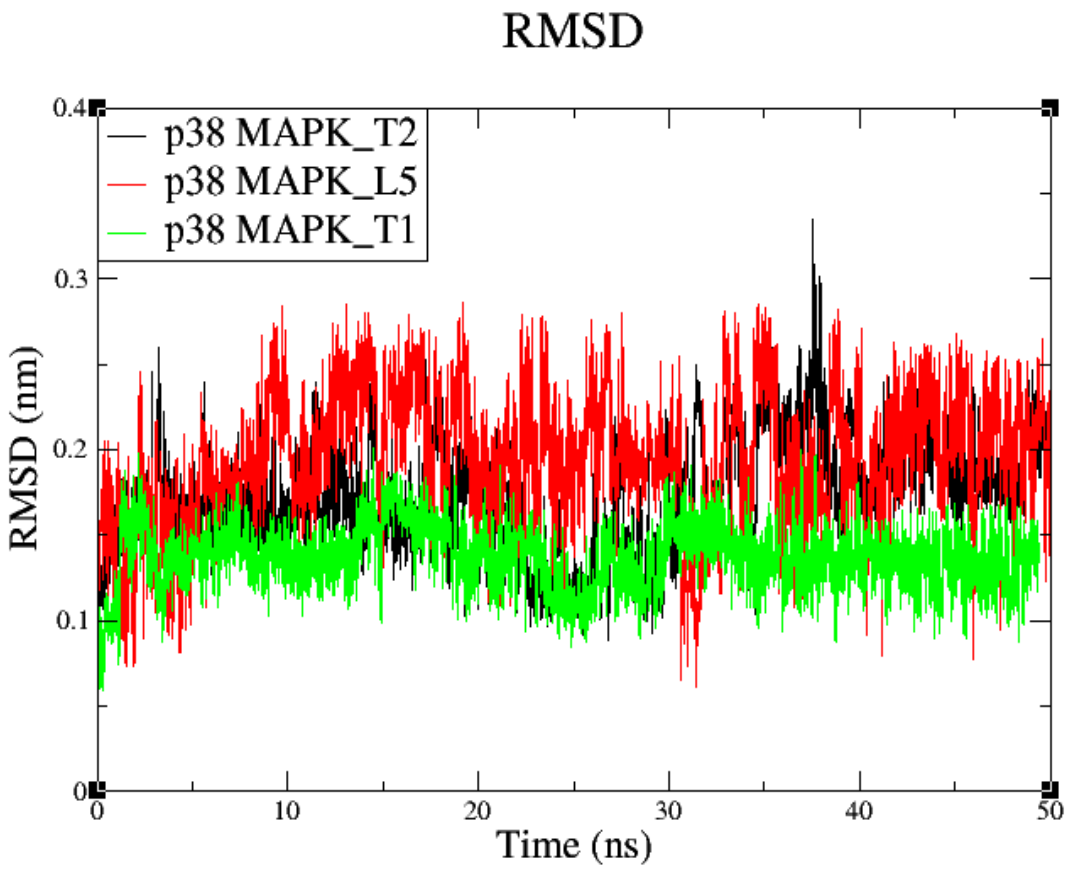


Figure 6. *Ligand RMSD plot of 50 ns of MD simulation. The green color represents the p38α MAPK _T1 complex, the black color represents the p38α MAPK _T2 complex, and the red color represents the p38α MAPK _L5 complex*

The RMSD of MAPK _T1 complex was stabilized at approximately 0.12 nm after 5 ns and remained undeviated for the rest of the time. Whereas the RMSD of the MAPK _T2 and MAPK _L5 complexes attained stability at 0.18 nm after 20 ns and remained stable for the rest of the time. The average RMSD values of p38α MAPK _T1, p38α MAPK _T2 and p38α MAPK _L5 complexes are 0.12 nm, 0.16 nm, and 0.18 nm, respectively.

Rg analysis

Rg is used to measure the compactness, conformational change, and folding behavior of the protein. The Rg plot of each studied system is presented in **Figure 7.** The average Rg values of p38α MAPK _T1, p38α MAPK _T2 and p38α MAPK _L5 complexes are 2.21 nm, 2.23 nm, and 2.24 nm, respectively. These results showed that compounds T1, T2, and L5 have a significant effect on the p38α MAPK receptor.

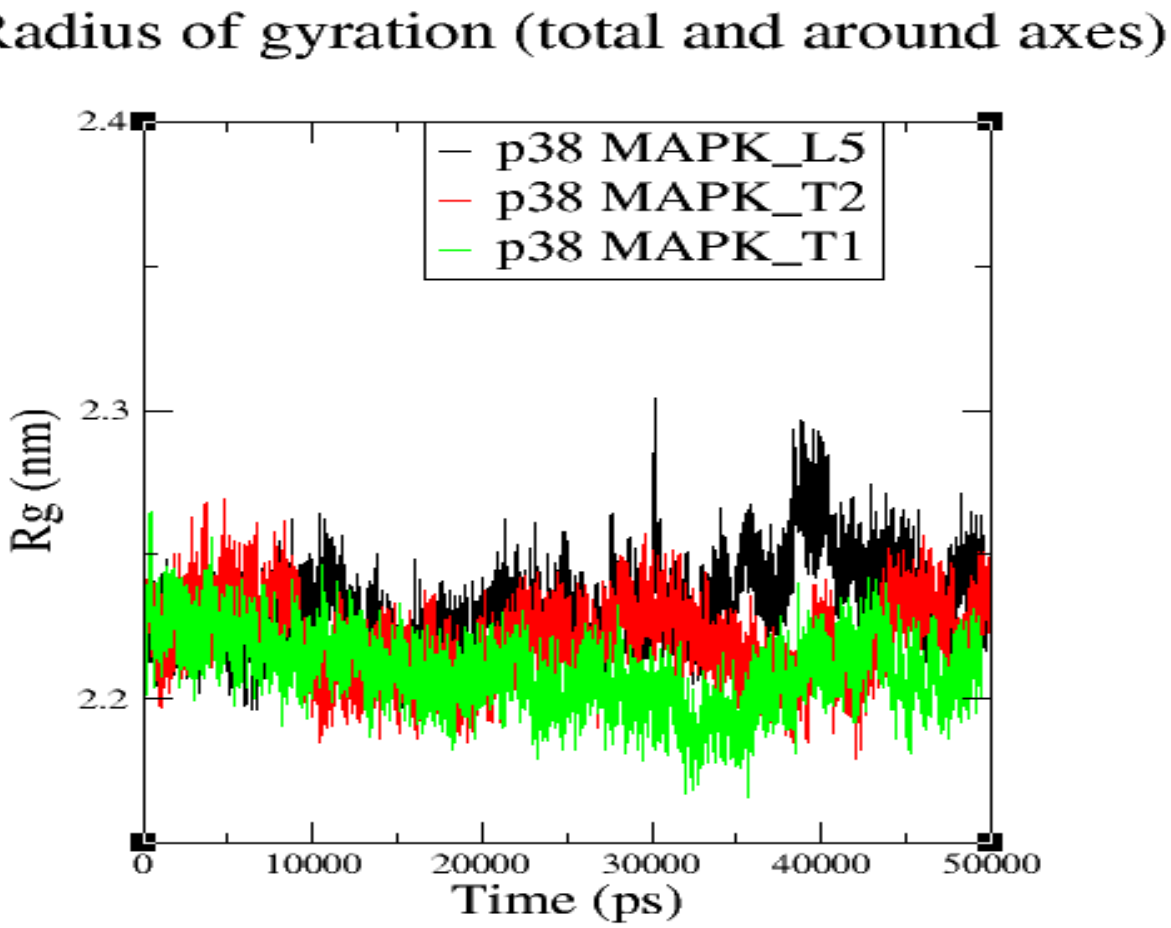


Figure 7. *Plot of Rg over the 50 ns of MD simulation at 300 K. The green color represents the p38α MAPK _T1 complex, the black color represents the p38α MAPK _T2 complex, and the red color represents the p38α MAPK _L5 complex*

Hydrogen bonding analysis

The formation of hydrogen bonds between the protein and ligand plays a key role in the binding and stability of the complex. The hydrogen bonds between the protein and ligand were analyzed using the g_hbond tool of Gromacs. The results are presented in **Figure 8**. The highly predicted compound (T1) forms a maximum number of hydrogen bonds of 6 with p38α MAPK binding pocket. The maximum h-bonds of T2 and L5 are 7 and 3, respectively. These results indicate that our predicted compounds (T1 and T2) are highly bonded in the p38α MAPK binding pocket.

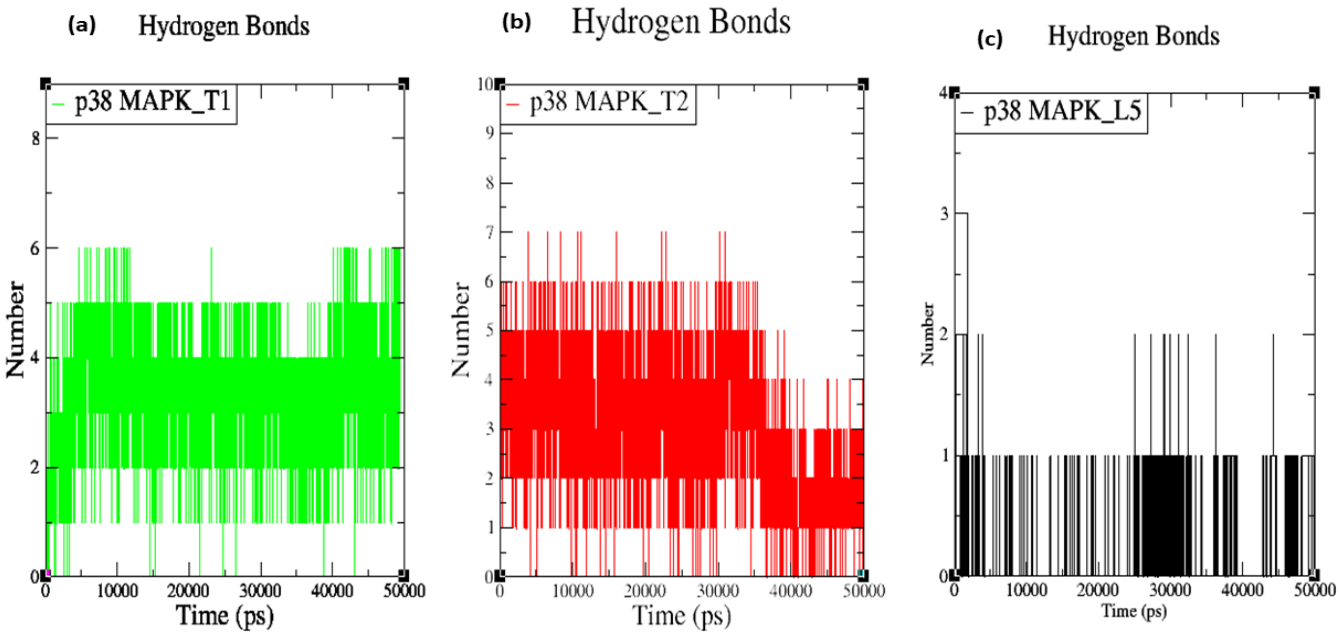


Figure 8*. Graphical representation of hydrogen bonding of p38α MAPK _T1 (green), p38α MAPK _T2 (red) and p38α MAPK _L5 (black) complexes*

5. ADMET analysis

The pharmacokinetic and toxicity parameters of our studied compounds were evaluated using the preADMET web server. The results are presented in **Table 6**.

The absorption is the limiting factor for the bioavailability and biological effects of a drug. For Caco2 cells and MDCK cells, all compounds (T1, T2, T3, and L5) had moderate and low permeability, respectively. For Caco2 cells and MDCK cells, all compounds (T1, T2, T3, and L5) had moderate and low permeability, respectively. The negative values of in-vitro skin permeability indicate that molecules are not suitable for transdermal administration. The predicted compounds (T1, T2 and T3) as well as the most active compound in the dataset (L5) were showing more than 93% predicted human intestinal absorption (HIA). Drugs targeting brain must cross the BBB, all compounds show high BBB permeability. Furthermore, the predicted plasma protein binding of all compounds was shown more than 87%, indicating reduced excretion rate and enhanced half-life. None of the selected molecules were predicted to inhibit of liver enzymes like CYP2C19, CYP2D6, and CYP3A4 (**Table 6**).

Table 6. *PreADMET server predicted pharmacokinetic properties*

No	HIA (%)	Caco2 (nm/s)	MDCK (nm/s)	SP (log Kp, cm/h)	BBB	PPB (%)	CYP2C19	CYP2D6	CYP3A4
							Inhibition	Inhibition	Inhibition
L5	94.94	51.86	0.05	-2.80	4.28	91.52	No	No	No
T1	93.92	46.14	2.52	-3.30	3.79	86.39	No	No	No
T2	93.90	44.18	0.84	-3.28	3.33	87.14	No	No	No
T3	94.85	27.85	0.05	-2.95	4.15	93.28	No	No	No

HIA = Human Intestinal Absorption [0-20 (poor), 20-70 (moderate), 70-100 (well)], **Caco2** = In-vitro Caco2 cell permeability [70 (high)], **MDCK** = Maden Darby Canine Kidney cell permeability [500 (high)], **SP** = Skin permeability, **BBB**= Blood brain barrier permeability (Cbrain/Cblood), **PPB** = In-vitro plasma protein binding.

The selected compounds (T1, T2 and L5) were also predicted as mutagens by the Ames mutagenicity test, except compound T3. Furthermore, all compounds were predicted as negative by carcinogenicity tests in animals (mouse).

Table 7. *PreADMET server predicted toxicity parameters*

No	Ames Test	Carcinogenicity (Mouse)
L5	Mutagen	negative
T1	Mutagen	negative
T2	Mutagen	negative
T3	non-mutagen	negative

Table 8 summarizes the drug-likeness and synthetic accessibility (SA) predictions of our predicted compounds (T1, T2, and T3) as well as the most active compound in the dataset (L5). These molecules were filtered against three drug-likeness filters viz. Lipinski's rule of five, Veber filter and Muegge filter. The compounds T1 and T2 were not found to violate any filter, but T3 was found to violate Log Po/w (MLOGP) AND Log Po/w (XLOGP3) parameters in Lipinski and Muegge filters, respectively. Compound L5 was found to violate MW and Log Po/w (MLOGP) parameters in Lipinski filter and Log Po/w (XLOGP3) parameter in the Muegge filter. Also, the synthetic accessibility of all predicted compounds is close to 1, which means that they are very easy to make.

Table 8. *Drug-likeness and Synthetic accessibility of most active hits*

Drug-likeness parameters	Compound			
	T1	T2	T3	L5
MW	486.48	468.49	498.55	512.57
Log Po/w (XLOGP3)	4.96	4.86	6.02	5.96
Log Po/w (MLOGP)	4.12	3.75	4.65	4.17
HBD	2	2	2	2
HBA	7	6	5	5
TPSA	95.97	95.97	95.97	95.97
MR	122.41	122.46	136.54	142.18
NRB	8	8	8	9
Bioavailability Score	0.55	0.55	0.55	0.17
Lipinski	Yes; 0 violation	Yes; 0 violation	Yes; 1 violation: MLOGP>4.15	No; 2 violations: MW>500, MLOGP>4.15
Veber	Yes	Yes	Yes	Yes
Muegge	Yes	Yes	No; 1 violation: XLOGP3>5	No; 1 violation: XLOGP3>5
SA	3.24	3.23	3.42	3.61

MW = Molecular Weight, **Log Po/**w = Lipophilicity**, HBD** = Hydrogen bond donor, **HBA** = Hydrogen bond acceptor, **TPSA** = Topological polar surface area, **MR** = Molecular refractivity, **NRB** = No. of rotatable bonds, **SA** = Synthetic accessibility (from 1 (very easy) to 10 (very difficult)).

Conclusion

In the current study, we built a 2D-QSAR model using a new series of pyridinylimidazole derivatives as p38α MAPK inhibitors. The 2D-QSAR model was built based on MOE descriptors and shows the importance of PEOE_RPC+ and SMR_VSA2 descriptors in describing the inhibitory activity of p38α MAPK. The generated model by the PLS method explains 85% and predicts 92% of the variance. We also applied molecular docking to analysis the binding sites and interactions between our studied compounds and p38α MAPK receptor. The results showed the importance of hydrogen bonds in the stability of ligand-receptor complexes. Based on these results, three new compounds with high potency were designed, and their activity was predicted using the 2D-QSAR model. The predicted compounds were also evaluated using MD simulation and ADMET investigations. All predicted compounds showed high stability in the binding pocket of p38α MAPK receptor with moderate pharmacokinetic proprieties. This generated work may pave the way to design and synthesize new p38α MAPK inhibitors with potential inhibitory activity for treating Alzheimer's disease.

Conflict of interest

All authors declare that they have no conflict of interest in this work.

Funding

This research received no external funding

Acknowledgements

I gratefully thank and deeply acknowledge anyone who gave a hand to make this work coming out to light.

References

1 W. R. Markesbery, *Arch. Neurol.*, 1999, **56**, 1449–1452.

2 S. J. Harper and P. Lograsso, *Cell. Signal.*, 2001, **13**, 299–310.

3 E. Martín-Blanco, *BioEssays*, 2000, **22**, 637–645.

4 J. K. Lee and N. J. Kim, *Molecules*, , DOI:10.3390/molecules22081287.

5 Q. Yu, F. Du, J. T. Douglas, H. Yu, S. S. Yan and S. F. Yan, *J. Alzheimer's Dis.*, 2017, **59**, 223–239.

6 V. Haller, P. Nahidino, M. Forster, S. A. Laufer, V. Haller, P. Nahidino, M. Forster and S. A. Laufer, *Expert Opin. Ther. Pat.*, 2020, **0**, 1.

7 F. Heider, F. Ansideri, R. Tesch, T. Pantsar, U. Haun, E. Döring, M. Kudolo, A. Poso, W. Albrecht, S. A. Laufer and P. Koch, *Eur. J. Med. Chem.*, 2019, **175**, 309–329.

8 A. El Aissouq, O. Chedadi, R. Kasmi, L. Elmchichi, F. En-nahli, A. Goudzal, M. Bouachrine, A. Ouammou and F. Khalil, *J. Solution Chem.*, 2021, **50**, 808–822.

9 A. El Aissouq, O. Chedadi, M. Bouachrine and A. Ouammou, *J. Chem.*, , DOI:10.1155/2021/1901484.

10 C. Oussama, *Biointerface Res. Appl. Chem.*, 2022, **12**, 5100–5115.

11 N. L. Allinger, *J. Am. Chem. Soc.*, 1977, **99**, 8127–8134.

12 A. El Aissouq and H. Toufik, *2019 Int. Conf. Intell. Syst. Adv. Comput. Sci.*, 2019, 1–7.

13 A. Goudzal, A. El Aissouq, H. El Hamdani and A. Ouammou, *Mater. Today Proc.*, , DOI:10.1016/j.matpr.2020.08.044.

14 L. El Mchichi, A. El Aissouq, R. Kasmi, A. Belhassan, R. El-Mernissi, A. Ouammou, T. Lakhlifi and M. Bouachrine, *Mater. Today Proc.*, , DOI:10.1016/j.matpr.2021.03.152.

15 A. Tropsha and A. Golbraikh, *Curr. Pharm. Des.*, 2007, **13**, 3494–3504.

16 J. Halekotte, L. Witt, C. Ianes, M. Krüger, M. Bührmann, D. Rauh, C. Pichlo, E. Brunstein, A. Luxenburger, U. Baumann, U. Knippschild, J. Bischof, C. Peifer, P. Koch and S. Laufer, *Optimized 4,5-diarylimidazoles as potent/selective inhibitors of Protein Kinase CK1δ and their structural relation to P38α MAPK*, 2017, vol. 22.

17 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindah, *SoftwareX*, 2015, **1**–**2**, 19–25.

18 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. Mackerell, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.

19 A. D. M. J. K. VANOMMESLAEGHE, E. HATCHER, C. ACHARYA, S. KUNDU, S. ZHONG, J. SHIM, E. DARIAN, O. GUVENCH, P. LOPES, I. VOROBYOV, *J. Comput. Chem*, 2010, **31**, 671–690.

20 S. Nosé, *Comput. Simul. Mater. Sci.*, 1991, **2384**, 21–41.

21 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

22 A. El Aissouq, M. Bouachrine, A. Ouammou and F. Khalil, *Neurosci. Lett.*, 2022, **786**, 136803.

23 A. EL El AISSOUQ, M. BOUACHRINE, A. OUAMMOU and F. KHALIL, *Turkish J. Chem.*, 2022, **46**, 687–703.

24 A. El El Aissouq, O. Chedadi, M. Bouachrine and F. Khalil, *J. Biomol. Struct. Dyn.*, 2022, **0**, 1–14.

Conclusion générale

Dans ce travail, nous avons présenté une étude détaillée sur la recherche et développement de nouveaux agents thérapeutiques par les méthodes de la modélisation moléculaire in silico. En effet, les méthodes de la modélisation *in silico* telles que les méthodes QSAR, le docking moléculaire, la dynamique moléculaire et autres ont été appliquées pour développer des nouvelles molécules candidates médicamenteuses destinées au traitement des maladies neurodégénératives.

Dans ce contexte, nous avons présenté dans ce travail cinq applications, parmi d'autres, que nous avons réalisées au cours de ces trois années de recherche.

Dans la première application, les méthodes de la modélisation *in silico*, y compris 2D-QSAR, docking et dynamique moléculaire, ont été appliquées à une série de 25 molécules dérivées de chalcone comme inhibiteurs de l'enzyme MAO-B. Dans cette partie nous avons étudié les paramètres physico-chimiques et biologiques qui influencent l'activité étudiée. Les résultats ont révélé l'importance des descripteurs BCUT_SMR_2, logP (o/w), SlogP_VSA4 et vsurf_IW3 dans la variation des valeurs de l'activité. Le mode de liaison de ces composés et leur relation avec l'activité biologique a été aussi étudié par les méthodes de docking moléculaire. Une grande corrélation a été trouvée (r^2 = 0.73) entre les valeurs de l'activité expérimentale et les valeurs de l'affinité de liaison calculées. La capacité de ces composés d'être actifs par voie orale a été déterminée à l'aide des propriétés ADMET. Les résultats indiquent que la plupart des composés ont des caractéristiques pharmacocinétiques modérées sans aucun effet secondaire. Les résultats de ce travail indiquent clairement l'importance de dérivés de chalcone dans l'inhibition de l'activité enzymatique de l'enzyme MAO-B, et elles peuvent ouvrir la voie pour le développement d'autres dérivés avec une activité inhibitrice potentielle.

Dans la deuxième application, nous avons poursuivi notre première étude que nous appelons cette fois-ci l'étude d'alignement basée sur le docking et l'étude 3D-QSAR. Tout d'abord, l'analyse de docking moléculaire a été appliquée pour déterminer la conformation optimale de chaque ligand. Ensuite, l'alignement basé sur le docking a été utilisé pour générer des modèles CoMFA et CoMSIA. Les résultats ont révélé l'importance des champs stériques, des champs électrostatiques et des accepteurs de liaisons hydrogène dans le contrôle de l'activité des composés. Sur la base de ces différentes propriétés issues de ces études, on a pu concevoir 4 molécules qui peuvent être d'un grand intérêt lors du processus de découverte de médicaments. Les études de la dynamique moléculaire et de l'énergie libre de liaison par la méthode MM_GBSA ont montré que ces composés prédits avaient une grande stabilité dans le site de liaison du récepteur de la MAO-B.

Dans la troisième application, un ensemble de 142 dérivés de chalcone, récoltés dans diverses plantes naturelles, a été analysé en combinant la méthode de la modélisation par homologie, le criblage virtuel et la simulation par dynamique moléculaire. L'objectif est de découvrir de nouvelles chalcones naturelles pour inhiber l'activité enzymatique de la MAO-B. Dans un premier temps, la structure 3D de la MAO-B humaine (hMAO-B) a été construite par la

méthode de la modélisation par homologie. Ensuite les structures 3D de toutes les molécules ont été générées par la méthode de DFT. Tous les composés ont été analysés en utilisant AutoDock Vina et ADMET. Parmi les 142 composés naturels, 10 composés ont été évalués en utilisant AutoDock 4.2. Enfin, une simulation par dynamique moléculaire a été réalisée pour évaluer la stabilité des meilleurs composés dans la poche de liaison du récepteur hMAO-B.

Dans de la quatrième application, une étude théorique par l'approche QASR a été réalisée sur une série de molécules dérivées de la 2-(4-aryl-2méthylpipérazine-1-yl)-pyrimidine-4-one en tant qu'inhibiteurs de GSK-3β. Les modèles 2D-QSAR ont été généré en utilisant les descripteurs de logiciel MOE (Molecular operating environment). La méthode PLS a été utilisée pour corréler les valeurs de l'activité étudiée avec les descripteurs calculés. Le mode de liaison et la sélectivité des ces composés ont été examinés par la méthode de docking moléculaire. Ces résultats, nous ont orientés vers une modification structurale dans la structure de base dans le but d'améliorer l'activité étudiée et de favoriser la formation de liaisons hydrogène avec le site de liaison du récepteur. Les cinq nouvelles molécules prédites ont montré une excellente inhibition et meilleures affinités en comparaison avec l'inhibiteur de référence (la molécule la plus active de la série étudiée).

Dans la cinquième application, nous avons construit un modèle 2D-QSAR en utilisant une nouvelle série de molécules, dérivées de pyridinyl imidazole comme inhibiteurs de p38a MAPK. Le modèle 2D-QSAR a été construit en utilisant un ensemble de 354 descripteurs moléculaires. Les résultats montrent l'importance des descripteurs PEOE_RPC+ et SMR_VSA2 dans la variation de l'activité étudiée. Nous avons également appliqué la méthode de docking moléculaire pour analyser les sites de liaison et les interactions entre les composés étudiés et le site de liaison du récepteur p38a MAPK. Les résultats ont montré l'importance des liaisons hydrogène dans la stabilité des complexes ligand-récepteur. Ces résultats ont permis de concevoir trois nouveaux composés avec une activité inhibitrice puissante. Les composés prédits ont également été évalués à l'aide de la simulation par dynamique moléculaire et les études de propriétés ADMET. Tous les composés prédits ont montré une grande stabilité dans le site de liaison du récepteur MAPK p38a avec des propriétés pharmacocinétiques modérées.

Les résultats obtenus dans cette thèse nous ont encouragés et motivés à compléter nos recherches sur les maladies neurodégénératives en :

- Validant les résultats obtenus au cours de cette thèse par des études in vitro/in vivo, afin de compléter le processus de la recherche et développement des médicaments (composés nouvellement prédits).
- Initiant des recherches sur d'autres molécules ayant une activité inhibitrice contre les enzymes responsables des maladies neurodégénératives.

Annexes

Tableau II.1. *Différents types de chalcones isolées à partir de différentes plantes médicinales*

Chalcone	Nom de la plante médicinale	Référence
Butein (3, 4,2',4'-tetrahydroxychalcone)	*Rhus verniciflua*	[242, 144, 243, 24]
Isoliquiritigenin (4, 2,4'-trihydroxychalcone)	*Glycyrrhiza glabra* - *Nepalese propolis*	[242, 144, 243] [245]
Trans-stilbene Piceatannol (3,5,3',4'-tetrahydroxy-trans-stilbene)	*Blueberries (Vaccinium sp.)*	[242, 144, 243]
Isobavachalcone 4-hydroxyderricin Xanthoangelol Xanthoangelol F Xanthoangelol D Xanthoangelol E Xanthoangelol B Xanthoangelol G Xanthokeistal A 2',4',4-Trihydroxy-3'-prenylchalcone	*Angelica keiskei*	[246, 247]
2',4-dihydroxy-3',4'-(2,2-dimethylchromene)-chalcone 2',4'-dihydroxy-4-methoxy-3'-prenyldihydrochalcone	*Artocarpus lowii*	[248]
Galiposin	*Galipea granulosa*	[249]
2'-Hydroxy-4',6'-dimethoxy-3'-methylchalcone	*Leptospermum scoparium*	[250]
Rhuschalcone	*Rhus pyroides*	[251, 252]
Azobechalcone	*Lophira alata*	[253]
(+)-Tephrosone Tephropurpurin	*Tephrosia purpurea*	[254] [255]
Munsericin	*Mundulea sericea*	[256]
Xanthohumol	*Humulus lupulus L. (Cannabaceae)* - *Hops Humulus*	[257] [258]
Millepachine	*Millettia pachycarpa.*	[259]
Licochalcone A	*Glycyrrhiza inflate*	[251]
Bavachalcone	*Psoralea corylifolia*	[260]
2',4',4,2"-Tetrahydroxy-3'-[3"- Methylbut-3"-Enyl] Chalcone	*Maclura tinctoria (Moraceae)*	[261]

Bakuchalcone Bavachromanol Echinatin Lichochalcone C	*Glycyrrhiza inflata*	[262]
abyssinone-VI-4-O-methyl ether	*Erythrina mildbraedii*	[263]
Crotaramosmin Crotaramin Crotin	*Crotolaria rosmosissima*	[264]
Myrigalone G	*Leptospermum recurvum*	[265]
Iryantherin K Iryantherin L	*Iryantheria megistophulla*	[266]
2’, 3, 4,4’-tetrahydroxy-2-prenylchalcone 3-methoxy-2’,4,4’-trihydroxy-2-prenylchalcone	*Shuteria sinensis*	[267]
Kuwanon J Kuwanon R Kuwanon V	*Morus bombycis*	[268]
Broussochalcone	*Broussonetiapapyrifa*	[269]
2', 4'-dihydroxy-3'-methoxychalcone 2', 4'-dihydroxychalcone	*Zuccagnia punctata*	[270]
Aurentiacin A Aurentiacin B	*Dydimocarpus aurentica*	[271]
Rubone	*Derris robusta*	[272]
2’,6’-Dihydroxy-4’-methoxydihydrochalcone 2’,4’,6’-Trihydroxydihydrochalcone Isoliquiritin	*Lindera umbellate*	[273]
Licuraside	*Glycyrrhizae radix*	[274]
2'-hydroxy-4, 4'-dimethoxychalcone	*Bidens tripartitus*	[275]
Ponganone I Ponganone II	*Pongamia pinnata*	[266]
Stipulin	*Dalbergia stipulacea*	[276]

3,3'-dihydroxychalcone	*Primula macrophylla*	[277]
Spinochalcone A Spinochalcone B	*Tephrosia spinosa*	[278]
Calythropsin Dihydrocalythropsin	*Calythropsis aurea*	[279]
2', 4,4',6'-tetrahydroxy-5-(E-3,7-dimethylocta-2,6- dienyl)-3-(3-methylbut-2-enyl) dihydrochalcone 2', 4,4',6'- tetrahydroxy-3,5-di (3-methylbut-2-enyl) dihydrochalcone	*Boronia inconspicua*	[280]
Pedicin Fissistin Isofissistin	*Fissistigma lanuginosum*	[281]
4',6'-dihydroxy-2'-methoxy-3',5'-dimethyl chalcone	*Syzygium samarangense*	[282]
Munchiwarin	*Crotalaria trifoliastrum*	[283]
1-[4-[(2E)-3,6-dimethylhepta-2,5-dienoxy]-2-hydroxyphenyl]-2-hydroxy-3-(4-hydroxyphenyl) propan-1-one	*Millettia usaramensis*	[284]
2',3'-Dihydroxy-4',6'-dimethoxychalcone	*Uvaria dulcis*	[285]
Prorepensin	*Dorstenia prorepens*	[286]
Dorsmanin A	*Dorstenia zenkeri*	[286]
4,2,4- trihydroxychalcone 4,2,4-trihydroxy-3-prenylchalcone L-hydroxypanduratin A Panduratin A Sakuranetin Pinostrobin	*Boesenbergia pandurate*	[287]
Dihydro-5,6-dehydrokawain Artoindonesianin J	*Artocarpus bracteata*	[288]
Lonchocarpin		[289]
Derricin Xanthoangelol H Laserpitin Isolaserpitin 3'-senecioyl khellacone	*Lonchocarpus sericeus*	[290]

4'-senecioyl khellactone Selinidin Trilobatin 2''-acetate Phloridzin Trilobatin	*Lithocarps pachyphylus*	[291]
Mallotophilippen C Mallotophilippen D Mallotophilippen E	*Mallotus philippinensis*	[292]
2′,6′-dihydroxy-4-isopenteniloxy-3,4-(3′′,3′′-dimethylpyrano) chalcone 4,2′,6′-trihydroxy-3′,4′-metilenodioxo-3-isopentenilchalcone	*Beilschmiedia tovarensis*	[293]
2',4'-dihydroxy-5'6'-dimethoxychalcone	*Cryptocarya costata*	[294]
Xanthokeismin A Xanthokeismin B Xanthokeismin C	*Angelica keiskei*	[295]
Garcinol	*Garcinia indica Choisy*	[296]
Isocordoin 2', 4'-dihydroxy-3'-(dimethylallyl)-dihydrochalcone	*Lonchocarpus xuul*	[297]
2,4,6-trihydroxy-5-(3,3-dimethyl propenyl)-3-(4-hydroxyphenyl) propiophenone	*Bacopa monnierie*	[298]
Morachalcone B Morachalcone C	*Mora alba L.*	[299]
Licochalcone G Echinantin 5-prenylbutein Licochalcone D Licoagrochalcone A	*Glycyrrhiza inflata*	[300]
Kanzonol C 4'-O-(6''-O-galloyl-β-d-glucopyranosyl)- 2',4-dihydroxychalcone 4′-O-β-D-glucopyranosyl-2′-hydroxy-4-methoxychalcone 4′-O-(6″-O-galloyl-β-D-glucopyranosyl)-2′-hydroxy-4-methoxychalcone	*Entada phaseoloides*	[301]
4-methoxylonchocarpin Isobavachromene Dorspoinsettifolin	*Millettia pachycarpa*	[302]

2',4'-dihydroxychalcone 2',4',4-trihydroxychalcone 2'-hydroxy-4'-methoxychalcone 2',4'-dihydroxy-4-methoxychalcone	*Oxytropis falcata*	[303]
Bractelactone	*Fissistigma bacteolatum*	[304]
Nardokanshone A	*Nardostachys chinensis*	[305]
Elastichalcone A Elastichalcone B	*Artocarpus elasticus*	[306]
3,2'-dihydroxy-4,3'-dimethoxychalcone-4'-glucoside 2',4',3-trihydroxy-3',4-dimethoxychalcone	*Coreopsis lanceolate*	[307]
(E)-1-(2,4-dihydroxy-6-methoxy-3-(3-methylbut-2-en-1-yl)phenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one (E)-1-(2,4-dihydroxy-3-(1-hydroxy-2-methylallyl)-6-methoxyphenyl)-3-(4-hydroxyphenyl) prop-2-en-1-one (E)-1-(5-hydroxy-7-methoxy-2,2-dimethyl-2H-chromen-6-yl)-3-(4-hydroxyphenyl)prop-2-en-1-one (E)-1-(3-(1,2-dihydroxy-2-methylpropyl)-2,5-dihydroxy-6-methoxyphenyl)-3-(4-hydroxyphenyl)prop-2-en-1-one (E)-1-(5-hydroxy-7-methoxy-2,2-dimethyl-2H-chromen-6-yl)-3-(4-hydroxyphenyl)prop-2-en-1-one (E)-1-(4-hydroxy-2-(2-hydroxypropan-2-yl)-6-methoxy-2,3-dihydrobenzofuran-7-yl)-3-(4-hydroxyphenyl) prop-2-en-1-one Xanthohumol-M	*Humulus lupulus*	[308]
Cardamonin Uvangoletin	*Piper aduncum L.*	[309]
Sappanchalcone 3-deoxysappanchalcone	*Haematoxylum campechianum L.*	[310]
Flavokawain B Pashanone	*Persicaria lapathifolia*	[311]

Liste des publications

Liste des publications en relation avec la thèse, publiées dans des journaux indexés et impactés :

Publication 1: New Journal of Chemistry, IF = 3.925 (2023) / H-index = 131 / Q2/ Scopus.

Abdellah El Aissouq*, Aissam Lachhab, Said El Rhabori, Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Computer-aided drug design applied to a series of pyridinylimidazole derivatives targeting p38α MAP kinase: 2D-QSAR, docking, MD simulation, and ADMET investigations"*. New Journal of Chemistry (2023). doi.org/10.1039/D2NJ03686J

Publication 2: Biomolecular structure and dynamics, IF = 5.235 (2023) / H-index = 81 / Q1/ Scopus.

Abdellah El Aissouq*, Mohammed Bouachrine, Lahcen Bouayyadi, Abdelkrim Ouammou and Fouad Khalil. *"Structure-based virtual screening of novel natural products as chalcone derivatives against SARS-CoV-2 M^{pro}"*. Biomolecular structure and dynamics (2023). doi.org/10.1080/07391102.2023.2172456

Publication 3: Neuroscience Letters IF = 3.198 (2022) / H-index = 184 / Q2/ Scopus.

Abdellah El Aissouq*, Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Homology modeling, virtual screening, molecular docking, molecular dynamic (MD) simulation, and ADMET approaches for identification of natural anti-Parkinson agents targeting MAO-B protein"*. Neuroscience Letters, Vol. 786, (2022). Doi: https://doi.org/10.1016/j.neulet.2022.136803.

Publication 4: Biomolecular structure and dynamics, IF = 5.235 (2022) / H-index = 81 / Q1/ Scopus.

Abdellah El Aissouq*, Oussama Chedadi, Mohammed Bouachrine, Abdelkrim Ouammou, Fouad Khalil. *"Development of novel monoamine oxidase B (MAO-B) inhibitors by combined application of docking-based alignment, 3D-QSAR, ADMET*

prediction, molecular dynamics simulation, and MM_GBSA binding free energy". Journal of Biomolecular Structure and Dynamics, (2022). Doi: https://doi.org/10.1080/07391102.2022.2071341.

Publication 5: Turkish journal of Chemistry, IF = 2.669 (2022) / H-index = 51 / Q3/ Scopus.

Abdellah El Aissouq*, Mohammed Bouachrine, Abdelkrim Ouammou, and Fouad Khalil *"Computational investigation of unsaturated ketone derivatives as MAO-B inhibitors by using QSAR, ADME/Tox, molecular docking and molecular dynamics simulations"*. Turkish journal of Chemistry, Vol. 46 (3), (2022), p. 687-703. Doi: 10.55730/1300-0527.3360.

Publication 6: Journal of Solution Chemistry, IF = 2.000 (2021) / H-index = 70 / Q3/ Scopus.

Abdellah El Aissouq, Oussama Chedadi, Rania Kasmi, Larbi Elmchichi, Fatima En-nahli, Amina Goudzal, Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Molecular modeling studies of C-Glycosylfavone derivatives as GSK-3β inhibitors based on QSAR and docking analysis"*. Journal of Solution Chemistry, Vol. 50 (5), (2021). Doi: https://doi.org/10.1007/s10953-021-01083-6

Publication 7: RHAZES: Green and Applied Chemistry

Abdellah El Aissouq*, Said El Rhabori, Hicham Zaitan, Mohammed Bouachrine, Abdelkrim Ouammou , and Fouad Khalil "*Molecular docking and ADMET investigation of a novel series of Pyridinylimidazole derivatives as potential p38α MAP kinase inhibitors"* RHAZES: Green and Applied Chemistry, Vol. 16, 2022, pp 72-82.

Liste des articles de collaboration publiées dans des journaux indexés et impactés :

Publication 1: Physical Chemistry Research, IF = 0.269 (2023) / H-index = 16 / Q4/ Scopus.

Lhoucine Naanaai, **Abdellah El Aissouq**, Hicham Zaitan, Fouad Khali "*3D QSAR, Molecular Docking, and ADMET Studies of a Series of 2-Acetylphenol-Rivastigmine Hybrids against Monoamine Oxidase A Inhibitors"* Physical Chemistry Research, Vol 12 (1), (2023). Doi: 10.22036/PCR.2023.394159.2326

Publication 2: Anti-Cancer Drugs, IF = 2.389 (2022) / H-index = 98 / Q3/ Scopus.

Said El Rhabori, **Abdellah El Aissouq**, Samir Chtita, Fouad Khalil. "Design of novel quinoline derivatives as anti-breast cancer using 3D-QSAR, molecular docking and ADMET studies" Anti-Cancer Drugs, (2022). DOI: 10.1097/CAD.0000000000001318

Publication 3: Journal of the Indian Chemical Society, IF = 0.243 (2022) / H-index = 33 / Q4/ Scopus.

Said El Rhabori, **Abdellah El Aissouq**, Samir Chtita, Fouad Khalil. "*3D-QSAR, Molecular docking and ADMET studies of thioquinazolinone derivatives against breast cancer*" Journal of the Indian Chemical Society, Vol. 99 (10), (2022). DOI: https://doi.org/10.1016/j.jics.2022.100675.

Publication 4: Structural Chemistry, IF = 1.795 (2022) / H-index = 50 / Q3/ Scopus

Said El Rhabori, **Abdellah El Aissouq**, Samir Chtita, Fouad Khalil *"QSAR, molecular docking and ADMET studies of quinoline, isoquinoline and quinazoline derivatives against Plasmodium falciparum malaria"*. Structural Chemistry, (2022). Doi: https://doi.org/10.1007/s11224-022-01988-y

Publication 5: RHAZES: Green and Applied Chemistry

Said El Rhabori, **Abdellah El Aissouq**, Samir Chtita and Fouad Khalil *"DFT, ANN and QSAR studies of quinoline, isoquinoline and quinazoline derivatives against Plasmodium falciparum"* RHAZES: Green and Applied Chemistry, Vol. 16, 2022, pp 63-71.

Liste des communications orales

- **Orale 1:** Said El Rhabori, **Abdellah El Aissouq,** Yassine El Allouche, Samir Chtita and Fouad Khalil **"***Discovery of new aromatase inhibitors for breast cancer treatment using computer-aided drug design (CADD) methods*". Cukurova 10th International Scientific Researches Conference, April-2-4, 2023, Adana, Turkey.

- **Orale 2:** Yassine El Allouche, **Abdellah El Aissouq,** Said El Rhabori, Hicham Zaitan and Fouad Khalil **"***Identification of Novel Benzodiazepine-1,2,3-triazol Derivatives as Butyrylcholinesterase inhibitors through 2D-QSAR and Docking Studies".* Cukurova 10th International Scientific Researches Conference, April-2-4, 2023, Adana, Turkey.

- **Orale 3: Abdellah El Aissouq*,** Abdelkrim Ouammou, and Fouad Khalil. *"Quantitative structure-activity relationship (QSAR) and molecular docking studies a novel series of 2-(4-aryl-2-methylpiperazin-1-yl)-pyrimidin-4-one derivatives as GSK-3β inhibitors".* International Conference on Global Practice of Multidisciplinary Scientific Studies help, July 26-28, 2022, Batumi, Georgia.

- **Orale 4: Abdellah El Aissouq**[*], Said El Rhabori, Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"In silico prediction of novel Glycogen synthase kinase-3 beta (GSK-3β) inhibitors: 3D-QSAR, molecular docking and ADMET prediction studies"***.** Doctoral day of LPME 2022, June 23, 2022, FST- USMBA-Fès.

- **Orale 5: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Discovery of new Glycogen synthase kinase-3 beta (GSK-3β) inhibitors through structure-based virtual screening"*. International ICONTECH SYMPOSIUM-3 on Innovative Surveys in Positive Sciences, 28-29 January 2021, Mohammed First University Oujda, Morocco.

- **Orale 6: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Discovery of novel 2-(4-aryl-2-methylpiperazin-1-yl)-pyrimidin-4-one derivatives for Alzheimer's disease treatment: Structure and ligand-based drug design approaches".*

MAS 14th International European Conference on Mathematics, Engineering, Natural & Medical Sciences, March 26-28, 2021 Széchenyi István University, Hungary.

- **Orale 7: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Virtual Identification of novel monoamine oxidase B (MAO-B) inhibitors by combining docking-based alignment, 3D-QSAR and ADMET prediction studies".* ENRSC2021, 24-25th, November 2021, Khenifra, Morocco.

- **Orale 8: Abdellah El Aissouq**, Abdelkrim Ouammou and Fouad Khalil. *"Inhibition of monoamine oxidase by chalcone analogues: docking-based alignment, 3D-QSAR and molecular dynamics simulation"*. International Siirt Conference on Scientific Siirt University, November 5-7, 2021.

- **Orale 09: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"In silico study of novel glycogen synthase kinase-3β (GSK-3β) inhibitors for Alzheimer's disease: 3D-QSAR, molecular docking and ADME/Tox prediction studies"*. International IGDIR Conference on Applied Science, April 14-15, 2021, Igdır, Turkey.

- **Orale 10: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil**.** *"Computational evaluation of some unsaturated ketone derivatives as MAO-B inhibitors, application of QSAR, ADMETand docking studies"*. International ICONTECH SYMPOSIUM-4 on Innovative Surveys in Positive Sciences, July 18-19, 2021, Adana, Turkey.

- **Orale 11:** Said El Rhabori, **Abdellah El Aissouq,** Samir Chtita and Fouad Khalil**.** *"3D-QSAR, Molecular docking and ADMET studies of thioquinazolinone derivatives against breast cancer".* Cinquième Colloque Maghrébin sur la Chimie Hétérocyclique CMCH 5-2022, organisé durant les 13, 14 & 15 Juillet 2022, La Faculté des Sciences, Université Moulay Ismail, Meknès au Maroc et L'Association Tunisienne de Chimie Hétérocyclique et ses Applications (STCHA) en Tunisie.

- **Orale 12:** Said EL RHABORI, **Abdellah EL AISSOUQ**, Samir CHTITA, and Fouad KHALIL. *"Molecular modeling and QSAR studies of quinolone, isoquinoline and quizoline derivatives against the* Plasmodium falciparum Chloroquine-sensitive *and Mefloquine-*

resistant strain protozoan Parasite". ENRSC2021, 24-25th, November 2021, Khenifra, Marocco.

- **Orale 13:** Said EL RHABORI, **Abdellah EL AISSOUQ**, Samir CHTITA, and Fouad KHALIL. *"In silico design of novel thioquinazolinone derivatives as anti-breast cancer agents using: 3D QSAR, molecular docking studies and ADMET prediction"*. Doctoral day of LPME 2022, June 23, 2022, FST- USMBA-Fès.

- **Orale 14**: Said El Rhabori, **Abdellah El Aissouq**, Samir Chtita and Fouad Khalil. *"Application of 2D-QSAR method on the study of new quinoline, isoquinoline, quinazoline derivatives against the Plasmodium falciparum CQ-sensitive strain W2 protozoan parasite"*. International IGDIR Conference on Applied Science, April 14-15, 2021 / Igdır, Turkey.

- **Orale 15:** Fatima EN-NAHLI, **Abdellah EL AISSOUQ**, Halima HAJJI, Fouad KHALIL, Mohammed Aziz AJANA, Tahar LAKHLIFI, Mohammed BOUACHRINE. *"Isoxazoline-containing Podophyllotoxin/2'(2',6')-(DI) Halogenopodo - phyllotoxin Derivatives as Acaricidal Activities against Tetranychus cinnabarinus. 2D-QSAR Study by using Molecular Operating Environment (MOE)"***.** International ICONTECH SYMPOSIUM-3 on Innovative Surveys in Positive Sciences, 28-29 January 2021, Mohammed First University Oujda, Morocco.

- **Orale 16:** Fatima EN-NAHLI, Halima HAJJI, **Abdellah EL AISSOUQ**, Fouad KHALIL, Mohammed Aziz AJANA, Tahar LAKHLIFI, Mohammed BOUACHRINE. *"Mechanism of inhibition of Escherichia coli by using Molecular Docking technics and ADMET studies".* MAS 14th International European Conference on Mathematics, Engineering, Natural & Medical Sciences, March 26-28, 2021 Széchenyi István University, Hungary.

- **Orale 17:** Said El Rhabori, **Abdellah El Aissouq,** Samir Chtita and Fouad Khalil. *"Application of 2D-QSAR method on the study of new quinoline, isoquinoline, quinazoline derivatives against the Plasmodium falciparum CQ-sensitive strain W2 protozoan parasite"*. International IGDIR Conference on Applied Science, April 14-15, 2021, Igdır, Turkey.

Liste des communications affichées

- **Poster 1: Abdellah El Aissouq**[*], Said El Rhabori, Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Discovery of novel GSK-3β inhibitors by combined application of 2D-QSAR and molecular docking studies"*. Doctoral day of LPME 2022, June 23, 2022, FST- USMBA-Fès.

- **Poster 2: Abdellah El Aissouq,** Mohammed Bouachrine, Abdelkrim Ouammou and Fouad Khalil. *"Potential inhibitors of Glycogen synthase kinase-3 beta (GSK-3β): 2D-QSAR and molecular docking studies".* 11th International Congress on Mathematic, Engineering and Natural Sciences, January 30-31, 2021, Karak, JORDAN.

- **Poster 3:** Said EL RHABORI, **Abdellah EL AISSOUQ**, Samir CHTITA and Fouad KHALIL. *"Design of new anti-breast cancer using 3D-QSAR, Molecular docking and* ADMET studies". Doctoral day of LPME 2022, June 23, 2022, FST- USMBA-Fès.

- **Poster 4:** Said EL RHABORI, **Abdellah EL AISSOUQ**, Samir CHTITA and Fouad KHALIL. *"3D-QSAR, molecular docking and ADMET studies of thioquinazolinone derivatives against cancer"*. Cinquième colloque Maghrébin sur la chimie Hétérocyclique CMCH 5-2022, 13, 14 et 15 Juillet 2022, Meknès, Maroc.