

Université Mohamed khider de Biskra

Faculté des sciences exactes et des science de la nature et de la vie

Département des sciences de la matière



MÉMOIRE DE MASTER

Science de la matière

Chimie

Chimie pharmaceutique

Mémoire de fin d'étude en Master Intitulé

Réf.n:

Présenté et soutenu par :

Belouerghi Souhila Azri Souad

Le: Septembre 2020

Etude qualitative et quantitative des relations structure-activité d'une série de composés phénoliques

Jury :

M. Daoud Ismail	MCA Université Mohamed Khider de Biskra	Président
Mme. Zekri Afaf	MAA Université Mohamed Khider de Biskra	Examinatrice
Mme.Lehraki faiza	MAA Université Mohamed Khider de Biskra	Rapporteur

Année universitaire : 2019 – 2020.

Remerciements

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

Je tiens à exprimer mes vifs remerciements à mon encadreur Madame Lehraki Faiza pour son aide, ses conseils, sa patience et sa grande disponibilité tout au long de la réalisation de ce mémoire, ainsi pour le temps qu'il a bien voulu nous consacrer et sans lui ce mémoire n'aurait jamais vu le jour.

Nous exprimons nos profonds remerciements aux membres de jury qui vont juger notre recherche.

J'adresse, enfin et surtout, ma plus profonde gratitude et tout mon amour à ma mère, mon père, mes sœurs et mes frères, qui ont su me faire confiance et me soutenir en toutes circonstances.

Merci

Dédicace

Je dédie:

A celle qui m'a commandé, la Très Miséricordieuse, à moi du Paradis sous ses pieds à celle qui a épargné sa vie pour me voir dans l'orgueil de la santé et du bonheur, même si sur elle-même, elle mérite de lui donner ma joie, mais ma vie lui est donnée.

À mon cher père, je vous répète mes remerciements du fond du cœur, vous étiez à mes côtés.

Vers la mer d'amour, de tendresse et le poulx qui habite le cœur de ma grand-mère

Aux étoiles de mon ciel scintillant et au soutien vital de mon frère

Au soutien de l'avenir et à la bougie du don, mes professeurs, à lehraki Faiza, Boucettea farida, ghamri meryam.

A chaque cœur qui m'a battu d'amour et de peur pour moi, je leur dédie le fruit de mon humble effort à amira, oum Kaltum et Lamia.

À ceux qui ont passé de beaux jours avec eux, à ceux qui aiment vivre avec eux, mes amis : Donia. Farah, Boutheyra, Noor elHouda, Rania, Sana, Nada, Nassima, Oum elhana, Samira. Amira, Hassna, Latifa, Samra, Khawla..

Merci à tous ceux qui m'ont appris une lettre afin d'atteindre ce que j'ai atteint, merci à tous ceux qui m'ont appelé au dos de l'invitation invisible.

Souhila

Dédicace

A l'aide de dieu tout puissant, qui m'a tracé le chemin de ma vie, j'ai pu réaliser ce travail que je dédie :

A ma très chère Mère Zohra Azri et à mon cher père Ali Azri paix à son âme, qui mes apporté leur appui durant toutes mes années d'étude, pour leur sacrifice et soutien qui m'ont donné confiance, courage et sécurité, pour tout cela et pour ce qui ne peut être dit, mes affections sans limite.

A mes très chères sœurs, Fadhila, Nacira, Fatma, Aicha et Hadjira.

A mes frères Mohammed, Hamza, Yassine et Amine.

A mes nièces et neveux et toute la famille pour leur appui et leur encouragement.

A mon professeur Mme Lehraki Faiza pour m'avoir encouragé et pour son engagement et son soutien ainsi que pour la pertinence de ses remarques et de ses feed-back.

A mes amis ceux que j'aime.

A tous ceux qui ont participé de près ou de loin à la réalisation de ce travail.

Souad

Sommaire

Liste des figures

Liste des tableaux

Liste des abréviations

Introduction générale.....1

Références bibliographiques.....3

Chapitre I: Généralités sur les composés phénoliques et l'activité antioxydante

I.1. Introduction	4
I.2. Composés phénoliques	5
I.2.1. Classification des polyphénols:.....	5
I.2.1.1. Les phénols simples et acides phénoliques:.....	5
I.2.1.2. Les flavonoïdes:.....	7
I.2.1.3. Les tannins:	8
I.2.1.4. Les coumarines:	9
I.2.1.5. Les quinones:.....	10
I.2.2. Propriétés biologiques des polyphénols.....	10
I.3. Radicaux libres et Stress oxydatif.....	11
I.3.1. Radicaux libres.....	11
I.3.1.1. Définition.....	11
I.3.2. stress oxydant.....	11
I.3.2.1. Définition.....	11
I.3.2.2. Les maladies liées au stress oxydant	12
I.3.3. Différents types des radicaux libres.....	12
I.3.4. Origine de production des radicaux libre.....	13
I.3.5. Dommages causés par les radicaux libres.....	13
I.3.5.1. Action sur l'ADN	14
I.3.5.2. Action sur les lipides	14
I.3.5.3. Action sur les glucides	14
I.3.5.4. Action sur les protéines.....	14
I.3.6. Systèmes de défense contre les ERO	15
I.4. Les anti oxydants.....	15
I.4.1. Antioxydants enzymatiques.....	15
I.4.2. Antioxydants non enzymatiques.....	16
I.4.3. Modes d'action anti-oxydante des phénols.....	17
I.4.3.1. Piégeage direct des radicaux libres :.....	17
I.4.3.2. Chélation des ions métalliques :	17
I.4.3.3. L'inhibition des enzymes responsables de la génération des ERO:	17
I.4.4. Méthodes de quantification de l'activité anti-oxydante.....	18

Références bibliographiques.....	20
----------------------------------	----

Chapitre II : Bases théoriques

II.1. Introduction.....	24
II.2. Méthodes de calculs.....	25
II.2.1. Méthodes quantiques.....	25
II.2.1.1. Méthodes semi empiriques.....	25
II.2.1.2. Méthodes ab initio.....	25
II.2.1.3. La théorie de la fonctionnelle de la densité DFT.....	26
a. L'objectif.....	26
b. Approximations en DFT.....	27
b.1. Approximation locale de la densité (LDA).....	27
b.2. L'approximation du gradient généralisé (GGA).....	27
b.3. Fonctionnelles hybrides pour le calcul moléculaire.....	28
c. La fonctionnelle B3LYP.....	28
d. Bases d'orbitales atomiques.....	28
II.2.2. Mécanique moléculaire.....	30
II.2.2.1. Différents champs de force en mécanique moléculaire.....	31
II.2.2.2. Principes de la minimisation.....	33
II.2.3. Dynamique moléculaire.....	33
II.3. QSAR.....	34
II.3.1. Principe de QSAR.....	34
II.3.2. Les applications de l'étude QSAR.....	35
II.3.3. Outils et Méthodologie de QSAR.....	35
II.3.3.1. Paramètres biologiques.....	35
II.3.3.2. Utilisation des descripteurs moléculaires.....	36
II.3.3.3. Sélection et réduction du nombre de descripteurs.....	40
II.4. Méthodes statistiques.....	41
II.4.1. Définition.....	41
II.4.2. Méthodes statistiques.....	41
II.4.2.1. La statistique descriptive.....	42
II.4.2.2. La statistique décisionnelle ou prédictive.....	42
II.4.3. La régression linéaire multiple.....	43
II.4.4. Techniques de validation.....	44
II.4.4.1. Coefficients et tests statistiques standards.....	44
a. Coefficient de corrélation r (et coefficient de détermination r^2).....	44

<i>b. Le coefficient de détermination ajusté r^2_{adj}</i>	45
<i>c. L'erreur type résiduel « s » (La déviation standard)</i>	45
<i>d. Le test de Fisher F</i>	46
<i>II.4.4.2. Pouvoir de prévision</i>	47
<i>a. Interne</i>	47
<i>b. Externe</i>	48
<i>Références bibliographiques</i>	50

Chapitre III: Etude qualitative et quantitative des relations structure-activité d'une série de composés phénoliques

III.1. Introduction.....	54
III.2. Présentation de la série de molécule, Méthodes et Matériel.....	54
III.2.1. Série de molécule.....	54
III.2.2. Méthodes et matériels.....	54
<i>III.3. Etude qualitative des relations structure/activité (SAR) dans les dérivés du phénol</i>	56
III.4. Les méthodes d'optimisation multiparamètres (MPO)	59
III.4.1. Règle de lipinski.....	59
III.4.2. Règle de Veber.....	61
III.4.3. Triangle d'or (Golden triangle).....	63
<i>III.5. Étude quantitative de la relation structure-activité d'une série de dérivés de phénol</i>	65
III.5.1. Descripteurs.....	65
III.5.2. Corrélation entre différent descripteurs.....	67
III.5.3. Equation du modèle QSAR.....	68
III.5.4. Validation de modèle.....	70
<i>Références bibliographiques</i>	74
<i>Conclusion générale</i>	77

Liste des figures et tableaux

<i>Liste des figures</i>		<i>Pages</i>
Chapitre I		
<i>Figure .1:</i>	<i>Structure des phénols simples.</i>	6
<i>Figure .2:</i>	<i>Structure générale des flavonoïdes.</i>	7
<i>Figure .3:</i>	<i>Différents classes des flavonoïdes.</i>	7
<i>Figure .4:</i>	<i>Structure de tannin hydrolysable et tannin condensé.</i>	9
<i>Figure .5:</i>	<i>Structures de l'acide Ellagique et l'acide Gallique.</i>	9
<i>Figure .6:</i>	<i>Structure de base des quinones.</i>	10
<i>Figure .7:</i>	<i>Déséquilibres de La balance des espèces pro oxydants et des systèmes de défense antioxydants.</i>	12
<i>Figure .8:</i>	<i>Origine extracellulaire et intracellulaire des radicaux libres dérivés de l'oxygène.</i>	13
<i>Figure .9:</i>	<i>Mécanismes de la peroxydation lipidique.</i>	15
Chapitre II		
<i>Figure .1:</i>	<i>Organigramme des méthodes utilisées en chimie quantique.</i>	30
<i>Figure .2:</i>	<i>Schéma de la méthodologie utilisée dans nos travaux.</i>	49
Chapitre III		
<i>Figure .1:</i>	<i>Structure chimiques des composés phénoliques avec les valeurs du potentiel redox expérimental (E_7).</i>	56
<i>Figure .2:</i>	<i>Sites donneur et accepteur du composé 20.</i>	59
<i>Figure .3:</i>	<i>Schéma sur les Règles de Lipinski.</i>	60
<i>Figure .4:</i>	<i>Triangle d'or (Golden triangle).</i>	64
<i>Figure .5:</i>	<i>Corrélation entre l'activité biologique expérimentale et prédite.</i>	72
<i>Figure .6:</i>	<i>Corrélation entre l'activité biologique expérimentale et résiduelle.</i>	73

<i>Liste des tableaux</i>		<i>Pages</i>
Chapitre I		
Tableau.1	<i>Structure des dérivés de l'acide benzoïque.</i>	6
Tableau.2	<i>Structure des dérivés de l'acide cinnamique.</i>	6
Tableau.3	<i>Structure des dérivés de coumarines.</i>	10
Tableau.4	<i>Potentiel redox.</i>	19
Chapitre III		
Tableau.1:	<i>Propriétés physicochimiques des dérivés du phénol.</i>	57
Tableau.2:	<i>Les règles de Lipinski des composés phénoliques.</i>	60
Tableau.3:	<i>Les règles de Veber des composés phénoliques.</i>	62
Tableau.4:	<i>Triangle d'or.</i>	63
Tableau.5	<i>Type des descripteurs moléculaires utilisés pour l'étude qualitative et quantitative des structure-activité.</i>	65
Tableau.6:	<i>Descripteurs utilisés dans le développement de modèle QSAR.</i>	66
Tableau.7:	<i>La matrice de corrélation entre différent descripteurs.</i>	67
Tableau.8:	<i>Les descripteurs sélectionnés par MLR.</i>	68
Tableau.9:	<i>Table de Fisher.</i>	69
Tableau.10:	<i>La matrice de corrélation du modèle..</i>	69
Tableau.11:	<i>Paramètres de validation croisée.</i>	70
Tableau.12:	<i>Valeurs expérimentales, prédites et résiduelles de E_7 des dérivées de phénol.</i>	71

Liste des Abréviations

¹O₂: Oxygène singulet.

3D : Trois dimensions.

ABTS: Acide 2,2'-azino-bis(3-éthylbenzothiazoline-6-sulphonique).

ACP: Analyse en Composantes Principales.

ADN: Acide Désoxyribonucléique.

AGPI: Acides Gras Poly-insaturés.

AM1: Austrin Model 1.

AMBER: Assisted Model Building with Energy Refinement.

A-T/C-G: Bases complémentaires (ADN).

B3LYP: Becke, Three-Parameter, Lee-Yang-Parr

B88: Becke 1988.

CAH: Classification Ascendante Hiérarchique.

CHARMM: Chemistry Harvard Macromolecular Mechanic.

CNDO: Complete Neglect of Differential Overlap.

CV: Validation Croisée.

DFT: Density Functional Theory.

DPPH : 2,2-diphényl-1-picrylhydrazyl.

E_{gap} : Gap énergétique (écart entre HOMO- LUMO).

ERO: Espèces Réactives Oxygénées.

E_T: Energie totale.

F: Indice de Fisher.

GROMOS : Groningen Molecular Simulation Program Package.

H₂O₂ : Peroxyde d'hydrogène.

HF : Hartree – Fock.

HOCl : Acide hypochloreux.

HOMO : Orbitale la plus haute occupée.

LCAO : Combinaison Linéaire d'Orbitale Atomique.

Log P : Coefficient de partage.

LOO : Leave-One-Out.

LUMO : Orbitale la plus basse vacante.

MLR : Régression Linéaire Multiple.

MM : La Mécanique Moléculaire.

MNDO : Modified Neglect of Differential Overlap.

MP n : Moller-Plesset perturbation theory.

MP3 : Parametric Method 3.

MR : Réfractivité Molaire.

MSE: Mean Square Error.

MV : Volume Moléculaire.

MW : Molecular Weight (Poids moléculaire).

NA : Nombre Avogadro.

NDDO : Neglect of Diatomic Differential Overlap.

NHA : Nombre d'Accepteur de liaison Hydrogène.

NHD : Nombre de Donneur de liaison Hydrogène.

NO : Monoxyde d'azote.

NROT : Nombre de Liaisons Rotatives

O₂⁻ : Anion superoxide.

OH : Fonction Hydroxyle.

OM-CLOA : Combinaison Lineaires des Orbitales Atomiques ou Moléculaires.

ONOOH : Nitroperoxyde.

OPLS : Optimised Potentials for Liquid Simulation.

P86 : Perdew 1986.

PBE : Perdew, Burke et Ernzerhof

PM3 : Parametrization Model 3.

PRESS: Predicted Residual Error Sum of Squares.

PSA : Superficie de la Surface Polaire.

PW86 et **PW91** : Perdew et Wang

Qi : La charge.

QSAR : Relation Structure Activité Quantitative.

R : Coefficient de Corrélation.

R² adj : Coefficient de Détermination Ajusté.

RO· : Radical alkoxy.

ROO· : Radical peroxy

ROOH : Hydroperoxydes.

ROS : Reactive Oxygen Species.

S : Erreur type résiduel.

SOD : Superoxyde Dismutase.

SSY : Sum of Squares in Y.

TEAC : Capacité antioxydante en équivalent Trolox.

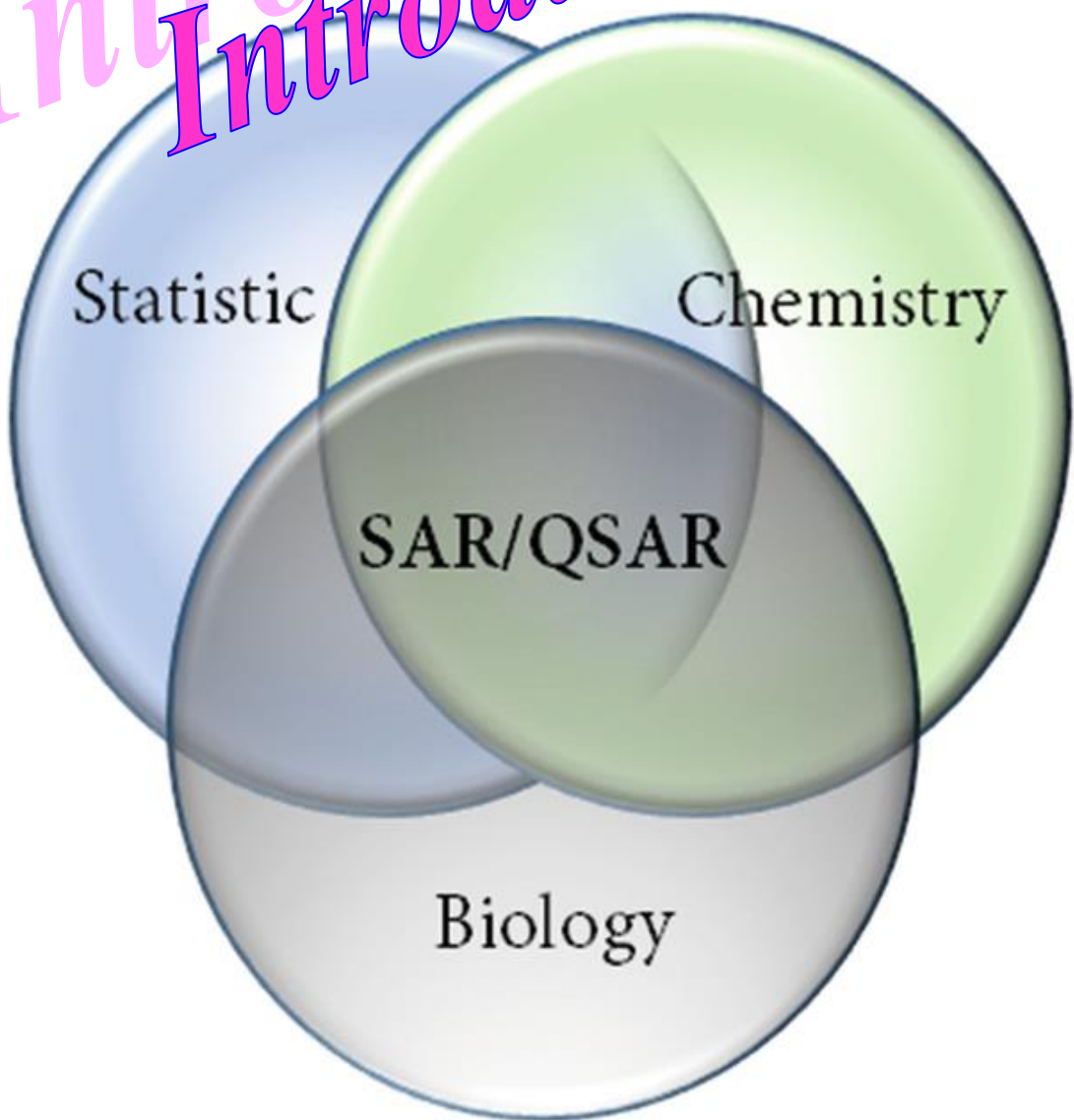
UV : Ultra Violet.

α : Polarisabilité.

μ : Moment dipolaire.

Introduction

Introduction



Découvrir de nouveaux médicaments de la manière la plus efficace et la moins coûteuse possible constitue un enjeu majeur pour les années à venir, Il s'agit en effet de trouver une molécule qui doit à la fois présenter des propriétés thérapeutiques particulières, et posséder le minimum d'effets secondaires indésirables.

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

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

Les antioxydants font l'objet de nombreuses recherches en vue dans une exploitation des métabolites secondaires en général et les composés phénoliques en particulier.

Les composés phénoliques sont des substances organiques caractérisés par la présence d'au moins un motif phénolique dans leurs structures chimiques. Il existe de nombreuses structures, ils sont divisés en plusieurs classes, les flavonoïdes, les acides phénoliques, ... ils sont distribués dans les fruits (ex : pomme, Kiwi, ...), les légumes (ex : brocoli, artichaut...), les huiles (ex : huile d'olive, huile d'argan). Ils sont des antioxydants puissants, leur activité anti-oxydante est attribuée : à la capacité de piéger les ERO, inhibition de la peroxydation des lipides et chélation les métaux et enfin ils peuvent inhiber les enzymes impliqués dans le stress oxydant [7].

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

simultanément et de tester leurs actions sur des cibles thérapeutiques et donner des résultats très attirants.

La modélisation QSAR (Quantitative Structure-Activity Relationship) c'est une méthode qui tente à identifier une relation quantitative entre l'activité mesurée expérimentalement et des descripteurs théoriques ou empiriques liés à la structure moléculaire. Elle a pu être mise en place dans les laboratoires et utilisée dans l'industrie pharmaceutique. L'objectif d'une modélisation QSAR est de trouver des modèles précis, applicables et robustes afin de trouver une relation entre la structure et l'activité dans un but de prédiction mais également d'interprétation.

L'activité anti-oxydante a été discutée dans plusieurs ouvrages [8-10], et de nombreux auteurs ont essayé de construire des modèles QSAR afin de corrélérer l'activité anti-oxydante des composés phénoliques avec des descripteurs physico-chimiques calculées.

L'objectif de ce travail est d'étudier des relations structure-activité sur les dérivés de phénol et d'élaborer un modèle QSAR robuste et fiable, en respectant toutes les étapes d'une étude QSAR pour expliquer et prédire l'activité anti-oxydante d'une série de ces composés en utilisant des descripteurs cibles et pertinents.

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

- ✚ Le premier chapitre intitulé : "généralités sur les composés phénoliques et l'activité antioxydant" est consacré à une étude bibliographique des composés phénoliques, les radicaux libres et les antioxydants.
- ✚ Le deuxième chapitre intitulé : " Bases théoriques " présente un aperçu général sur la base théorique des différentes méthodes de la modélisation moléculaire, méthode QSAR, son principe, son application et la méthode statistique, nous avons utilisé la régression linéaire multiple MLR.
- ✚ Le troisième chapitre nous présentons et nous discutons les résultats de notre étude qualitative et quantitative de la relation structure-activité d'une série de dérivés de phénol et aussi une application des méthodes de sélection (*Lipinski*, Veber et Golden triangle).

Enfin, nous finirons par une conclusion générale

Références bibliographiques

- [1] D. Harman, A Theory Based on Free Radical and Radiation Chemistry, *J.Geron*, 11, 298–300, 1956.
- [2] J. R. Roede, D. P. Jones, Reactive Species and Mitochondrial Dysfunction: Mechanistic Significance of 4-hydroxynonenal, Environmental and molecular mutagenesis, 51, 380-39, 2010.
- [3] K. Sachidanandam, S. C. Fagan, A. Ergul, Oxidative Stress and Cardiovascular Disease: Antioxidants and Unresolved Issues, *Cardiovas. Drug. Rev*, 23, 115-132, 2005.
- [4] D. Winyard, R. Blake, C. H. Evans, Reactive oxygen/nitrogen species and acute inflammation: A physiological process, *Free Radicals and Inflammation: Springer Basel AG*, 11-16, 2000.
- [5] S. Kawanishi, Y. Hiraku, M. Murata, S. Oikawa, Advances in Inorganic Chemistry: Homogeneous Biomimetic Oxidation Catalysis, *Am. cancer Soc*, 32, 822, 2002.
- [6] F. Zerargui, Activité antioxydante des extraits de racines *Tamus communis L.* et caractérisation des substances bioactives, Thèse-Doctorat, 2015.
- [7] R. Ksouri, W. Megdiche, A. Debez, H. Falleh, C. Grignon, C. Abdelly, Salinity effects on polyphenol content and antioxidant activities in leaves of the halophyte *Cakile maritima* *Physio. Biochem.*, 45, 244, 2007.
- [8] T. R. Sarkar, A. Middy, J. Dipnagar, *J. Mol. Model*, 18, 2621, 2012.
- [9] S. Ekins, Systems Approaches for Pharmaceutical Research and Development, *Computer Applications in Pharmaceutical Research and Development*, dt. John Wiley & Sons, Inc. Etats Unies, 2006.
- [10] M. Filipovic, Z. Markovic, *C. R. Chimie*, 18, 492-498, 2015.

Chapitre I

Chapitre I

I.1. Introduction

Les plantes ont une importance capitale pour la survie de l'homme et des différents écosystèmes. Elles renferment une part importante des composés qui interviennent dans l'ensemble des réactions enzymatiques ou biochimiques ayant lieu dans l'organisme [1]. On distingue ainsi deux groupes de métabolites: les métabolites primaires et les métabolites secondaires [2]

- ✓ Les métabolites primaires sont des molécules organiques qui se trouvent dans toutes les cellules de l'organisme d'une plante pour y assurer sa survie. Ces composés sont classés en quatre principaux groupes: les glucides, les protéines, les lipides et les acides nucléiques.
- ✓ Les métabolites secondaires sont des molécules ayant une répartition limitée dans l'organisme de la plante. Ils sont nécessaires à sa défense contre les agressions extérieures. Cependant, ils ne sont pas toujours nécessaires à la survie de la plante. Les produits du métabolisme secondaire qui sont émis en très faible quantité, sont d'une grande variété structurale (plus de 200000 structures définies). Ils sont répartis en trois grandes familles chez les végétaux, sur la base de leurs voies de biosynthèses: les composés phénoliques, les terpénoïdes et les alcaloïdes [3]. Dans cette étude, nous nous intéressons plus particulièrement aux composés phénoliques.

Plusieurs études ont montré que les composés phénoliques possèdent une activité biologique intéressante liée à leurs propriétés anti-oxydantes et anti-radicalaires [4], grâce à la mobilité de l'hydrogène phénolique. Ils sont capables de piéger les différents radicaux générés en permanence par l'organisme et ou causés par des agressions externes (polluants, alcool, tabac, infection, ...). Cette activité anti-oxydante permet à l'organisme de se protéger contre les agressions de l'oxygène qui sont à l'origine de plusieurs maladies. Ce qui suscite un intérêt croissant à étudier l'activité de ces composés dans la prévention et le traitement de nombreuses maladies comme, le cancer, les maladies cardiovasculaires, les inflammations, l'Alzheimer, parkinson et autres.

I.2. Composés phénoliques

Les composés phénoliques ou les polyphénols sont l'un des plus importants groupes rencontrés dans les plantes, ou leur distribution est considérablement large comprenant au moins 8000 différentes structures connues.

Les composés phénoliques sont des métabolites secondaires, d'un poids moléculaire élevé. Ils sont largement répandus chez les plantes et sont très importants dans la contribution à la couleur et au goût des fruits et végétaux [5]. Ils sont présents partout dans les racines, les tiges, les fleurs, les feuilles de tous les végétaux, les polyphénols prennent une importance croissante, notamment grâce à leurs effets bénéfiques sur la santé. Selon leurs caractéristiques structurales, ils se répartissent en une dizaine de classes chimiques, qui présentent toutes un point commun : la présence dans leur structure d'au moins un cycle aromatique à 6 carbones, lui-même porteur d'un nombre variable de fonctions hydroxyles (OH) [6, 7]. Ces espèces sont des monomères, des polymères ou des complexes dont la masse moléculaire peut atteindre 9000 [8]. Ils sont divisés en plusieurs catégories :

- ⇒ Les acides phénoliques.
- ⇒ Les flavonoïdes.
- ⇒ Les tanins condensés et hydrolysables.
- ⇒ -Les coumarines.
- ⇒ -Les quinones.

1.2.1. Classification des polyphénols:

1.2.1.1. Les phénols simples et acides phénoliques: Les acides phénoliques sont largement distribués dans les fruits. [9], les tiges et les feuilles des légumes. [10]. Ils sont solubles dans les solvants organiques polaires [11]. Les acides phénoliques sont divisés en trois classes:

- ⇒ Acides phénoliques simples
- ⇒ Acides phénoliques dérivés de l'acide benzoïque
- ⇒ Acides phénoliques dérivés de l'acide cinnamique.

- **Les phénols simples** : Les composés phénoliques simples comprennent un noyau aromatique avec un ou plusieurs substituant hydroxyles (figure .1). Tels que le phénol, catéchol, guaiacol, phloroglucinol [12].

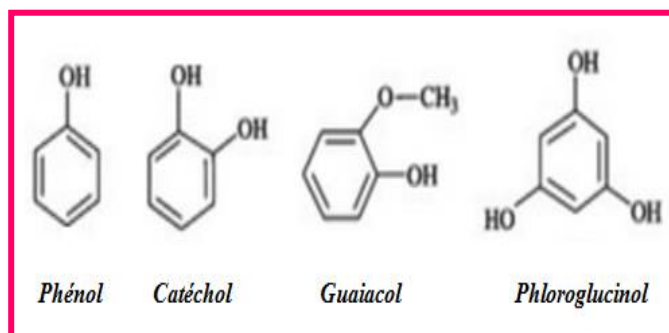
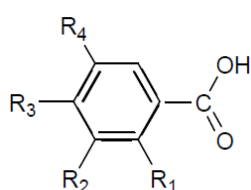


Figure.1: Structure des phénols simples.

- **Acide phénols dérivés d'acide benzoïque** : [13]

Tableau.1: Structure des dérivés de l'acide benzoïque.

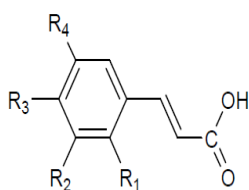


Acides Benzoliques	R1	R2	R3	R4
Ac. Salicylique	OH	H	H	H
Ac. P-hydroxy -benzoïque	H	H	OH	H
Ac. Gallique	H	OH	OH	OH
Ac. Protocatechique Ac. Vanilique	H	OCH ₃	OH	H

- **Acide phénols dérivés d'acide cinnamique**: [13]

Ces composés phénoliques comprennent une structure en commun constituée d'un noyau aromatique, un ou plusieurs substituant hydroxyles et une chaîne de 3 carbones avec un groupe carboxyle (tableau. 2).

Tableau.2: Structure des dérivés de l'acide cinnamique.



Acides cinnamiques	R1	R2	R3	R4
Ac. cinnamique	H	H	H	H
Ac. o-coumarique	OH	H	H	H
Ac. m-coumarique	H	OH	H	H
Ac. P-coumarique	H	H	OH	H
Ac. caféique	H	OH	OH	H

I.2.1.2. Les flavonoïdes: Ils sont caractérisés par la présence de deux cycles benzoïques (A et B) et d'un hétérocycle oxygéné (C) [14] (figure.2). Les flavonoïdes, peuvent empêchés les dommages oxydatifs par différentes mécanismes d'actions : soit par capture des radicaux hydroxyles, superoxydes, alkoxyles et peroxydes; soit par chélation des métaux (le fer et le cuivre) qui sont d'importance majeure dans l'initiation des réactions radicalaires ; soit l'inhibition des enzymes responsables de la génération des radicaux libres [15].

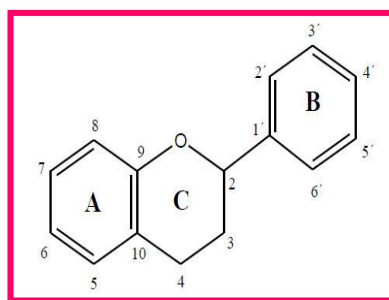


Figure. 2: Structure générale des flavonoïdes.

Les flavonoïdes peuvent être divisés en différentes classes : anthocyanidines ; flavonoles; flavones ; isoflavones ; flavanols ; flavanones ; isoflavanones (figure .3)[16] .

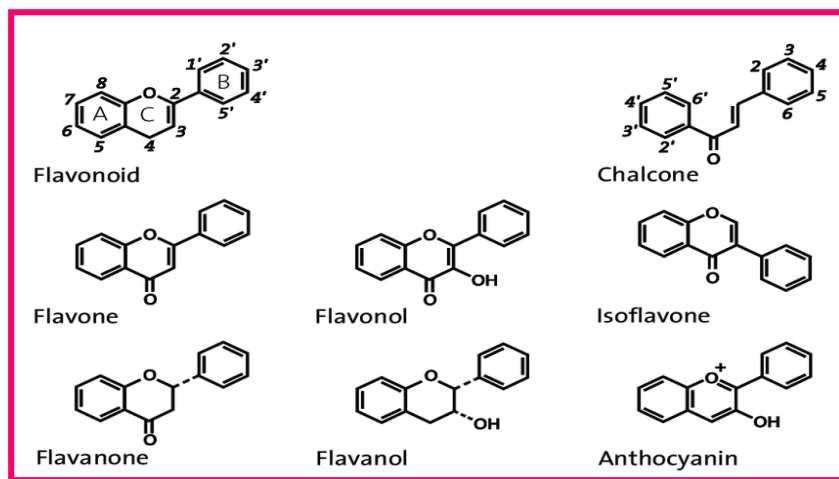


Figure.3: Différents classes des flavonoïdes.

I.2.1.3. Les tannins: Le mot Tanin est un terme provenant d'une pratique ancienne qui utilisait des extraits de plantes pour tanner les peaux d'animaux, autrement dit transformer une peau en cuir.

Les tannins sont caractérisés par une saveur astringente et sont trouvés dans toutes les parties de la plante : l'écorce, le bois, les feuilles, les fruits et les racines [17]. Ce sont des polymères (polyphénols) présents sous forme polymérisés, solubles dans l'eau et insolubles dans les solvants organiques apolaires [18], et ayant des poids moléculaires compris entre 500-3000 [19].

Concernant le pouvoir antioxydant des tannins, cette propriété est très remarquable due à leurs noyaux phénols et la présence des groupes di- ou trihydroxyles sur le cycle B et les groupes méta 5, 7 dihydroxyles sur le cycle A. Les tanins ont la capacité de fixer les protéines et par conséquent ont le pouvoir d'inhiber la peroxydation lipidique. Il existe deux classes de tanins

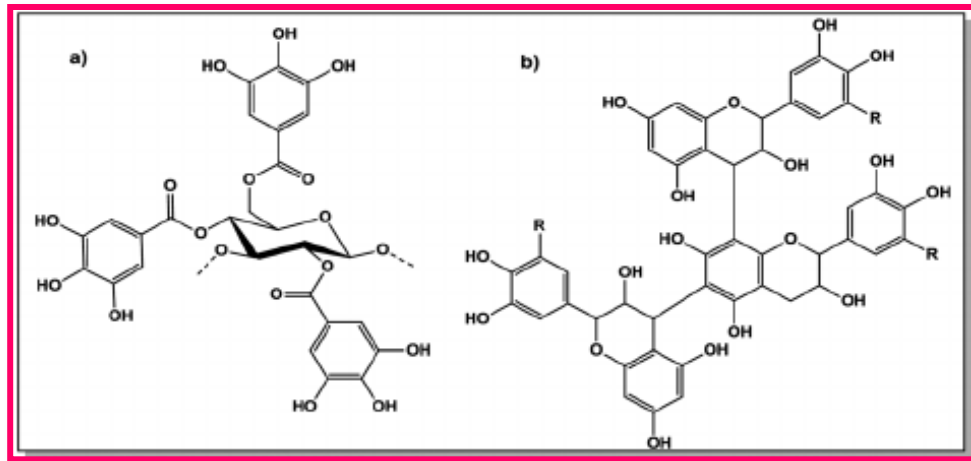
- **Les tanins condensés:**

Les tannins condensés ou les proanthocyanidines sont des polymères constitués d'unités flavane reliées par des liaisons entre les carbones C4, C8 et C6 (figure. 4). Ils ont la propriété de coaguler les protéines du derme, d'où leur utilisation dans le tannage des peaux [20].

- **Les tannins hydrolysables:**

Les tanins hydrolysables sont des hétéropolymères dont l'hydrolyse chimique libère un sucre, généralement le glucose, et un acide phénolique.

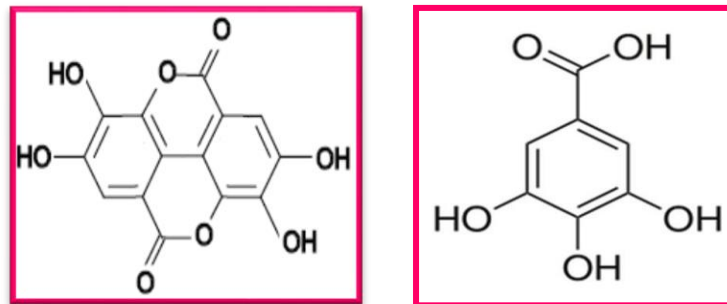
Le sucre est très généralement le D-glucose et l'acide phénol est soit l'acide gallique dans le cas des gallotannins soit l'acide ellagique dans le cas des tannins classiquement dénommés ellagitannins [11, 12] (figure.4 et 5).



a- structure de tannin hydrolysable

b- structure de tannin condensé

Figure.4: Structure de tannin hydrolysable et tannin condensé.



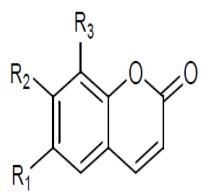
Structure de l'acide Ellagique

Structure de l'acide Gallique

Figure.5: Structures de l'acide Ellagique et l'acide Gallique.

1.2.1.4. Les coumarines: Les coumarines sont des hétérocycles oxygénés ayant comme structure de base le benzo-2-pyrone (figure.6). Les coumarines et leurs dérivés (tableau.3) ont la capacité de capter les ERO. Ils préviennent aussi la peroxydation lipidique de membranes [20]. Elles se trouvent dans la nature soit à l'état libre ou bien combiné avec des sucres.

Tableau.3: Structure des dérivés de coumarines



Coumarines	R1	R2	R3
<i>Umbelliférol</i>	H	OH	H
<i>Aesculol</i>	OH	OH	H
<i>Scopolétole</i>	OCH ₃	OH	H
<i>Fraxétole</i>	OCH ₃	OH	OH
<i>Daphnétole</i>	H	OH	OH

1.2.1.5. Les quinones: Les quinones sont des composés aromatiques contenant un noyau de benzène avec deux substitutions cétones, attachées aux dérivés phénoliques simples. Elles sont caractérisées par un motif 1,4-dicéto cyclohexa- 2,5-diénique (para-quinones) ou, éventuellement, par un motif 1,2-dicéto cyclohexa-3,5- diénique (ortho-quinones) [11] (figure.6). Elles ont des caractéristiques intéressantes dont le transport d'électrons dans les membranes mitochondriales.

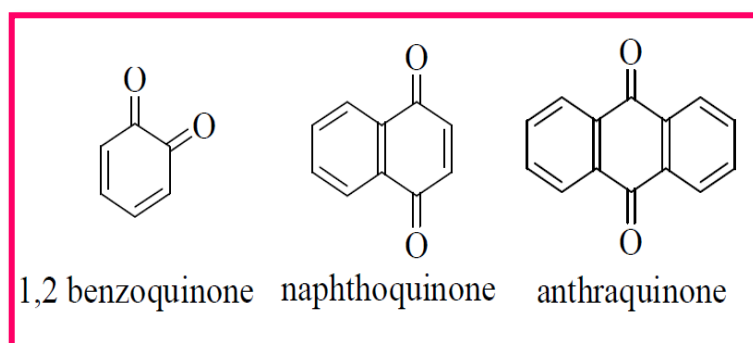


Figure.6: Structure de base des quinones.

1.2.2. Propriétés biologiques des polyphénols

Les recherches récentes sur les composés phénoliques sont très poussées en raison de leurs divers propriétés physiologiques comme les activités antiallergique, anti-atherogénique, anti-inflammatoire, hépatoprotective, antimicrobienne, antivirale, antibactérienne, anticarcinogénique, anti-thrombotique, cardioprotective et vasodilatatoire [21, 22]. Ces actions sont attribuées à leur effet antioxydant qui est due à leurs propriétés

redox en jouant un rôle important dans la destruction oxydative par la neutralisation des radicaux libres, piégeage de l'oxygène, ou décomposition des peroxydes [23].

Les effets bénéfiques des polyphénols intéressent particulièrement deux domaines: la phytothérapie et l'hygiène alimentaire [24]. D'après les études multiples attestant de l'impact positif de la consommation de polyphénols sur la santé et la prévention des maladies, les industriels commercialisent maintenant des aliments enrichis en polyphénols ou des suppléments alimentaires. De plus, leur activité antioxydante assure une meilleure conservation des denrées alimentaires en empêchant la peroxydation lipidique. Dans l'industrie cosmétique, les composés phénoliques trouvent leur application pratique en luttant contre la production des radicaux libres néfastes dans la santé et la beauté de la peau. En phytothérapie, même si certaines indications sont communes à plusieurs classes (les propriétés vasculoprotectrices, sont par exemple aussi bien attribuées aux flavonoïdes qu'aux anthocyanes, tanins et autres coumarines), chaque classe chimique semble être utilisée pour des bénéfices spécifiques [6].

I.3. Radicaux libres et Stress oxydatif

I.3.1. Radicaux libres

I.3.1.1. Définition

Un radical libre est défini comme toute molécule possédant un ou plusieurs électrons non appariés [25], cette molécule est très instable et réagit rapidement avec d'autres composants, essayant de capturer l'électron nécessaire pour acquérir la stabilité, une réaction en chaîne débute lorsqu'un radical libre attaque la molécule stable la plus proche en lui arrachant son électron, et la molécule attaquée devient elle-même un radical libre [26, 27].

I.3.2. stress oxydant

I.3.2.1. Définition

Le stress oxydant, qui est défini comme un déséquilibre entre les oxydants et les antioxydants [28, 29]. Il peut se produire en raison de la surproduction d'oxydants, la diminution de la défense antioxydante ou une combinaison de ces deux facteurs [30].



Figure.7 : Déséquilibres de La balance des espèces pro oxydants et des systèmes de défense antioxydants

1.3.2.2. Les maladies liées au stress oxydant

La plupart des maladies induites par le stress oxydant apparaissent avec l'âge, car le vieillissement diminue les défenses anti-oxydantes et augmente la production mitochondriale de radicaux. En faisant apparaître des molécules biologiques anormales et en sur-exprimant certains gènes, le stress oxydant sera la principale cause initiale de plusieurs maladies: cataracte, sclérose latérale amyotrophique, syndrome de détresse respiratoire aigu, oedème pulmonaire, vieillissement accéléré.

1.3.3. Différents types des radicaux libres

La majorité des espèces radicalaires dans le corps est dérivée de l'oxygène, certaines peuvent être dérivées de l'azote ($\text{NO}\cdot$) qui possède un électron partagé entre son atome d'azote et son atome d'oxygène [31].

Parmi toutes les espèces réactives oxygénées (ERO en English ROS: Reactive Oxygen Species), on distingue un ensemble restreint de ces composés qui jouent un rôle particulier en physiologie et que nous appelons les *radicaux primaires* (sont formes directement a partir de l'oxygène) à savoir : l'anion superoxyde ($\text{O}_2\cdot^-$), le radical hydroxyle ($\cdot\text{OH}$), le monoxyde d'azote ($\text{NO}\cdot$), le radical peroxyde ($\text{ROO}\cdot$) et le radical alkoxyde ($\text{RO}\cdot$).

Les autres radicaux libres, dits *radicaux secondaires* telles que l'oxygène singulet $^1\text{O}_2$, le peroxyde d'hydrogène (H_2O_2) et le nitroperoxyde (ONOOH), se forment par réaction de

ces radicaux primaires sur les composés biochimiques de la cellule (acide amine, acide gras polyinsaturé, protéine, ou acide nucléique) [32].

1.3.4. Origine de production des radicaux libre

Les radicaux libres nocifs sont produits dans l'organisme au cours du métabolisme normal. Cette production augmente en rapport avec l'élévation de la consommation d'oxygène [33]. Plusieurs mécanismes et systèmes responsables de la production de radicaux libres ont été identifiés jusqu'à présent, parmi eux nous citons (figure. 8) [34] :

- ⇒ des fuites d'électrons au niveau de la chaîne respiratoire de la mitochondrie [35].
- ⇒ des processus inflammatoires produits par les cellules phagocytaires activées [36 , 37].
- ⇒ d'exposition à des agressions de l'environnement, comme les agents infectieux, la pollution, les UV, la fumée de cigarette et le rayonnement [38].

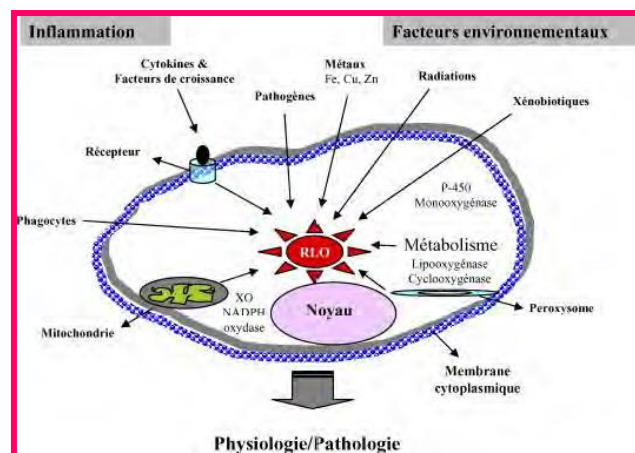


Figure.8: Origine extracellulaire et intracellulaire des radicaux libres dérivés de l'oxygène

1.3.5. Dommages causés par les radicaux libres

Les radicaux libres sont à l'origine de réactions en chaîne qui conduisent à des destructions cellulaires. Leurs structures cibles essentielles sont l'ADN, les membranes cellulaires, les lipides, les glucides et les protéines. Toutes ces altérations augmentent le risque de plus de 30 processus de différentes maladies [39]. Parmi les, nous citons, les

maladie d'Alzheimer [40, 41], de Parkinson [42], les maladies cardiovasculaires et déficience cardiaque [43] et vieillissement prématuré de la peau [44] et le cancer [45].

I.3.4.1. Action sur l'ADN : Les radicaux et en particulier .OH, s'attaquent aux différents nucléotides de l'ADN par modification des bases azotées, par rupture des brins de la chaîne nucléotidique, ou par déstabilisation de la forme géométrique de l'ADN en rompant les liaisons hydrogène entre bases complémentaires (pontage entre bases A-T ou C-G)*. Ils peuvent également être à l'origine de multiples lésions cellulaires en attaquant les membranes cellulaires. Toutes ces actions conduisent au développement de maladies telles que le cancer, le vieillissement cellulaire,...., etc.

I.3.4.2. Action sur les lipides : Les lipides sont une cible privilégiée des radicaux libres qui provoquent l'oxydation des acides gras poly-insaturés (AGPI) des phospholipides membranaires. Le phénomène d'auto-oxydation ou peroxydation lipidique (Figure.9) [46], consiste en l'attaque par un radical libre, d'origine exogène ou endogène, de dérivés lipidiques.

Cette réaction est à l'origine de la formation d'hydroperoxydes (ROOH) lipidiques. Les peroxydes lipidiques (ROO) peuvent aussi réagir avec d'autres composés cellulaires et être à l'origine de molécules très toxiques. Par exemple, ils peuvent réagir avec l'ADN et être à l'origine de substances mutagènes.

I.3.4.3. Action sur les glucides : Les radicaux libres peuvent aussi agir sur le glucose et générer des intermédiaires réactifs. Les dommages oxydatifs peuvent alors se propager via l'attaque de radicaux formés sur d'autres molécules.

I.3.4.4. Action sur les protéines : Les radicaux libres peuvent réagir avec les différents acides aminés et donc altérer la structure des protéines. Les fonctions de multiples enzymes, de récepteurs et de protéines de transport cellulaire peuvent ainsi être modifiées. C'est donc toute la machinerie cellulaire qui peut être affectée.

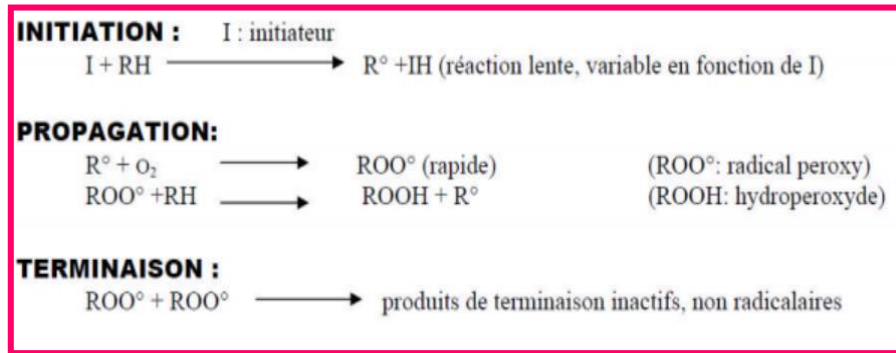


Figure. 9 : Mécanismes de la peroxydation lipidique

I.3.5. Systèmes de défense contre les ERO

Comme nous venons de le voir, les radicaux libres se produisent spontanément et d'une manière continue. Cependant, ils sont rigoureusement contrôlés par des antioxydants. Par conséquent, un antioxydant peut être défini comme toute substance capable, à concentration relativement faible, d'entrer en compétition avec d'autres substrats oxydables et ainsi retarder ou empêcher l'oxydation de ces substrats [47].

I.4. Les anti oxydants

Selon Halliwell et Gutteridge 1989 [48] un antioxydant est défini comme étant une substance qui, présente a faible concentration par rapport a un substrat oxydable, retarde ou prévient significativement l'oxydation de ce substrat. Les antioxydants par leurs pouvoirs d'offrir leurs propres électrons sont capables de ralentir ou d'inhiber l'oxydation des ERO et de détruire les radicaux libres propages dans l'organisme. Ils constituent une ligne de défense pour les cellules pour maintenir des concentrations non cytotoxiques de radicaux libres.

Une forme endogène d'antioxydants sous forme d'enzymes permette au corps de réguler constamment la détoxification des cellules. Une autre forme exogène apportée par l'alimentation est nécessaire pour maintenir l'équilibre.

I.4.1. Antioxydants enzymatiques

Différents systèmes enzymatiques sont connus pour leur rôle antioxydant [49] :

- ⇒ **Superoxyde dismutase SOD** qui catalyse la dismutation de l'anion superoxydes en oxygène moléculaire et peroxyde d'hydrogène.

- ⇒ **Catalase** une enzyme qui intervient dans la décomposition de la molécule peroxyde d'hydrogène en eau et oxygène moléculaire.
- ⇒ **Glutathion peroxydase**, avec l'apport du Sélénium la glutathion peroxydase peut réduire le peroxyde d'hydrogène et les hydroperoxydes lipidiques.

1.4.2. Antioxydants non enzymatiques

Constituent un groupe de molécules exogènes apportées essentiellement par l'alimentation. Certains de ces composés ont le pouvoir de complexer des métaux de transition et empêcher la production des radicaux libres ce sont des chélateurs métalliques.

D'autres antioxydants agissent en captant les électrons célibataires des radicaux libres et les transformer en molécules neutres ou ions stables ce sont des piègeurs. Différents vitamine, composés phénoliques et

- ⇒ **Vitamine C** : Un puissant antioxydant connu sous le nom d'acide ascorbique. La vitamine C intervient aussi dans la protection des vitamines liposolubles A et E et les acides gras polyinsaturés de l'oxydation.
- ⇒ **Vitamine E** : Connue sous le nom des α -tocopherols. Elle possède un pouvoir antioxydant très intéressant. La nature lipidique de l' α -tocopherol lui donne la capacité d'agir sur le site de la peroxydation lipidique membranaire. Le radical tocopheroxyle génère est peu réactif stabilise de par sa structure cyclique. La vitamine E agit conjointement avec la vitamine C pour donner lieu a une chaine de détoxications des radicaux libres.
- ⇒ **Composés phénoliques** : Ce sont des composés organiques qui font partie des métabolites secondaires des plantes. Ils sont caractérisés par la présence d'un ou plusieurs groupes phénoliques. Leur pouvoir antioxydant varie en fonction de leurs structures et leurs caractéristiques physico-chimiques.
- ⇒ **Métaux** : des métaux oligo-éléments comme le Cuivre le Manganese et le Zinc, sont indispensables pour nombreuses réactions enzymatiques inhibitrices des ERO. Leur présence joue un rôle important dans la protection des cellules contre les stress oxydatifs.

⇒ **Antioxydants synthétiques** : il existe d'autres composés synthétiques qui possèdent un pouvoir antioxydants. Ce sont des dérivés phénoliques synthétiques hautement actifs. Leur utilisation en excès peut être toxique.

1.4.3. Modes d'action anti-oxydante des phénols

Plusieurs modes d'actions sont connus pour l'activité anti-oxydante des composés phénoliques [7] :

- ⇒ Le piégeage direct des radicaux libres.
- ⇒ Chélation des ions métalliques.
- ⇒ L'inhibition des enzymes responsables de la génération des ERO.

1.4.3.1. Piégeage direct des radicaux libres : Par leur faible potentiel redox les composés phénoliques peuvent réduire rapidement des radicaux libres comme l'anion superoxyde $O_2^{\cdot-}$, les alkoxydes $RO\cdot$, les peroxydes $ROO\cdot$ Et les hydroxydes $\cdot OH$ par transfert d'hydrogène.



1.4.3.2. Chélation des ions métalliques : Les acides phénoliques et les flavonoïdes sont connus pour leurs capacité à chélater les ions des métaux de transition comme le cuivre et le fer, en formant des complexes ou des composés de coordination avec ces ions. Les travaux mènes par Van Acker et al [50].

Ont mis en évidence 3 sites essentiels pour les chélationes comme points d'attache pour chez les flavonoïdes a savoir :

- La structure catechol (3,4 hydroxy) sur le cycle B
- La structure 4-ceto, 3-hydroxy sur le cycle C
- La structure 4-ceto et 5-hydroxy entre les cycles A et C [51].

1.4.3.3. L'inhibition des enzymes responsables de la génération des ERO: Les composés phénoliques révèlent aussi un pouvoir inhibiteur envers une grande variété d'enzyme. Cos et al. en 1998 [52] ont établi une relation entre la structure de quelques flavonoïdes

et leurs capacité d'inhiber la génération des radicaux libres superoxydes par la formation de complexes enzyme-inhibiteur avec la xanthine oxydase.

1.4.4. Méthodes de quantification de l'activité anti-oxydante

Plusieurs méthodes sont disponibles pour mesurer l'activité antioxydante des aliments et les systèmes biologiques [45, 53]. Elles peuvent être classées en deux groupes selon deux mécanismes : soit par le transfert d'atome d'hydrogène, soit par le transfert d'un simple électron [54, 55].

Les techniques du premier groupe sont employées pour évaluer la peroxydation lipidique en utilisant un substrat lipidique ou lipoprotéique. La quantification de cette propriété est exprimée par la mesure du degré d'inhibition de l'oxydation [56].

Alors, les méthodes du deuxième groupe sont celles qui interviennent dans la mesure de l'habilité du piégeage des radicaux libres. Elles comportent le balayage du peroxyde d'hydrogène (H_2O_2), de l'acide hypochloreux ($HOCl$), de l'hydroxyle ($\cdot OH$), des anions superoxyde ($O^{\cdot -2}$), du peroxyde ($ROO\cdot$) et de l'oxyde nitrique ($NO\cdot$) [54]. Parmi ces techniques, nous citons :

- ⇒ **Test DPPH** (Elle est basée sur la réduction du radical libre DPPH. C'est un radical stable centré sur l'Azote. Sa réduction par l'antioxydant donneur d'atome H, notamment les composés phénoliques, conduit à la formation du DPPH-H (composé incolore) et au radical AH) [57].
- ⇒ **Test TEAC** (basée quant à elle sur la capacité de l'antioxydant à inhiber le cation radical ABTS.⁺ comparativement à un antioxydant de référence, le trolox de structure similaire à la vitamine E. le ABTS.⁺ de coloration bleu-vert en contact avec un donneur de H. conduit au cation incolore ABTS) [58] .
- ⇒ **Test ORAC** (Capacité d'absorbance du radical de l'oxygène) [59].
- ⇒ **Potentiel Redox** : Le potentiel redox est une mesure électrochimique du pouvoir antioxydant exprimée en volt (tableau .4). C'est une technique utilisée dans l'évaluation du pouvoir antioxydant du phénol. En présence de radicaux libres

(agents oxydants) les composés phénoliques sont convertis en radicaux phénoxyles.

Tableau.4: Potentiel redox [60]

Composé	E ₀ (v)	Famille
myricétine	0,712	flavonols
quercétine	0,872	
kaempférol	0,929	
épicatéchine	0,937	flavanols
catéchine	0,957	
naringénine	0,961	flavanone

Le pouvoir antioxydant décroît avec l'augmentation du potentiel redox.

Références bibliographiques

- [1] A. Benayahoum, Analyse quantitative par Chromatographie Liquide à Haute Performance de molécules à intérêt thérapeutique dans les aliments et étude de leur relation structure/activité antioxydante, Thèse de Doctorat en chimie, 2013.
- [2] T. Hartmann, From waste products to ecochemicals: Fifty years research of plant secondary metabolism. *Photochemistry* 68, 2831 – 2846, 2007.
- [3] S. Krief, Métabolites secondaires des plantes et comportement animal: surveillance sanitaire et observations de l'alimentation des chimpanzés en Ouganda. Activités biologiques et étude chimique de plantes consommées, Museum national d'histoire naturelle, paris, 2003.
- [4] J. Kanner, E. Frankel, R. Granit, B. German, J. E. Kinselva, 42, 64, 1994.
- [5] F. Manolaraki, Propriétés anthelminthiques du sainfoin: Analyse des facteurs de variations et du rôle des composés phénoliques impliqués, Institut National Polytechnique de Toulouse, 2011.
- [6] T. Hennebelle, S. Sahpaz, F. Bailleul, Polyphénols végétaux, sources, utilisations et potentiel dans la lutte contre le stress oxydatif. *Phytothérapie*, 1, 3-6, 2004.
- [7] Y. Yousfi, Etude QSAR de l'activité anti-oxydante d'une série de composés phénoliques, Thèse de Master en Chimie Théorique et Modélisation, 2017.
- [8] J. B. Harbone, Introduction to Ecological Biochemistry, 4th Ed; Academic Press: London, 1993.
- [9] C. Garcia-vegura, P. Bridle, F. Ferreres, B. F. Tomas, Influence of variety, maturity and processing on phenolic compounds of apricot juices and jams. *Z Lebensm Unters Forsch.*199, 433-436, 1994.
- [10] L. W. Morton, A. Amsha, R. Caccetta, I. B. Puddy, K. D. Croft, Chemistry and biological effects of dietary phenolic compounds: Relevance to cardiovascular disease. *Clin Exp Pharmacol PHysiol.* 27, 152-59, 2000.
- [11] J. Bruneton, Pharmacognosie: Phytochimie, Plantes médicinales. 2ème édition, Lavoisier Techniques & Documentation, Paris, 1993.
- [12] M.M. Cowan, Plant Products as Antimicrobial Agents. *Clin. Microbiol Re*, 4, 564-582, 1999.
- [13] S. N. Brianceau, Vers une amélioration quantitative et qualitative de l'extraction des composés phénoliques du marc de raisin rouge à l'aide d'électrotechnologies, Compiègne, 2015.
- [14] A. El Kabouss, Z. Charrouf, H. Oumzil, M. Faid, D. Lamnaouer, Y. Miyata, and K. Miyahara, Caractérisation des flavonoïdes des feuilles d'Arganier (*Argania spinosa* (L.) Skeels, Sapotaceae) et étude de leur activité antimicrobienne. *Revue Marocaine des Sciences Agronomiques et Vétérinaires*, 21, 157-162, 2011.

- [15] N. Benhammou, Activité antioxydante des extraits des composés phénoliques de dix plantes médicinales de l'Ouest et du Sud-Ouest Algérien, Thèse de Doctorat en Biologie, 2011.
- [16] B. H. Havsteen, The biochemistry and medical significance of the flavonoids, *Pharmacol. Therapeut*, 96, 67–202, 2002.
- [17] A. Scalbert, Antimicrobial properties of tannins. *Phytochemistry*, 30, 3875-3883, 1991.
- [18] A. Aguilera-Carbo, C. Augur, L. A. Prado-Barragan, E. Favela-Torres, and C. N. Aguilar, Microbial production of ellagic acid and biodegradation of ellagitannins, *Applied microbiology and biotechnology*, 78(2), 189-199. 2008.
- [19] J. Doat, Les tanins dans les bois tropicaux, *Bois et Forêts des Tropiques*, 182, 37-54, 1978.
- [20] J. Bruneton, Pharmacognosie : Phytochimie, Plantes Médicinales. 3ème édition, Lavoisier Techniques & Documentation, Paris. 1999.
- [21] E. Middleton, C. Kandaswami, T. C. Theoharides, The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease and cancer. *Pharmacol Rev*, 52, 673-839, 2000.
- [22] R. Ksouri, W. Megdiche, A. Debez, H. Falleh, C. Grignon, C. Abdelly. Salinity effects on polyphenol content and antioxidant activities in leaves of the halophyte *Cakile maritima* Plant, *Physiol Bioch*, 45, 244-249, 2007.
- [23] R. J. Nijveldt, E. Nood, D. E. Hoorn, P. G. Boelens, K. Norren, P. Leeuwen, Flavonoids: A review of probable mechanisms of action and potential applications. *Am. J. Clin Nutr*, 74, 418–425, 2001.
- [24] L. P. Leong, G. Shui, An investigation of antioxidant capacity of fruits in Singapore markets. *Food Chem*, 76, 69-75, 2002.
- [25] B. Jacques, and R. André, Biochimie métabolique, Ed ellipses, Paris, 2004.
- [26] M. Martinez-Cayuela, Oxygen free radicals and human disease, *Biochem*, 77, 147-161, 1995.
- [27] N. Belkheiri, Dérives phénoliques à activités antiathérogènes, Doctorat en Chimie-Biologie-Santé, 2010.
- [28] V. D. Ratnam, D. D. Ankola, V. Baradwaj, D. K. Sahana, M. N. Ravi Kumar, Role of antioxidants in Sciences, 81, 895-905, 2006.
- [29] A. Chelabi, Composition phénolique et activité antioxydante d'extraits de feuilles de l'ivier, Thèse de Master en agroalimentaire et contrôle de qualité, 2019.
- [30] A. Ece, F. Gurkan, F. Celik, M. Boşnak, S. Yel, H. Balik, O. Erel, Paraoxonase total antioxidant activity and peroxide levels in marasmic children: relationships with leptin . *Clinical Biochemistry Journal*, 40, 634-639, 2007.

- [31] C. E. Benammar, Effets antioxydants et immuno-modulateurs d'une plante médicinale nord africaine: étude des différents extraits, Thèse de Doctorat en Biologie moléculaire et cellulaire, Tlemcen, 84, 2011.
- [32] A. Favier, Le stress oxydant. Intérêt conceptuel et expérimental dans la compréhension des mécanismes des maladies et potentiel thérapeutique. *L'actualité chimique*, 108-115, 2003.
- [33] E. Gauche, C. Hausswirth, Stress oxydant, complémentation nutritionnelle en antioxydants et exercice. *Science & Motricité*, 58, 43-66, 2006.
- [34] A. Guillouty, Plantes médicinales et antioxydants, Thèse de Doctorat en pharmacie, 2016.
- [35] B. Aurousseau, Les radicaux libres dans l'organisme des animaux d'élevage: conséquences sur la reproduction, la physiologie et la qualité de leurs produits, *INRA Prod. Anim*, 15, 67-82, 2002.
- [36] H. Milane, La quercétine et ses dérivés: molécules à caractère prooxydant ou capteurs de radicaux libres, études et applications thérapeutiques. Thèse de doctorat en Sciences domaine, Pharmacochimie, Université Louis Pasteur Strasbourg I, 2004.
- [37] P. Van Antwerpen, Contribution à l'étude du pouvoir antioxydant de divers agents d'intérêt thérapeutique: Ciblage du système myeloperoxydase / peroxyde d'hydrogène /chlorure. Thèse de doctorat en Sciences Pharmaceutiques, Académie universitaire Wallonie-Bruxelle, 2006.
- [38] F. Tamer, Free radicals, Types, sources and damaging reactions, *Internal Medicine Articles*. (<http://www.doctorslounge.com/primary/articles/>),2003.
- [39] O. I. Aruoma, Free radicals, oxidative stress and antioxidants in human health and disease. *J. Am. Oil Chem. Soc*, 75, 199–212, 1998.
- [40] M. A. Smith, G. Perry, P.L. Richey, L. M. Sayre, V. E. Anderson, M. F. Beal, et al, Oxidative damage in Alzheimer's, *Nature*, 382, 120, 1996.
- [41] A. R. Smith, S. V. Shenvi, M. Widlansky, J. H. Suh, T. M. Hagen, Lipoic acid as a potential therapy for chronic diseases associated with oxidative stress. *Curr. Med. Chem*, 11, 1135– 1146, 2004.
- [42] J. L. Bolton, M. A. Trush, T. M. Penning, G. Dryhurst, T. J. Monks, Role of quinones in toxicology, *Chem. Res. Toxicol*, 13, 135, 2000.
- [43] P. Jha, M. Flather, E. Lonn, M. Farkouh, S. Yusuf, The antioxidant vitamins and cardiovascular disease. A critical review of epidemiologic and clinical trial data. *Ann. Intern Med*, 123, 860, 1995.
- [44] S. R. Georgetti, R. Casagrande, V. M Di Mambro, E. C. S. Azzolini Ana, J. V. Fonseca Maria, Evaluation of the antioxidant activity of different flavonoids by the Chemi luminescence Method. *AAPS Pharm Sci*, 5, 1-5, 2003.

- [45] S. S. Ali, N. Kasoju, A. Luthra, A. Singh, H. Sharanabasava, A. Sahu, U. Bora, Indian medicinal herbs as sources of antioxidants. *Food Res Int*, 41, 1–15, 2008.
- [46] H. Louanas, Prédiction théorique de l'activité antioxydante de composés d'espèces naturelles, Thèse de Magister en physique chimie théorique chimie informatique, 2011.
- [47] M. M. Berger, Manipulations nutritionnelles du stress oxydant: état des connaissances. *Nutrition Clinique et Métabolisme*, 20, 48-53, 2006.
- [48] B. Halliwell, J. M. C. Gutteridge. *Anal. Biochem*, 1, 215, 1987.
- [49] F. M. Hocine, M. A. GORINE, Evaluation de l'exposition au plomb et cadmium et impact sur quelques paramètres du statut oxydant/anti oxydant chez les ouvriers exposés aux fumées de soudage, Master en Biologie, 2017.
- [50] S. A. Van Acker, D. G. Van Den Berg, D. H. Griffioen, W. P. Van Bennekom, W. J. Van Der Vijgh, A. Bast, *Free Radical. Biol. Med*, 20, 331, 1996.
- [51] C. A. Rice-Evans, N. J. Miller, G. Paganga. *Free Radic. Biol. Med*, 20, 93, 1996.
- [52] P. Cos, L. Ying, M. Calomme, J. H. Hu, K. Cimanga, B. Van Poel, L. Pieters, A. J. Vlietinck, D. Vanden Berghe, Structure-activity relationship and classification of flavonoids as inhibitors of xanthine oxidase and superoxide scavengers, *J. Nat. Prod*, 61, 71, 1998.
- [53] R. Scherer, H. T. Godoy, Antioxidant activity index (AAI) by the 2, 2-diphenyl-1-picrylhydrazyl method, *Food Chem*, 112, 654–658, 2009.
- [54] C. Sanchez-Moreno, Review: Methods used to evaluate the free radical scavenging activity in food and biological systems, *Food Sci Tech Int*, 8, 121-137, 2002.
- [55] D. Huang, B. Ou, R. I. Prior, The chemistry behind antioxidant capacity assays, *J. Agric. Food Chem*, 53, 1841-1856, 2005.
- [56] C. Sanchez-Moreno, J. A. Larrauri, F. Saura-Calixto, A procedure to measure the antiradical efficiency of polyphenols, *J. Sci Food. Agric*, 76, 270–276, 1998.
- [57] E. Vamanu, S. Nita, Antioxidant Capacity and the Correlation with Major Phenolic Compounds, Anthocyanin, and Tocopherol Content in Various Extracts from the Wild Edible *Boletus edulis* Mushroom, BioMed Research International, 2013.
- [58] N. J. Miller, C. Rice-Evans, M. J. Davies, V. Gopinathan, A. Milner, A novel method for measuring antioxidant capacity and its application to monitoring the antioxidant status in premature neonates. *Clin Sci*, 84, 407–412, 1993.
- [59] G. H. Cao, H. M. Alessio, R. G. Cutler, R.G. Oxygen-Radical Absorbency Capacity Assay for antioxidants, *Free Radical Biol Med*, 14, 303-311, 1993.
- [60] A. Fougerousse, Le potentiel redox E et le rH₂ deux approches de l'évaluation de la force des oxydants et des réducteurs, *Bull.Un.Phy.*, 781, 319, 1996.

Chapitre II

Chapitre II

II.1. Introduction

Le Développement important des moyens informatiques (mémoire plus importante, calcul scientifique, infographie, ...) au cours de ces dernières années a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules. Ainsi un nouvel outil est apparu permettant à tout chimiste de mieux appréhender les problèmes liés au domaine moléculaire "*la Modélisation Moléculaire*".

La modélisation moléculaire implique l'utilisation de méthodes de calculs théoriques (mécanique moléculaire, dynamique moléculaire, mécanique quantique ab-initio ou semi-empirique, ...) permettant de déterminer la représentation graphique de la géométrie et d'évaluer les propriétés physico-chimiques de la molécule étudiée. La modélisation moléculaire permet d'interpréter des phénomènes physico-chimiques, de suggérer des nouvelles expériences et d'analyser ainsi des résultats d'une façon plus critique que les expériences classiquement utilisées, elle est donc complémentaire à l'expérience [1, 2].

Les premiers essais de modélisation des relations structure-activité ont commencé à la fin du 19^{ème} siècle, lorsque *Crum-Brown* et *Frazer* ont postulé que l'activité biologique d'une molécule est une fonction de sa constitution chimique [3-5]. Mais ce n'est qu'au début des années 60 que les travaux de *Corwin Hanch* ont proposé un modèle mathématique reliant l'activité biologique à la structure chimique [6].

Aujourd'hui, l'utilisation de QSAR n'a cessé de progresser [7, 8]. Elle est devenue indispensable en chimie pharmaceutique et pour la conception de médicaments, notamment dans le cas où la disponibilité des échantillons est limitée ou les mesures expérimentales sont dangereuses, longues et chères [9, 10].

Sans l'utilisation de grands instruments analytiques, les résultats des études QSAR peuvent fournir des informations utiles pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent être alors utilisées dans la prédiction des activités biologiques de nouveaux composés ainsi que dans la conception de nouvelles structures [11].

II.2. Méthodes de calculs

II.2.1. Méthodes quantiques

II.2.1.1. Méthodes semi empiriques

Les méthodes semi empiriques sont des techniques de résolution de l'équation de Schrödinger de systèmes à plusieurs électrons. Contrairement aux *ab initio*, les méthodes semi empirique utilisent de données ajustées sur des résultats expérimentaux afin de simplifier les calculs. La longueur et la difficulté des calculs est en grande partie due aux intégrales biélectroniques qui apparaissent aux cours du processus de résolution.

Plusieurs méthodes semi-empiriques ont été développées telles que CNDO, MNDO, NDDO, [12] ...etc. Les méthodes les plus connues sont la méthode AM1 "*Austin Model 1*" [13] et la méthode PM3 "*Parametrization Model 3*" [14] qui offrent des avantages réellement significatifs concernant la rapidité du calcul et une plus grande flexibilité dans la dimension du système à étudier.

II.2.1.2. Méthodes *ab initio*

Le but ultime en chimie quantique est de pouvoir résoudre l'équation de Schrödinger, ce qui revient à chercher les énergies. Les méthodes *ab-initio* sont des méthodes non empiriques, il n'y'a pas d'approximation à faire sauf celle de Born Oppenheimer et l'approximation OM-CLOA [15]. Dans les méthodes *ab-initio*, toutes les particules (noyaux et électrons) sont traitées explicitement.

Cette méthode permet d'obtenir des résultats fiables et précis pour des molécules de petite taille. Cependant il est également possible d'obtenir de bons résultats à partir de méthodes quantiques moins exigeantes en moyens informatiques : les méthodes basées sur la théorie de la fonctionnelle de la densité (DFT) ont atteint un niveau de fiabilité compétitif avec des méthodes *ab initio* incluant la corrélation électronique, tout en permettant de réduire considérablement la charge de calculs.

II.2.1.3. La théorie de la fonctionnelle de la densité DFT

Au cours des trente dernières années, la Théories de la Fonctionnelle de le Densité DFT (Density Functional Theory), venant compléter les méthodes HF et Post HF, s'est révélée comme particulièrement performante. Cette approche présente plusieurs avantages qui facilitent son utilisation dans différents domaines, surtout en chimie et sciences des matériaux.

- D'abord, elle inclut dans son formalisme une grande part de la corrélation électronique.
- L'aspect mono-électronique permet de donner une interprétation chimique de la fonction d'onde.

Dans le cas des méthodes Hartree-Fock, il s'agissait d'exprimer l'énergie du système en fonction de la fonction d'onde Ψ , alors que dans le cas de la DFT, l'énergie sera une fonctionnelle de la densité électronique ρ du système [16].

a. Objectif

L'objectif principal de la théorie de la fonctionnelle de la densité est de remplacer la fonction d'onde multiélectronique par la densité électronique en tant que quantité de base pour les calculs. Alors que la fonction d'onde multiélectronique dépend de $3N$ variables (où N est le nombre total de particules du système), la densité est seulement fonction de trois variables ; il s'agit donc d'une quantité plus facile à traiter tant mathématiquement que conceptuellement.

Le principe de la DFT consiste en une reformulation du problème quantique à N corps en un problème monocorps (ou, à la rigueur, bi-corps si l'on considère les problèmes de spin) avec pour paramètre la densité électronique. L'idée centrale de la DFT est que la seule densité électronique de l'état fondamental du système détermine entièrement les valeurs moyennes des observables, comme l'énergie [17].

La DFT trouve ses origines dans le modèle développé par Llewellyn Thomas et Enrico Fermi à la fin des années 1920. Néanmoins il faudra attendre le milieu des années 1960 et les contributions de Pierre Hohenberg, Walter Kohn et Lu Sham pour que soit établi le formalisme théorique sur lequel repose la méthode actuelle [18].

b. Approximations en DFT

b.1. Approximation locale de la densité (LDA)

La fonctionnelle d'échange-corrélation est une grandeur physique dépendante de plusieurs paramètres comme la densité électronique, l'énergie cinétique, etc.

Afin d'obtenir une bonne approximation de cette fonction, elle doit reproduire au mieux les caractéristiques physiques du trou d'échange-corrélation. Elle se base sur les propriétés d'un gaz homogène d'électrons ; la densité électronique et la fonction d'onde sont considérées localement comme constantes.

Par le concept de LDA il est possible d'estimer l'énergie d'échange-corrélation d'un système inhomogène en utilisant les résultats d'un gaz homogène d'électrons de densité égale à la densité locale d'un système inhomogène. L'énergie de corrélation d'un gaz uniforme d'électrons n'est pas déterminée exactement. Cependant, certaines expressions approchées permettent de décrire les densités électroniques faibles ou élevées. Cette approximation considérée localement permet de décrire correctement les systèmes ayant une densité électronique variant faiblement dans l'espace [19].

b.2. L'approximation du gradient généralisé (GGA)

Ainsi une amélioration peut donc être apportée sur ce point en utilisant des corrections qui prennent en compte les variations locales de la densité (gradient). Ces approximations sont définies comme semi-locales et sont regroupées sous le nom de GGA (Generalized Gradient Approximation). Les variations semi-locales impliquent à la fois l'échange et la corrélation.

Ces fonctionnelles améliorent dans de nombreux cas les résultats structuraux et énergétiques. Les plus utilisées sont celles proposées par Perdew (P86), [20] Becke (B88), [21] Perdew et Wang (PW86 et PW91) [22] et par Perdew, Burke et Ernzerhof (PBE) sous sa forme initiale [23] ainsi que sous sa forme modifiée. La fonctionnelle d'échange corrélation reste la seule approximation liée à la DFT, même si sa qualité s'améliore constamment avec par exemple l'apparition de fonctionnelles non locales [24]. Mais d'autres approximations sont nécessaires pour l'étude de nos systèmes.

b.3. Fonctionnelles hybrides pour le calcul moléculaire

Les fonctionnelles hybrides, quant à elles, reprennent l'énergie d'échange calculée par méthode HF qui peut être calculée de façon exacte contrairement aux fonctionnelles approximatives d'échange disponible en DFT. Elles reprennent également une partie des énergies d'échange et de corrélation de fonctionnelles LDA et GGA, le tout pondéré par des coefficients bien déterminés.

c. La fonctionnelle B3LYP

La fonctionnelle B3LYP est comptée parmi les fonctionnels hybrides les plus utilisées pour l'étude de systèmes moléculaires.

d. Bases d'orbitales atomiques

Les orbitales moléculaires obtenues par les méthodes *ab-initio* ou DFT sont exprimés comme des combinaisons linéaires d'orbitales atomiques (LCAO) [25]. Une orbitale atomique de type 1s, 2s, 2px, 2py peut être représentée mathématiquement par une ou plusieurs fonctions (proches des solutions de l'équation de Schrödinger pour l'atome H). L'ensemble de ces fonctions pour tous les atomes d'une molécule est appelé «base d'orbitales atomiques». Cette dernière est nommée « base simple zêta » si chaque orbitale atomique d'un atome est représentée par une seule fonction, « base double zêta » si elle est représentée par deux fonctions.

Plusieurs catégories de bases peuvent être distinguées, les plus développées sont :

1- Les bases minimales : dans ces bases, les orbitales de coeur et de valence comprennent le même nombre de primitives gaussiennes. Les résultats obtenus sont moins coûteux en temps de calcul mais restent insuffisants pour la recherche et les comparaisons avec les analyses expérimentales.

2- Les bases de Pople : [26] ce sont des bases à valence séparée qui ont été introduites par l'équipe de John Pople et sont typiquement en forme de « X-YZG ». Dans ce cas, X représente le nombre de gaussienne utilisé pour décrire chaque orbitale atomique de coeur. Le Y et le Z indiquent que les orbitales de valence sont composées chacune de deux fonctions, la première étant composée d'une combinaison linéaire de Y fonctions gaussiennes primitives, l'autre d'une combinaison linéaire de Z fonctions gaussiennes.

Dans ce cas, la présence de deux nombres après le trait d'union indique que la base est une base double zêta à valence séparée. La différence entre deux fonctions Y et Z réside essentiellement dans les exposant ξ des fonctions gaussiennes, plus grands pour Y (la fonction Y est plus contractée que Z). Les bases les plus utilisées dans cette catégorie sont la 6-31G et la 3-21G.

La base 6-31G (d, p) est celle utilisée pour tous les calculs de cette mémoire.

Base 6-31G:

La base 6-31G est constituée de deux familles de fonctions :

- (i) une fonction de base approchée par 6 gaussiennes pour décrire les électrons de cœur.
- (ii) Deux fonctions de base approchées respectivement par 3 et 1 gaussiennes pour décrire les électrons de valence.

La qualité des bases peut encore être améliorée en ajoutant :

- ⇒ des fonctions de polarisation.
- ⇒ des fonctions diffuses.

Pour une plus grande flexibilité on peut encore rajouter des fonctions de polarisation (déformation du nuage électronique, et des distorsions dues à l'environnement). La dénomination la plus ancienne est l'ajout d'un astérisque sur la base en question (par exemple 6-31G*), et dans une désignation plus récente, le caractère de la fonction ajoutée est explicitement donné : 6-31G(d). La base 6-31G(d) signifie ainsi qu'un jeu de fonctions « d » a été ajouté à tous les atomes (sauf H) dans la molécule, alors que 6-31G** ou 6-31G(p, d) signifie qu'un jeu de fonctions « p » a été ajouté aux hydrogènes et que des fonctions « d » ont été ajoutées aux autres atomes.

Un autre type de fonctions est indispensable à inclure dans la base d'orbitale atomique chaque fois que le phénomène physique décrivant la propriété étudiée nécessite une bonne description de l'espace situé au-delà des orbitales de valence (espace diffus). Ce sont les fonctions diffusées, qui augmentent la taille du nuage électronique. Pour les espèces ayant des doublets libres et les espèces chargées (anions), la présence d'orbitales diffusées est indispensable. On note par le signe (+) la présence d'orbitales diffusées, et par

un astérisque(*) les orbitales de polarisation.

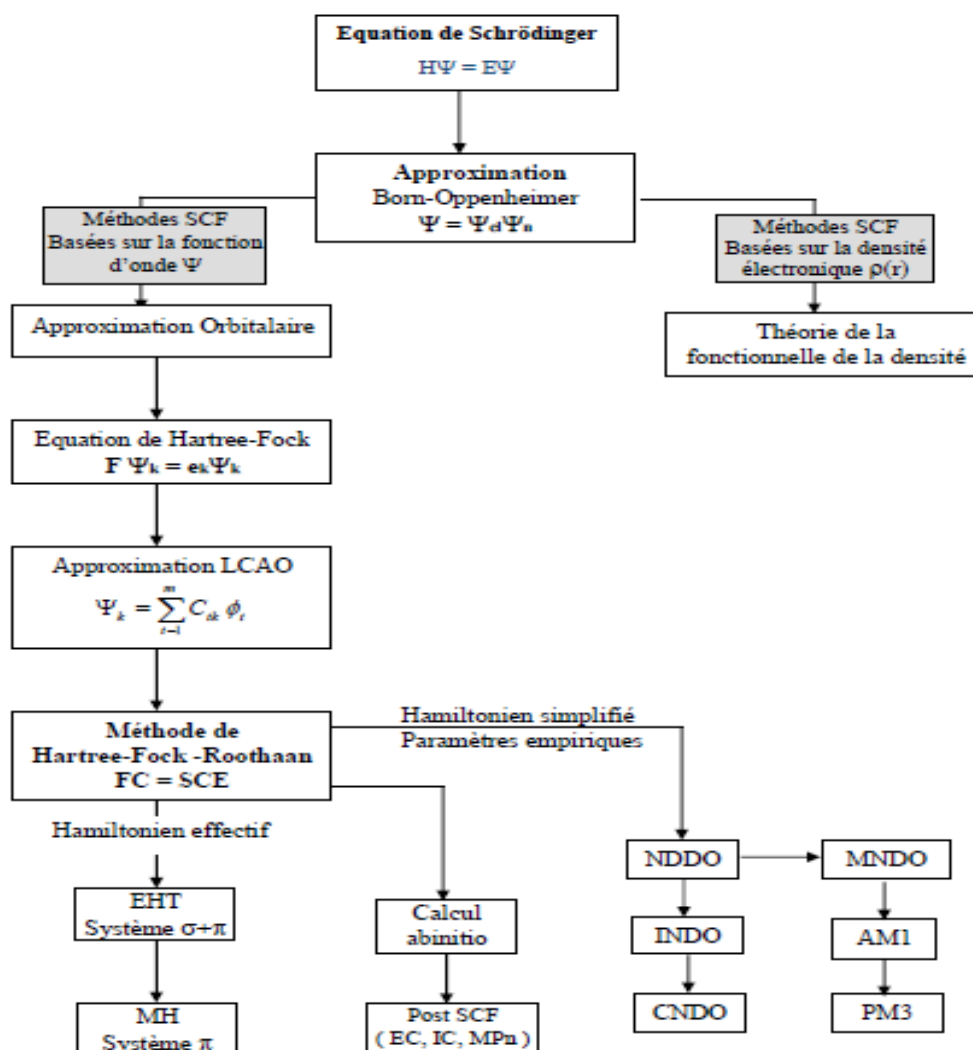


Figure. 1: Organigramme des méthodes utilisées en chimie quantique.

II.2.2. Mécanique moléculaire

L'expression "Mécanique moléculaire" désigne actuellement une méthode de calcul qui permet, d'obtenir des résultats de géométrie d'énergie moléculaires en se basant sur la mécanique classique.

La mécanique moléculaire est apparue en 1930 [27], mais s'est développée à partir des années soixante, quand les ordinateurs furent plus accessibles et plus performant.

Les méthodes de la mécanique moléculaire permettent le calcul de propriétés structurales et thermodynamiques de systèmes moléculaires comportant jusqu'à plusieurs milliers d'atomes.

Les électrons n'y sont pas traités explicitement comme dans un calcul de mécanique quantique, mais les atomes y sont représentés par des masses ponctuelles chargées reliées les unes aux autres par des ressorts auxquelles sont associés des paramètres dérivant de l'expérience ou de calculs précis quanto-chimiques. Le « champ de force » établi par cette méthode représente aussi bien que possible les variations de l'énergie potentielle avec la géométrie moléculaire. L'énergie de la molécule est exprimée sous la forme d'une somme de contributions associées aux écarts de la structure par rapport à des paramètres structuraux de référence :

$$E = E_{\text{Liaison}} + E_{\text{Angle}} + E_{\text{Dièdre}} + E_{\text{Van der Waals}} + E_{\text{Electrostatique}}$$

Les variables du calcul sont les coordonnées internes du système : longueurs de liaisons, angles de valence, angles dièdres, ainsi que les distances entre atomes non liés dont les interactions sont représentées par un potentiel de van der Waals et un potentiel électrostatique le plus souvent de type Coulombien.

Habituellement, on distingue dans l'équation de l'énergie du champ de force les termes intramoléculaires concernant les atomes liés chimiquement (liaisons, angle, dièdres, etc.) des termes intermoléculaires concernant les interactions entre les atomes non-liés chimiquement (électrostatiques, van der Waals, etc.).

II.2.2.1. Différents champs de force en mécanique moléculaire

Différents champs de force sont proposés dans la littérature, ils se distinguent les uns des autres par les termes dans le développement de l'expression de l'énergie de la molécule. Chacun a un domaine d'application spécifique de sorte que le choix d'un champ de force

dépend des propriétés et de l'application du système que l'on veut étudier. Les champs de forces en MM peuvent être groupés en trois classes principales [28].

- Champs de force contenant uniquement les termes harmoniques.
- Champs de force utilisant les termes d'ordre supérieur (cubique, quadratique,...).
- Champs de force suggérés par *Allinger et col.* [29] ne considérant pas que les termes de la mécanique moléculaire classique mais aussi les effets chimiques comme l'électronégativité.

MM2: Il a été développé par Allinger en 1976 et c'est le champ de force le plus utilisé par la communauté des chimistes organiciens [30]. Il a été conçu au début pour les molécules simples (alcanes, alcènes et alcynes non conjugués, les composés carbonylés, les sulfures, les amines...), mais ses version améliorées MM3 (1989) [31] et MM4 (1996) [32] lui permet de traiter des molécules organiques de plus en plus complexes.

MM+: Est une extension du champ de force MM2, avec l'ajout de quelques paramètres additionnels [33]. MM+ est un champ de force robuste, il a l'aptitude de prendre en considération les paramètres négligés dans d'autres champs de force et peut donc s'appliquer pour des molécules plus complexes tels que les composés inorganiques [34].

AMBER: (*Assisted Model Building with Energy Refinement*), ce champ de force a été paramétrisé pour les protéines et les acides nucléiques. Il a été utilisé pour les polymères et pour d'autres petites molécules [35].

OPLS: (*Optimised Potentials for Liquid Simulation*), comme l'indique son nom, est conçu pour optimiser le potentiel qui permet la description des propriétés de solvation [36, 37].

GROMOS: (*Groningen Molecular Simulation Program Package*), est conçu spécialement pour les biomolécules en milieu aqueux en vue de l'étude des interactions entre les molécules d'eau et les groupements polaires des protéines [38].

CHARMM (Bio+): (*Chemistry Harvard Macromolecular Mechanic*), Son concept est semblable à celui d'AMBER [39].

MMFF94: C'est l'un des champs de force de mécanique moléculaire les plus récents, il a été développé par Tom Halgren [40, 41]. Les termes d'énergie d'élongation, de valence et de torsion sont les mêmes que ceux de MM3 (expansion de MM2). MMFF94 comprend des termes croisés de liaison-valence [42].

II.2.2.2. Principes de la minimisation

La mécanique moléculaire nous permet de minimiser l'énergie calculée [43]. Le rôle de la minimisation est essentiellement de relaxer la structure initiale et d'éliminer les mauvais contacts interatomiques.

Le principe de base de la minimisation est qu'à partir d'une géométrie très approximative, on recherche le jeu de coordonnées cartésiennes qui réduit à son minimum la somme de toutes les contributions énergétiques dues aux déformations des coordonnées internes et aux interactions entre atomes non liés. En principe, il suffit de prendre la dérivée première de l'énergie stérique par rapport à chacun des degrés de liberté de la molécule et de trouver l'endroit sur l'hypersurface énergétique où pour chaque coordonnée r_i : $\frac{\partial E}{\partial r_i} = 0$

Les différentes méthodes de minimisation les plus utilisées sont :

- La méthode de la plus grande pente appelée «Steepest descent».
- La méthode du gradient conjuguée (Polak-Ribiere). (Cette méthode que nous avons utilisée)
- La méthode de Newton Raphson.

II.2.3. Dynamique moléculaire

Les premiers pas de la dynamique moléculaire n'ont pu se faire que grâce à l'arrivée des premiers ordinateurs (1957). Mais les premières réelles simulations ont été faites par Rahman, grâce à ses travaux sur la simulation de l'argon liquide en 1964 avec un temps de simulation de 10-11 s, puis de l'eau liquide [44] en 1971.

La mécanique moléculaire permet de calculer la position des atomes et des molécules dans leurs états énergétiques minima, à 0 °K c'est à dire sans la moindre vibration ni le moindre mouvement.

La dynamique moléculaire tente à simuler le mouvement des atomes et des molécules au cours du temps pour une température supérieure à 0 °K. Elle consiste à étudier la trajectoire d'une molécule en appliquant les lois de la mécanique classique Newtonienne et elle permet de simuler les mouvements intramoléculaires que l'on peut visualiser ensuite en temps réel. Ces mouvements correspondent à des vibrations autour d'un minimum, ou au passage d'un minimum à un autre minimum d'énergie [44].

II.3. QSAR

II.3.1. Principe de QSAR

Le principe des méthodes QSAR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative la structure moléculaire, codée par des propriétés moléculaires appelées descripteurs, avec une activité en utilisant des méthodes d'analyse de données [45]. Par ces relations nous pouvons développer des modèles prédictifs de la forme générale suivante:

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

L'objectif de ces méthodes est alors d'analyser les données structurales afin de déterminer les facteurs influençant l'activité mesurée [46].

En pratique, le développement d'un modèle commence par la collecte de données expérimentales fiables et en nombre important, ensuite il est nécessaire de générer un nombre de descripteurs, caractérisant les structures moléculaires, utilisés dans le développement du modèle QSAR [47]. Une fois le modèle construit, il doit être évalué par des méthodes de validation (validations interne et externe) afin d'estimer sa robustesse et son pouvoir prédictif [48-50]. En fin, pour tout modèle, il est important de connaître quel type de molécules ce modèle est utilisable (connaître le domaine d'applicabilité) [51].

Il y a plusieurs raisons pratiques qui justifient l'utilisation des études QSAR [9].

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

II.3.2. Les applications de l'étude QSAR

Certaines études QSAR semblent être un peu plus d'études universitaires, il existe un grand nombre d'applications de ces modèles tels que [52]:

- L'optimisation de l'activité pharmacologique.
- La conception rationnelle de nombreux autres produits tels que des agents tensioactifs, des parfums, des colorants et des produits chimique fins.
- L'identification des composés dangereux dans les premiers stades de développement de produits ou la projection des stocks de composés existants.
- La prédiction de la toxicité et les effets secondaires de nouveaux composés.
- La prédiction de la toxicité pour les espèces environnementales.
- La sélection des composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.
- La prédiction d'une variété de propriétés physico-chimique des molécules.
- La prédiction des effets conjugués de molécules, que ce soit dans des mélanges ou des formulations.

II.3.3. Outils et Méthodologie de QSAR

II.3.3.1. Paramètres biologiques

Les modèles QSAR sont dépendants des données expérimentales utilisées pour leur construction. Le modélisateur doit tenir compte des données à modéliser. Le choix de la

base de données est donc une étape très importante dans le développement des modèles QSAR.

Ces données devraient, idéalement, être de grande qualité, ce qui signifie qu'elles devraient être fiables et cohérentes. Il est donc important de les choisir parmi celles présentant des incertitudes faibles afin de limiter les barres d'erreurs expérimentales. De plus, le modélisateur doit s'assurer que les données expérimentales utilisées ont été obtenues selon le même protocole. En effet les conditions expérimentales ont, généralement, une forte influence sur les valeurs obtenues [53].

Il faut également que la distribution des données soit la plus homogène et normale que possible, car la plupart des méthodes statistiques sont basées sur ce type de distribution [54].

L'efficacité d'un modèle QSAR dépend également du type de molécules qui y sont incluses, plus le modèle présentera des composés de structures proches et similaires, plus il aura de chance d'être performant.

Les données biologiques sont généralement exprimées en logarithmes inverses ($\log 1/C$) afin d'obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces [55, 56].

II.3.3.2. Utilisation des descripteurs moléculaires

L'activité biologique ne peut pas être reliée directement à la structure moléculaire ; cette dernière est codée par des grandeurs qui représentent d'une manière quantitative les informations contenues dans la structure moléculaire telles que les caractéristiques physico-chimiques et structurelles. Ces grandeurs sont appelées descripteurs [57]. Une fois ces derniers disponibles, il est possible d'établir des relations entre ceux-ci et une activité à l'aide des outils de la modélisation [58]. On en dénombre aujourd'hui des milliers, qui peuvent être calculés ou obtenus de manière empirique. De nombreux logiciels ont été développés pour calculer les différents descripteurs moléculaires tels que : *Gaussian, ChemOffice, ChemSketch, Marvin Sketch, Dragon, QSARIS*.

Dans ce qui suit, nous allons présenter que les descripteurs moléculaires qui ont été utilisés dans l'ensemble de nos travaux.

Le poids moléculaire, noté MW (appelé aussi le poids de formule), mesuré en daltons (Da). C'est la somme des poids atomiques des différents atomes constituant la molécule. Il est utilisé dans l'étude de transport dont la diffusion et le mode de fonctionnement. Les composés avec des poids plus élevés sont moins susceptibles d'être absorbés et donc ne peuvent pas atteindre le site d'action. Ainsi, essayer de garder des poids moléculaires aussi bas que possible devrait être l'objectif pour établir un médicament. Pour les médicaments délivrés par voie orale le poids moléculaire doit être inférieur ou égal à 500 daltons (optimum autour de 300 daltons) [59].

Le volume moléculaire, noté MV, en cm³, est défini par la formule suivante :

$$MV = \frac{MW}{d}$$

Avec : MW est le poids moléculaire et d la densité.

Le coefficient de partage Octanol/Eau. Le transport, le passage à travers les membranes et l'activité pharmacologique d'une molécule peuvent être conditionnés par son partage entre une phase lipidique et une phase aqueuse, c'est-à-dire son caractère hydrophile. Celui-ci peut être quantifié par le coefficient de partage Octanol-Eau, noté (log P), qui mesure la solubilité différentielle d'un soluté dans ces deux solvants non miscibles. C'est une mesure importante pour l'identification de la similarité médicamenteuse, selon la règle de Lipinski, les médicaments délivrés par voie orale doivent avoir des valeurs de log P supérieures ou égales à -2 et inférieures ou égales à 5) [60]. Il est défini par la formule suivante:

$$\log P = \frac{[Octanol]}{[H_2O]}$$

[Octanol] et [H₂O] sont les concentrations du soluté dans l'Octanol et l'eau.

Les composés qui ont les valeurs de log P > 0 sont dites lipophiles, et les composés qui ont les valeurs de log P < 0 sont dites hydrophiles. Si le Log P est positif et très élevé, cela

exprime le fait que la molécule est plus soluble dans l'Octanol que dans l'eau, ce qui reflète son caractère lipophile, et inversement, si le Log P est négatif cela signifie que la molécule est hydrophile.

Un Log P nul signifie que la molécule est aussi soluble dans un solvant que dans l'autre.

La réfractivité moléculaire, notée (MR), en m³/mol, est le volume de la substance absorbée par mole de cette substance. Elle est définie par Lorentz-Lorenz par la formule suivante:

$$MR = \frac{n^2 - 1}{n^2 + 2} \frac{MW}{d} = \frac{n^2 - 1}{n^2 + 2} MV$$

Où : MW est le poids moléculaire ; d est la densité ; n est l'indice de réfraction ; MV est le volume molaire.

La réfractivité moléculaire est également proportionnelle à la polarisabilité α_e , par la relation suivante [41] : $MR = \frac{4}{3} \pi N_A \alpha_e$

Où : NA est le nombre d'Avogadro qui est, le nombre de molécules dans une mole de substance, NA = 6.022 10²³

La polarisabilité, notée (α_e), en (m³), est l'aptitude à la déformation du nuage électronique de la molécule sous l'influence d'un champ électrique uniforme. C'est l'un des paramètres qui traduisent les propriétés moléculaires liées à l'hydrophobie et par conséquent aux activités biologiques. Elle est calculée à partir de la réfractivité molaire ou du volume molaire comme suit :

$$\alpha_e = 0,3964308 \times MR = 0,3964308 \times \frac{n^2 - 1}{n^2 + 2} MV$$

L'énergie totale. Pour une molécule isolée à l'état fondamental, l'énergie totale calculée, notée Et, mesurée en eV, peut être utilisée comme descripteur moléculaire quantique. Cette énergie approximative a été calculée pour une conformation optimisée de la géométrie la plus stable dont la structure d'énergie est minimale.

Le moment dipolaire, noté μ , mesuré en debye (D), mesure la polarité nette moléculaire, et décrit la séparation de charge dans une molécule où la densité d'électrons est partagée inégalement entre les atomes. L'existence d'un moment dipolaire dans une molécule a son origine dans la différence d'électronégativité entre les atomes. La densité électronique est plus élevée au voisinage de l'atome le plus électronégatif. Ceci entraîne une dissymétrie dans la répartition des électrons de liaison. Ainsi, plus le moment dipolaire d'une molécule est élevé, plus la dissymétrie dans la molécule est importante.

Les énergies des orbitales frontières, jouent un rôle majeur dans de nombreuses réactions chimique et dans les mécanismes réactionnels. Les énergies de ces orbitales sont des paramètres très populaires dans la chimie quantique et dans les études RQSA:

L'énergie HOMO, notée E_{HOMO} , mesurée en eV, est le niveau d'énergie le plus élevé dans la molécule qui contient des électrons, il est directement lié au potentiel d'ionisation.

Lorsqu'une molécule agit comme une base de Lewis (un doublet d'électrons donneur) dans la formation d'une liaison, les électrons sont alimentés à partir de cette orbite. Il mesure la nucléophilie d'une molécule et caractérise la susceptibilité de la molécule à l'attaque par des électrophiles.

L'énergie LUMO, notée E_{LUMO} , mesurée en eV, est le niveau d'énergie le plus bas dans la molécule qui ne contient pas d'électrons, il est directement lié à l'affinité d'électron. Lorsqu'une molécule agit comme un acide de Lewis (un doublet d'électrons accepteur) dans la formation de liaisons, des doublets d'électrons entrants sont reçus dans cette orbite. Il mesure l'électrophilicité d'une molécule et caractérise la susceptibilité de la molécule à l'attaque par les nucléophiles.

Le Gap énergétique, ou l'écart HOMO-LUMO, noté E_{gap} , mesuré en eV, traduit l'énergie comprise entre l'orbitale moléculaire la plus haute occupée et la plus basse vacante, C'est

un indice de stabilité important. Cette différence d'énergie sert de mesure de l'excitabilité d'une molécule. Ainsi, plus l'intervalle d'énergie sera faible et plus la molécule pourra interagir avec l'environnement.

Un grand écart HOMO-LUMO implique une grande stabilité pour la molécule dans le sens de sa faible réactivité dans les réactions chimiques, et de même, un faible écart implique une grande réactivité de la molécule. L'écart HOMO-LUMO a également été utilisé comme une approximation de la plus faible énergie d'excitation de la molécule.

Descripteurs de charge, noté q_i , Ce sont des descripteurs électroniques définis en termes de charges atomiques utilisées pour décrire les aspects électroniques aussi bien de la molécule entière que de régions particulières, comme les atomes, les liaisons, les fragments moléculaires et les orbitales. Les charges renseignent sur l'importance de la localisation de la densité électronique dans une molécule: des valeurs q_i négatives signifient la localisation d'un excès de charge électronique au centre i , alors que des valeurs positives indiquent que le centre i est électroniquement déficitaire.

II.3.3.3. Sélection et réduction du nombre de descripteurs

Un grand nombre de descripteurs différents sont collectés pour la modélisation d'une grandeur donnée (activité), car les facteurs déterminants du processus étudié ne sont a priori pas connus. Cependant, les descripteurs envisagés n'ont pas tous une influence significative sur la grandeur modélisée, et les variables ne sont pas toujours indépendantes. De plus, le nombre de descripteurs, c'est-à-dire la dimension de la base de données d'entrée, détermine la dimension du vecteur des paramètres à ajuster. Si cette dimension est trop importante par rapport au nombre des observations (molécules) de la base d'apprentissage, le modèle risque d'être sur-ajusté à ces exemples, incapable de prédire la grandeur modélisée sur de nouvelles molécules et peut contenir des informations redondantes. Les descripteurs moléculaires employés doivent être porteurs de sens et interprétables d'un point de vue chimique. Et par conséquent, lorsque les descripteurs sélectionnés sont pertinents, ils offrent des idées sur les mécanismes, et les

modèles RQSA seront simples, transparents et compréhensibles [61].

Avant d'entamer le développement effectif des équations de régression RQSA, il est hautement recommandé d'examiner la qualité statistique des données de départ, à la fois les données à corrélérer (variable dépendante) et les descripteurs utilisés dans la corrélation (variables indépendantes).

Pour un ensemble de descripteurs différents, il est nécessaire d'effectuer une analyse des données bivariée, c'est-à-dire de calculer le coefficient de corrélation entre chacune des paires de l'ensemble des descripteurs. Si ce coefficient est statistiquement significatif ($R > 0,95$), ces deux descripteurs sont considérés comme fortement corrélés et ne peuvent être utilisés simultanément lors de l'analyse RQSA [62] et en pratique, ils seront alors enlevés dans le procédé de sélection. Ce type d'analyse est appelé l'analyse objective qui permet de réduire le nombre de descripteurs sans faire participer la variable dépendante (l'activité).

Finalement, pour que les relations RQSA ne soient pas statistiquement non significatives ou en cas d'erreur ponctuelles, il faut que le rapport composés/descripteurs doive être supérieur à 5 [63, 64].

II.4. Méthodes statistiques

II.4.1. Définition

Par définition, la statistique est « la science dont l'objet est de recueillir, de traiter et d'analyser des données issues de l'observation de phénomènes dans lesquels le hasard intervient (phénomène aléatoire) ». Par conséquent, l'objectif principal de la statistique est d'*Approches de* maîtriser au mieux l'incertitude pour extraire des informations utiles des données, par l'intermédiaire de l'analyse des variations dans les observations. En outre, l'analyse des données est utilisée pour décrire, comprendre et gérer les phénomènes étudiés, faire des prévisions et prendre des décisions [59].

II.4.2. Méthodes statistiques

Faire de la statistique suppose que l'on étudie un ensemble d'objets équivalents sur lesquels on observe des caractéristiques appelées « variables ». Dans notre cas, les objets

(ou individus) sont les molécules et les variables sont les descripteurs moléculaires précédemment décrits dans ce chapitre.

Après le recueil des descripteurs, la démarche statistique consiste à traiter et interpréter les informations recueillies sur ces molécules. Cette démarche comporte deux grandes classes : la statistique descriptive et la statistique décisionnelle ou prédictive.

II.4.2.1. La statistique descriptive

La statistique descriptive (appelée aussi l'analyse des données), a pour but d'extraire le maximum de l'information contenue dans les données d'une façon efficace, simple et compréhensible. Elle permet de résumer les caractéristiques essentielles du phénomène étudié et de suggérer des hypothèses pour des études plus sophistiquées. Elle utilise pour cela des représentations de données sous forme de graphiques, de tableaux et d'indicateurs statistiques. Elle est utilisée aussi pour diviser et classer les données dans des classes homogènes.

- L'analyse en composantes principales (ACP) comme technique pour l'analyse des données
- La méthode du partitionnement en k-moyennes (ou k-means en anglais)
- La classification ascendante hiérarchique (CAH) pour la classification des données.

II.4.2.2. La statistique décisionnelle ou prédictive

Contrairement à la statistique descriptive, dans ce type de statistiques les probabilités jouent un rôle fondamental. Cette statistique a pour but de prendre des décisions et de faire des prévisions au vu des observations. En général, il faut pour cela proposer des modèles probabilistes du phénomène aléatoire étudié et savoir gérer les risques d'erreurs. Dans notre cas, il faut rechercher une relation approximative entre une activité ou propriété et plusieurs variables quantitatives (descripteurs moléculaires), la forme de cette relation peut être linéaire ou non linéaire.

Dans l'ensemble de nos travaux, nous avons utilisé la régression linéaire multiple MLR pour la construction des modèles RQSA.

II.4.3. La régression linéaire multiple

La régression linéaire multiple MLR est l'une des méthodes de modélisation les plus populaires grâce à sa simplicité d'utilisation et facilité d'interprétation. L'avantage important de la régression linéaire multiple est qu'elle est très transparente, puisque l'algorithme est disponible, et que les prédictions peuvent être réalisées facilement. Cette méthode a été utilisée aussi pour la sélection des descripteurs moléculaires utilisés dans les autres méthodes statistiques [65]. La méthode MLR se base sur l'hypothèse que la propriété y dépend linéairement des différentes variables (les descripteurs) $x_1, x_2, x_3, \dots, x_i$, selon la relation:

$$Y = a_0 + \sum_{i=0}^n a_i X_i$$

Avec : y est la variable dépendante (à expliquer ou à prédire).

X_i : sont les variables indépendantes (explicatives).

n : est le nombre de variables explicatives.

a_0 : est la constante de l'équation du modèle.

a_i : sont les coefficients de descripteurs dans l'équation du modèle.

La taille de ces coefficients indique le degré d'influence des descripteurs moléculaires correspondants sur l'activité cible. Un coefficient positif indique que le descripteur moléculaire correspondant contribue positivement à la cible, tandis qu'un coefficient négatif suggère la contribution négative.

La méthode de progression que nous avons utilisée c'est la progression ascendante.

La MLR progressive ascendante, qui consiste à incorporer les variables au modèle une à une, en sélectionnant, à chaque étape, la variable dont la corrélation partielle avec la grandeur modélisée est la plus élevée.

II.4.4. Techniques de validation

II.4.4.1. Coefficients et tests statistiques standards

Afin de déterminer la qualité d'un modèle, différents paramètres statistiques sont employés, tels que les erreurs quadratiques moyennes (Mean Square Errors), les coefficients de corrélation qui sont régulièrement utilisés dans les études RQSA/RQSP, sont décrits en détail dans cette partie.

a. Coefficient de corrélation r (et coefficient de détermination r^2)

C'est l'indicateur statistique le plus répandu est le coefficient de corrélation qui évalue la part de la variance de l'activité / la propriété cible expliquée par le modèle.

$$r = \sqrt{1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}}$$

Avec : r : est le coefficient de corrélation.

y_i, \hat{y}_i : sont, respectivement, les valeurs observées et calculées de la variable dépendante.

\bar{y}_i : est la valeur moyenne des valeurs observées.

Ces coefficients ne sont pas affectés par l'unité de mesure choisie et traduisent une bonne corrélation entre l'activité cible et l'activité initiale si r^2 est proche de 1 (cas idéal).

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

La valeur de r^2 dépend de la taille de l'échantillon et le nombre de variables prédictives dans l'équation. Il garde la même valeur ou augmente lors d'une nouvelle variable de prédiction est ajoutée à l'équation de régression, même si la variable ajoutée ne contribue pas à la réduction de la variance inexpliquée. Par conséquent, un autre paramètre statistique peut être utilisé, appelé r^2 ajusté (r^2_{adj}). Bien entendu, un autre indicateur est l'erreur quadratique moyenne (MSE, pour Mean Square Error), à laquelle est parfois préférée la déviation standard s .

b. Le coefficient de détermination ajusté r^2_{adj}

Ce coefficient est tient compte du degré de liberté :

$$r^2_{adj} = \frac{r^2(n-1) - p}{n - p - 1}$$

Avec : n : est le nombre des observations (les molécules).

p : est le nombre de variables indépendantes (les descripteurs).

r^2 : est le coefficient de détermination du modèle.

c. L'erreur type résiduel « s » (La déviation standard)

$$s = \sqrt{\frac{\sum(\hat{Y}_i - Y_i)^2}{n - p - 1}}$$

Avec:

y_i , \hat{y}_i : sont, respectivement, les valeurs observées et calculées de la variable dépendante.

n : est le nombre des observations (les molécules).

p : est le nombre de variables indépendantes (les descripteurs).

Ces paramètres mesurent la variation de l'activité cible non expliquée par le modèle RQSA. En particulier, plus la déviation standard est petite et plus la corrélation est meilleure. Sa valeur est toujours fonction de l'unité de mesure de l'activité cible et tient également compte des erreurs expérimentales ce qui explique qu'une valeur trop petite n'ait aucune signification.

d. Le test de Fisher F

L'indice de Fisher F-test est employé afin de mesurer le niveau de signification statistique du modèle à « x% » (le niveau usuel est 95%), c'est-à-dire la qualité du choix du jeu de paramètres [59].

Hypothèses :

H₀: les variances des échantillons sont homogènes.

H₁: les variances des échantillons ne sont pas homogènes.

La valeur à calculer est:

On calcule le F (observé) à partir de la formule :

$$F(\text{observée}) = \frac{\sum(\hat{Y}_i - \bar{Y}_i)^2 n - p - 1}{\sum(Y_i - \hat{Y}_i)^2 p}$$

Avec: F est l'indice de Fisher.

y_i , \hat{y}_i : sont, respectivement, les valeurs observées et calculées de la variable dépendante.

n: est le nombre des observations (les molécules).

p: est le nombre de variables indépendantes (les descripteurs).

\bar{y}_i : est la valeur moyenne des valeurs prédites.

Après le calcul de F (observé) on le compare avec le F théorique obtenu à partir des tables statistiques usuelles (la table de Fisher).

- Si F observé est plus grand que le F théorique : refus de l'hypothèse nulle H_0 et cela signifie que les variances des échantillons sont trop différentes pour être considérées comme homogènes.
- Si F observé est plus petit que le F théorique : acceptation de l'hypothèse nulle H_1 et cela signifie que les deux variances ont des valeurs suffisamment proches pour qu'on accepte l'idée qu'elles soient homogènes.

II.4.4.2. Pouvoir de prévision

a. Interne

Quel que soit le but ultime du modèle QSAR, il doit être obligatoirement validé avant d'être interprété ou utilisé à fins prédictives. Il existe différentes solutions pour s'assurer de validité d'un modèle. La validation croisée LOO (leave-one-out) est un processus qui permet de tester la précision prédictive d'un modèle, cette méthode basée sur le calcul de quelques paramètres statistiques tels que : $PRESS$, SSY , $Spress$, R^2_{CV} et R^2_{adj} , ces paramètres sont calculés à partir des équations suivantes [66] :

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

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

$$R^2_{adj} = (1 - r^2) \left(\frac{n - 1}{n - p - 1} \right)$$

$$R^2_{CV} = 1 - \frac{PRESS}{TSS}$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n}}$$

Avec: n : nombre de molécules de la série étudiée.

$PRESS$: Somme des carrés des écarts entre les valeurs observées et les valeurs prévues à partir du modèle.

SSY: Somme des carrés des distances des valeurs observées pour une variable par rapport à la moyenne de cette variable, la somme des carrés permet de mesurer

la variation totale dans une variable.

S_{PRESS}: La capacité prédictive des modèles est évaluée par l'erreur de racine-carrée-moyenne.

R²_{adj}: Estimation de la qualité d'ajustement du modèle à la population.

R²_{cv}: Est la variation de la statistique du R² obtenue en ajoutant ou en enlevant une variable indépendante.

b. Externe

Cette méthode consiste à prédire l'activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modèle, cette validation est caractérisée par le paramètre R²_{test}. Récemment plusieurs études [66] ont montré l'insuffisance des paramètres R², R²_{cv} pour vérifier le pouvoir prédictif des modèles RQSA. Par conséquent, d'autres paramètres doivent être vérifiés pour cet objectif. Ces paramètres sont connus sous le nom «critères de validation externe» ou souvent appelés «critères de Trophsa »

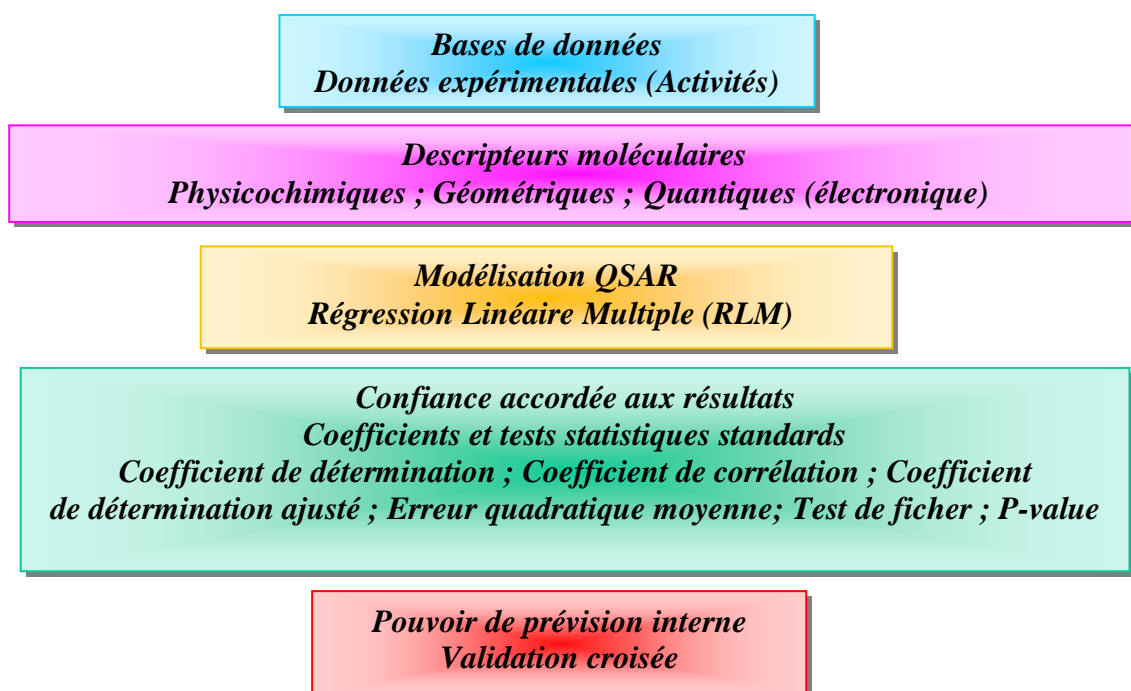


Figure .2 : Schéma de la méthodologie utilisée dans nos travaux.

Références bibliographiques

- [1] E. Derety, J. Mol. Structr. (Theochem), 459, 273, 1999.
- [2] S. Belaidi, M. Omari, T. Lanez, A. Dibi, Contribution a l'étude de la relation structure-activité dans des nouveaux macrolides antibiotiques, J. Soc. Alger. Chim, 14, 27, 2004.
- [3] A. C. Brown, T.R. Fraser, On the connection between chemical constitution and physiological action. Trans. Roy.Soc. 25, 151-203, 1868.
- [4] Hansch, Quantitative approach to biochemical structure-activity relationships, Acc. Chem. Res, 2, 232-239, 1969.
- [5] Hansch, T. Fujita, Activity and Chemical Structure, J. Am. Chem. Soc, 86, 1616-1626, 1964.
- [6] A. K. Debnath, Quantitative structure-activity relationship (QSAR) paradigm--Hansch era to new millennium, Mini Rev.Medic.Chem, I, 187-195, 2001.
- [7] M. Grover, B. Singh, M. Bakshi, S. Singh, Quantitative structure-property relationships in pharmaceutical research-Part 2, Pharm. Sci. Tech. Today, 3, 50-57, 2000.
- [8] M. Grover, B. Singh, M. Bakshi, S. Singh, Quantitative structure-property relationships in pharmaceutical research-Part 1, Pharm. Sci. Tech. Today, 3, 28-35, 2000.
- [9] T. I. Oprea, Chemoinformatics in Drug Discovery, Ed. WILEY-VCH Verlag. Allemagne, 2005.
- [10] E. A. Rekká ; P. N. Kourounakis, Chemistry and Molecular Aspects of Drug Design and Action, Ed. Taylor & Francis Group, LLC. Etats Unies, 2008.
- [11] A. K. Ghose, V. N. Viswanadhan, Combinatorial Library Design and Evaluation: Principles, Software Tools, and Applications in Drug Discovery, Ed. Marcel Dekker, Inc. Etats Unies, 2001.
- [12] J. L. Rivail, Elements de chimie quantique à l'usage des chimistes, CNRS Editions, Paris, 1999.
- [13] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, The development and use of quantum mechanical molecular models, J. Am. Chem. Soc, 107, 3902-3909, 1985.
- [14] J. J. P. Stewart, Optimization of parameters for semiempirical methods I. Method, J. Comput. Chem, 10, 209-220, 1989.
- [15] G. Boucekkine, Méthodes de la chimie quantique, Techn. Ing, 2007.
- [16] N. Merabet, Thèse de Doctorat: Université 8 mai 1945 Guelma, 2017.
- [17] E.Fermi, Springer Handbook of Atomic, Molecular and Optical Physics, Rend. Accad, 6, 602, 1927.
- [18] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys. Rev, 136, 864, 1964.
- [19] M. Rahim, Thèse de Doctorat : Université 8 mai 1945 Guelma, 2015.

- [20] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, **38**, 3098, 1988.
- [21] J.P. Perdew, Y. Wang, Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation, *Phys. Rev. B*, **33**, 8800, 1986.
- [22] J.P. Perdew, Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, *Phys. Rev. B*, **45**, 13244, 1992.
- [23] Y. Zhang, W. Yang, Comment on Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett*, **80**, 890, 1998.
- [24] J.P. Perdew, S. Kurth, A. Zupan, P. Blaha, Accurate Density Functional with Correct Formal Properties: A Step Beyond the Generalized Gradient Approximation, *Phys. Rev. Lett*, **82**, 2544, 1999.
- [25] R.S. Mulliken, Electronic population analysis on LCAO-MO molecular wave functions, *J. Chem. Phys*, **23**, 1833-1846, 1955.
- [26] W. J. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton, J. A. Pople, Gaussian 70 Quantum Chemistry Program Exchange, Program No. 237, 1970.
- [27] L.P. Hammett, The effect of structure upon the reactions of organic compounds. Benzene derivatives, *J. Am. Chem. Soc*, **59**, 96-103, 1937.
- [28] T. Shimanouchi, Force constants of small molecules, *Pure Appl. Chem*, **7**, 131-146, 1963.
- [29] N.L. Allinger, *Q. C. P. E*, **3**, 32, 1983.
- [30] N.L. Allinger, Y.H. Yuh, J-H. Lii, Molecular mechanics. The MM3 force field for hydrocarbons, *J. Am. Chem. Soc*, **111**, 8551-8566, 1989.
- [31] N.L. Allinger, K. Chen, J-H. Lii, An improved force field (MM4) for saturated hydrocarbons, *J. Comp. Chem*, **17**, 642-668, 1996.
- [32] G.B. Jones, B.J. Chapman, π stacking effects in asymmetric synthesis, *J. Syn. Org. Chem*, **5**, 475-497, 1995.
- [33] A. Hocquet, M. Langgard, An evaluation of the MM+ force field, *J. Mol. Model*, **4**, 94-112, 1998.
- [34] S.J. Weiner, P.A. Kollman, T. Nguyent, D.A. Case, An all atom force field for simulations of proteins and nucleic acids, *J. Comput. Chem*, **7**, 230-252, 1986.
- [35] W.L. Jorgensen, J.T. Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc*, **110**, 1657-1666, 1988.
- [36] J.C. Smith & M. Karplus, Empirical force field study of geometries and conformational transitions of some organic molecules, *J. Am. Chem. Soc*, **114**, 801-812, 1992.
- [37] W.F. Van Gunsteren, M. Karplus, Effect of constraints on the dynamics of macromolecules, *Macromolecules*, **15**, 1528-1544, 1982.

- [38] B.R. Brooks, R.E. Bruccoleri, B.D. O'falon, D.J. States, S. Swaminathan, M. Karplus, Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comp. Chem*, 4, 187-217, 2004.
- [39] F. Tristram, V. Durier, G. Vergoton, The structures and vibrational frequencies of a series of aliphatic alcohols as obtained using the spectroscopic potential SPASIBA, *J. Mol. Struct*, 378, 249-256, 1996.
- [40] T.A. Halgren, Merck molecular force field. IV. conformational energies and geometries for MMFF94, *J. Comp. Chem.*, 17, 490, 1996.
- [41] T.A. Halgren, R.B. Nachbar, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J.Comp. Chem*, 17, 525-552, 1996.
- [42] M. Baaden, Outils pour étudier la structure et la dynamique des peptides et des protéines, Institut de Biologie Physico-chimique, PARIS, CNRS UPR9080, 2003.
- [43] S.J .Weiner, P.A .Kollman, T .Nguyent, D.A.Cas. An all atom force field for simulations of proteins and nucleic acids, *J. Comput .Chem*, 7, 230-252, 1986.
- [44] F. Li, H. Wu, L. Li, X. Li, J. Zhao, W. J. G. M. Peijnenburg, Docking and QSAR study on the binding interactions between polycyclic aromatic hydrocarbons and estrogen receptor, *Ecotoxico. Envi. Saf*, 80, 273-279, 2012.
- [45] A. Srinivas Reddy, S. Kumar, R. Garg, Hybrid-Genetic Algorithm based Descriptor Optimization and QSAR Models for Predicting the Biological Activity of Tipranavir Analogs for HIV Protease Inhibition, *J Mol Graph Model*, 28, 852-862, 2010.
- [46] M. Zhao, Z. Li, L. Peng, Y. R. Tang, C. Wang, Z. Zhang, S. Peng, A new class of analgesic agents toward prostacyclin receptor inhibition: synthesis, biological studies and QSAR analysis of 1-hydroxyl-2-substituted phenyl-4,4,5,5-tetramethylimidazolines, *Bioorg. Med.Chem*,15, 2815-2826 , 2007.
- [47] A. A. Toropov, A. P. Toropova, E. Benfenati, G. Gini, D. Leszczynska, J. Leszczynski, G. De Nucci, QSAR models for inhibitors of physiological impact of *Escherichia coli* that leads to diarrhea, *Biochem. Biophys. Res. Comm*, 432, 214-225, 2013.
- [48] A. Golbraikh, A. Tropsha, Beware of q^2 !, *J. Mol. Graph. Model*, 20, 269-276, 2002.
- [49] A. Tropsha, P. Gramatica, K.V. Gombar, The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models, *QSAR. Comb. Sci*, 22, 69-77, 2003.
- [50] P. Gramatica, E. Giani, E. Papa, Statistical external validation and consensus modeling: a QSPR case study for Koc prediction, *J. Mol. Graph. Mod*, 25, 755-766, 2007.
- [51] T. Ziani, Prédiction de l'effet inhibiteur des dérivés d'imidazole sur l'enzyme Cox-2, Université Med Khider Biskra, 2013.

- [52] D. Fourches, E. Muratov, A. Tropsha, Trust, but verify: on the importance of chemical structure curation in cheminformatics and QSAR modeling research, *J. Chem. Inf. Model*, 50, 1189–1204, 2010.
- [53] A. Tropsha, Best practices for QSAR model development, validation, and exploitation, *Mol Inform*, 29, 476–488, 2010.
- [54] S. Chatterjee, A. S. Hadi, B. Price, *Regression analysis by example*, Wiley-Interscience, New York, 2000.
- [55] H. T. N. Phuong, Synthèse et étude des relations structure/activité quantitatives (QSAR/2D) d'analogues Benzo[c]phénanthridiniques, Thèse de doctorat, Université d'Angers, France, 2007.
- [56] N. L. Mai, Y-M Koo, Quantitative prediction of lipase reaction in ionic liquids by QSAR using COSMO-RS molecular descriptors, *Biochem. Engin. J.*, 87, 33-40, 2014.
- [57] J. Wu, Y. Wang, Y. Shen, Molecular docking and QSAR analysis on maleimide derivatives selective inhibition against human monoglyceride lipase based on various modeling methods and conformations, *Chemomet, Intel. Lab. Syst*, 131, 22-30, 2014.
- [58] S. Chtita, Modélisation de molécules organiques hétérocycliques biologiquement actives par des méthodes QSAR/QSPR. Recherche de nouveaux médicaments. Thèse de doctorat. Université Moulay Ismaïl, Meknès, 2017.
- [59] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adva.Dru.Deliv.Rev*, 6, 3–25, 1997.
- [60] S. Mannan, *Lee's Loss Prevention in Process Industries: Hazard Identification, Assessment and Controll*, Elsevier Butterworth-Heinemann, Burlington, 2005.
- [61] N. Trinajstić, S. Nikolić, S.C. Basak, I. Lukovits, Distances indices and their hypercounterparts: Intercorrelation and use in the structure-property modeling, SAR and QSAR. *Envir. Rese*, 12, 31–54, 2001.
- [62] P.P. Roy, S. Paul, I. Mitra, K. Roy, Two novel parameters for validation of predictive QSAR models, *Molecules*, 14, 1660–701, 2009.
- [63] J.G. Topliss, R.P. Edwards, Chance factors in studies of quantitative structure-activity relationships, *J. Med. Chem*, 22, 1238–1244, 1979.
- [64] K. Roy, S. Kar and R. Narayan Das, *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, Chapter 6: Selected Statistical Methods in QSAR, Acad. Press, Boston, 191–229, 2015.
- [65] A. Tropsha, P. Gramatica, and V.K. Gombar, The importance of Being Earnest: Validation is the Absolute Essential for Successful Application and interpretation of QSPR Models, *QSAR and Combinatorial Sciences*, 22, 69–77, 2003.
- [66] R. Pratim, P. Partha, Somnath, Mitra, Indrani, et al, On two novel parameters for validation of predictive QSAR models, *Molecules*, 14, 1660-1701, 2009.

Chapitre III

Chapitre III

III.1. Introduction

Vu les dégâts impliqués par la propagation des radicaux libres dans l'organisme causant ainsi de nombreuses maladies plus ou moins graves, l'activité anti-oxydante se présente parmi les activités biologiques qui suscitent le plus d'intérêt ces dernières années [1].

L'objectif de ce travail est d'étudier et de prédire l'activité anti-oxydante d'une série constituée de 25 composés phénoliques simples, en utilisant des descripteurs cibles qui peuvent expliquer le mécanisme de cette activité étudiée et en respectant toute la méthodologie d'une étude QSAR. De cette manière, nous espérons comprendre quelles caractéristiques moléculaires sont corrélées à l'activité antioxydante des composés phénoliques contre les radicaux libres.

III.2. Présentation de la série de molécule, Méthodes et Matériel

III.2.1. Série de molécule

Le potentiel redox déterminé expérimentalement à pH 7 (E_7) a été couramment utilisé comme mesure directe de l'activité anti-oxydante par plusieurs auteurs [1, 2-5]. Les 25 composés utilisés dans cette étude, ainsi que les valeurs du potentiel redox expérimental (E_7) sont présentés sur la figure 1.

III.2.2. Méthodes et matériels

Pour réaliser cette étude nous disposons des logiciels suivants :

- Tout d'abord, les vingt cinq molécules du phénol ont été pré-optimisées au moyen de la mécanique moléculaire, champ de force (MM+) en utilisant le logiciel HyperChem 8.03 [6]. Ensuite, ces dérivées ont été réoptimisées en utilisant le logiciel gaussien 09 [7], au niveau de la théorie de la densité fonctionnelle DFT en utilisant le Lee-Yang-Parr à trois paramètres de Becke (B3LYP), avec de base 6 - 311++G (d, p).
- Le module « propriétés QSAR » de l'HyperChem 8.03 a été utilisé pour calculer les paramètres suivants: La polarisabilité (Pol), la réfractivité molaire (MR), le

coefficient de partage octanol/eau ($\log P$), le volume molaire (MV), la surface Moléculaire (SAG) et le poids moléculaire (MW).

- Les descripteurs électroniques: le moment dipolaire (MD), l'énergie des orbitales frontières (E_{HOMO} , E_{LUMO}), les charges (q_{C1} et q_{O}), l'énergie totale (E_{T}) ont été calculés par la méthode DFT (B3LYP /6-311++G (d, p) et les charges partielles atomiques peuvent être calculées par le modèle Chelpg [8]) en utilisant logiciel Gaussian 09 et GaussView (5.0).
- Logiciel MarvinSketch [9] a été utilisé pour calculer les descripteurs suivantes: Le nombre des Liaisons hydrogène donneur et accepteur (HBD, HBA), Superficie de la surface polaire (PSA), le coefficient de partage octanol/eau ($\log P$), Lipophilie (LogD à $\text{pH}=7,4$) et aussi le nombre de liaison rotatif (nrotb). Ces descripteurs sont utilisés pour l'identification des composés "drug-like".
- La méthode de régression multilinéaire (MLR) a été utilisée pour développer de modèle QSAR en utilisant le logiciel SPSS 20 [10].

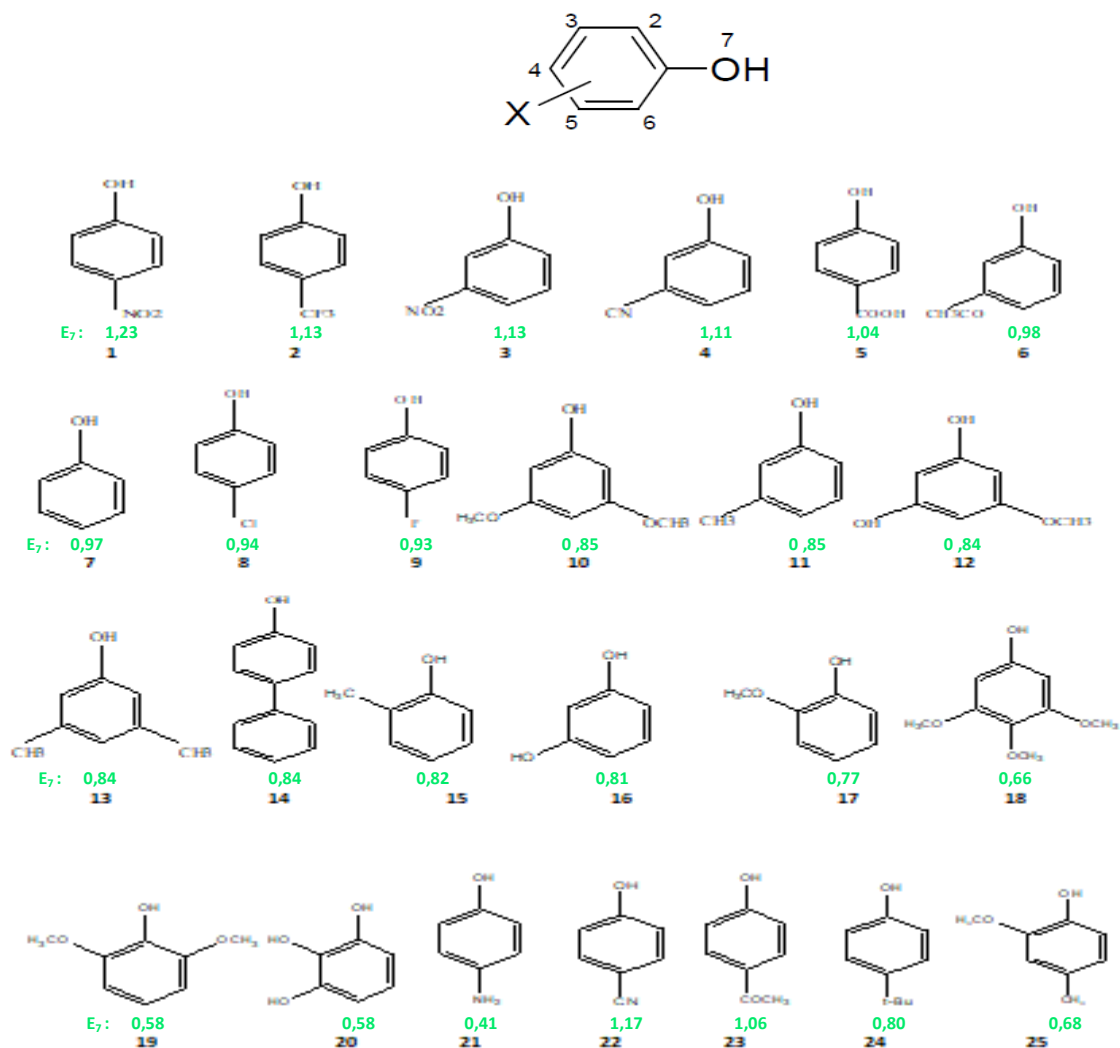


Figure.1. Structure chimiques des composés phénoliques avec les valeurs du potentiel redox expérimental (E_7)

III.3. Etude qualitative des relations structure/activité (SAR) dans les dérivés du phénol

Dans cette partie nous avons étudié sept propriétés physico-chimiques d'une série de vingt cinq de dérivés de phénol à l'aide du logiciel HyperChem 8.03 [6]. Le tableau 1 regroupe les propriétés suivantes: Surface (S), volume molaire (V), énergie d'hydratation (HE), coefficient de partage octanol / eau (LogP), réfractivité molaire (Réf), polarisabilité (Pol) et poids moléculaire (MW).

Tableau.1: Propriétés physicochimiques des dérivés du phénol.

Composés	Log P	Pol (Å ³)	MW (amu)	V (Å ³)	Réf (Å ³)	S (Å ²)	HE (Kcal/mol)
1	0,26	12.91	139.11	413.60	37,29	268.11	-13,81
2	2,57	12.63	162.11	437.13	33,47	280.02	-8,16
3	0,26	12.91	139.11	414.59	37,29	266.42	-13,51
4	0,79	14.26	123.15	435.13	33,18	244.40	-12,80
5	0,99	13.63	138.12	415.66	35,42	252.40	-13,75
6	1,12	14.83	136.15	445.77	38,66	274.93	-7,41
7	1,45	11.07	94.11	345.12	28,46	213.91	-8,96
8	2,07	13.00	128.56	388.85	33,48	247.73	-8,54
9	1,90	10.98	112.10	353.42	28,42	223.29	-8,62
10	0,87	16.02	154.17	491.95	41,45	313.76	-11,36
11	1,81	12.91	108.14	398.52	33,43	252.83	-7,45
12	0,53	14.18	140.14	444.09	36,98	276.95	-16,67
13	2,14	14.74	122.74	457.90	38,40	298.61	-6,15
14	3,07	20.73	170.21	544.27	53,90	285.64	-9,66
15	1,81	12.91	108.14	392.56	33,43	241.15	-6,49
16	0,79	11.71	110.11	365.45	30,49	222.91	-15,50
17	1,15	13.54	124.14	422.53	34,96	256.52	-8,59
18	0,60	18.49	184.19	569.22	47,94	335.73	-11,24
19	0,87	16.02	154.17	497.77	41,45	300.39	-9,20
20	0,18	12.35	126.11	378.97	32,51	212.09	-19,60
21	0,79	12.42	109.13	378.53	32,87	215.48	-14,31
22	0,79	12.92	119.12	398.32	33,18	264.48	-12,96
23	1,12	14.83	136.15	446.43	38,66	277.52	-7,75
24	2,76	18.41	150.22	535.46	47,73	341.83	-6,38
25	1,48	15.38	138.17	475.70	39,92	295.25	-7,35

Grande Valeur

Petite Valeur

Interprétation des résultats :

A lumière de ces résultats, on remarque que les valeurs de la polarisabilité sont généralement proportionnelles aux valeurs des surfaces et des volumes.

Presque toutes les valeurs de la polarisabilité et de la réfractivité molaire augmentent relativement avec la taille et le poids moléculaire des phénols étudiés. Ce résultat est en accord avec la formule de Lorentz-Lorenz qui donne une relation entre la polarisabilité, la réfractivité molaire et la taille moléculaire [11, 12].

Cette relation montre que la réfractivité molaire et la polarisabilité augmentent avec le volume et la masse moléculaire. Par exemple, les composés 14 et 18 portent des substituants volumineux ont des valeurs importantes de la polarisabilité ($20,73 \text{ \AA}^3$), ($18,49 \text{ \AA}^3$) et de la réfractivité molaire ($53,90 \text{ \AA}^3$), ($47,94 \text{ \AA}^3$) respectivement. Contrairement, les composés 7 et 9 sont des petites molécules dans la série étudiée ci-dessus; elles ont des petites valeurs de polarisabilité ($11,07 \text{ \AA}^3$) et ($10,98 \text{ \AA}^3$) et de réfractivité molaire ($28,46 \text{ \AA}^3$) et ($28,42 \text{ \AA}^3$).

L'énergie d'hydratation en valeurs absolues, les petites valeurs sont celles des composés 13 ($-6,15 \text{ Kcal/mol}$), 24 ($-6,38 \text{ Kcal/mol}$) et la plus importante valeur est celle du composé 20 ($-19,60 \text{ Kcal/mol}$) (Tableau. 1).

Les groupements hydrophobes dans les structures des dérivés de phénol induisent une diminution de l'énergie d'hydratation.

Le composé 20 possède trois sites donneurs de proton (OH) et trois sites accepteurs de proton (O) (Figure 2).

La lipophilie est une propriété qui a un effet majeur sur la solubilité, l'absorption, la distribution, le métabolisme et les propriétés d'excrétion, ainsi que l'activité pharmacologique. Hansch et Leo ont expliqué que des molécules hautement lipophiles seront distribuées dans l'intérieur lipidique des membranes et y seront retenues [13]. Pour une bonne biodisponibilité orale, et optimale le log P doit être ($0 < \log P < 3$). Pour log P trop élevé, le médicament a une faible solubilité et pour un log P trop faible; le médicament a des difficultés à pénétrer les membranes lipidiques. [14].

D'après les résultats (Tableau .1), La majorité des composés étudiés présentent des valeurs optimales de Log P, qui varient entre $0,26 \leq \text{Log P} \leq 3,07$, il n'existe pas des valeurs négatives. Ces molécules ont une bonne absorption intestinale due à une bonne balance entre la solubilité et la perméabilité par diffusion passive. En fait, le métabolisme est minimisé à cause de la faible fixation avec les enzymes métaboliques.

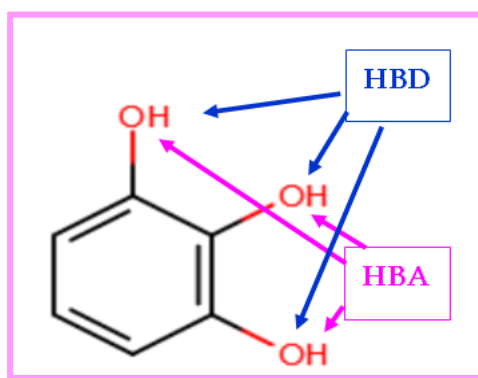


Figure.2: Sites donneur et accepteur du composé 20

III.4. Les méthodes d'optimisation multiparamètres (MPO) :

Pour prédire un médicament de haute qualité qu'il doit concilier de nombreuses propriétés, notamment la puissance, l'ADME et sécurité. Les méthodes d'optimisation multiparamètres (MPO) permettent de mieux utiliser ces données pour cibler rapidement des composés présentant un bon équilibre de propriétés, mais ils ont tous leurs forces et leurs faiblesses [15]. Nous pouvons donc utiliser les méthodes d'optimisation multiparamètres (MPO) pour prédire le meilleur équilibre des propriétés. Parmi ces méthodes : les règles de Lipinski , Veber [16,17] ainsi que Triangle d'or (Golden triangle). [18]

III.4.1. Règle de lipinski

Le médicament apparaît comme un paradigme prometteur pour coder l'équilibre entre les propriétés moléculaires d'un composé qui influence sa pharmacodynamie et sa pharmacocinétique et optimise finalement leur absorption, leur distribution, leur métabolisme et leur excrétion (ADME) dans le corps humain comme un médicament. Les conditions empiriques pour satisfaire la règle de Lipinski et manifester une bonne

biodisponibilité orale impliquent un équilibre entre la solubilité aqueuse d'un composé et sa capacité à diffuser passivement à travers les différentes barrières biologiques [16, 19]. Ces règles sont la plus utilisée pour l'identification des composés « drug-like ». D'après ces règles, les composés ne validant pas au moins deux des critères suivants ont de très fortes chances d'avoir des problèmes d'absorption ou de perméabilité par voie orale [20]. Ces règles, communément appelées « règles de Lipinski » ou « la règle de 5 », comportent cinq critères physico-chimiques qui décrivent la molécule, sont résumé sur le schéma suivant (Figure 3):

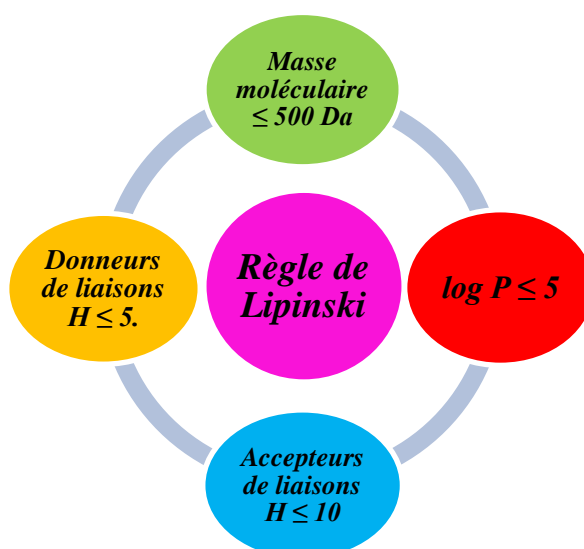


Figure.3: Schéma sur les Règles de Lipinski

Les propriétés de divers paramètres des composés (1-25) se trouvent dans le tableau 2.

Tableau. 2: Les règles de Lipinski des composés phénoliques.

Composés	MW	Log P	HBA	HBD
1	139.11	0,26	3	1
2	162.11	2,57	4	1
3	139.11	0,26	3	1
4	123.15	0,79	2	1
5	138.12	0,99	3	2
6	136.15	1,12	2	1
7	94.11	1,45	1	1
8	128.56	2,07	1	1
9	112.10	1,90	2	1
10	154.17	0,87	3	1
11	108.14	1,81	1	1

12	140.14	0,53	3	2
13	122.74	2,14	1	1
14	170.21	3,07	1	1
15	108.14	1,81	1	1
16	110.11	0,79	2	2
17	124.14	1,15	2	1
18	184.19	0,60	4	1
19	154.17	0,87	3	1
20	126.11	0,18	3	3
21	109.13	0,79	2	2
22	119.12	0,79	2	1
23	136.15	1,12	2	1
24	150.22	2,76	1	1
25	138.17	1,48	2	1

HBD : liaisons hydrogène donneurs, HBA : liaisons hydrogène accepteurs.

- A travers le tableau 2, tous les composés ont des valeurs moins de 5 pour la lipophilie, allant de 0,26 à 3,07. A cet égard, ces composés sont mieux solubilisés dans des solutions aqueuses et lipidiques.
- Nous pouvons également voir dans le tableau 2 que tous ces composés ont le nombre accepteurs d'hydrogène inférieurs à 10
- Le nombre donneurs d'hydrogène inférieurs à 5.
- Aussi, pour les poids moléculaires, tous les composés ont des valeurs inférieures à 500 Da, donc ils traversent facilement les membranes cellulaires, la valeur maximale de MW est de 184.19 Da pour le composé 18.

Donc, ces composés répondent aux règles de Lipinski (règles des cinq), en suggérant que ces composés théoriquement n'auraient pas de problèmes avec la biodisponibilité orale.

III.4.2. Règle de Veber

Les règles Veber suggèrent que la flexibilité moléculaire et la surface polaire (PSA) sont des déterminants importants de la biodisponibilité orale [17]. La réduction de la flexibilité moléculaire (mesurée par le nombre de liaisons rotatives) et la faible surface polaire sont des prédicateurs importants de la bonne biodisponibilité orale [21].

Les propriétés Veber des composés (1-25) se trouvent dans le tableau 3.

Tableau .3: Les règles de Veber des composés phénoliques

Composés	PSA	Nrotb
1	66,05	1
2	20,23	1
3	66,05	1
4	44,02	0
5	57,53	1
6	37,30	1
7	20,23	0
8	20,23	0
9	20,23	0
10	38,69	2
11	20,23	0
12	49,69	1
13	20,23	0
14	20,23	1
15	20,23	0
16	40,46	0
17	29,46	1
18	47,92	3
19	38,69	2
20	60,69	0
21	46,25	0
22	44,02	0
23	37,30	1
24	20,23	1
25	29,46	1

Pour la règle de Veber, nous remarquons que :

➤ Les valeurs de nombre de liaisons rotatives sont toutes inférieures à 10. Le faible nombre de liaisons rotatives (flexibilité réduite) dans les composés étudiés indique que ces ligands lors de la liaison avec une protéine ne changent que légèrement leur conformation.

➤ Pour les résultats de PSA, nous remarquons que tous les composés de la série étudiée ont des valeurs inférieures à 140 \AA^2 (PSA entre 20 et 130 \AA^2) ce qui montre la bonne prédiction de la biodisponibilité orale et le transport à travers les membranes.

Nous pouvons conclure que ces composés étudiés sont en accord avec la règle de Veber.

III.4.3. Triangle d'or (Golden triangle)

Le triangle d'or est un outil de visualisation développé par Pfizer à partir de données de perméabilité, de clairance in vitro et de calculs in vitro, conçu pour aider les chimistes à obtenir des candidats médicaments métaboliquement stables, perméables et puissants.[22]

Le classement des composés comme perméables et stables et la courbe de masse moléculaire (MW) en fonction de l'octanol: tampon (pH 7,4) (logD) ou du coefficient estimé d'octanol: tampon (pH 7,4) (logD) révèlent des tendances utiles. L'analyse d'au moins deux tendances orthogonales, telles que la perméabilité et la clairance, peut s'avérer extrêmement efficace pour équilibrer et optimiser plusieurs propriétés. De plus, des calculs de poids moléculaire et de puissance-efficacité logD impact permettant d'optimiser simultanément la puissance, la clairance et la perméabilité [18].

La lipophilie (LogD) et poids moléculaire (MW) ont été identifiés en tant que propriétés montrant la corrélation avec les données de perméabilité et de stabilité disponibles pour de grands ensembles de données. Toutes les tendances combinées conduisent à observer que la polarité et la perméabilité du poids moléculaire in vitro et les composés à faible clairance sont concentrés avec une ligne de base de $\log D = -2,0$ à $\log D = 5,0$ à $MW = 100$ et un sommet de $\log D = 1,0$ à $2,0$ et $MW = 450$. Ces tendances conduisent à une zone en forme connue sous le nom de Triangle d'or et de molécules dans cette zone qui: sont de faibles clairance et perméables devraient obéir à la règle du triangle d'or.

Tableau.4: Triangle d'or

Composés	MW	Log D (pH=7,4)
1	139.11	1,30
2	162.11	2,55
3	139.11	1,50
4	123.15	1,49
5	138.12	-1,20
6	136.15	1,22
7	94.11	1,67
8	128.56	2,27
9	112.10	1,81
10	154.17	1,35

11	108.14	2,18
12	140.14	1,21
13	122.74	2,70
14	170.21	3,32
15	108.14	2,18
16	110.11	1,36
17	124.14	1,51
18	184.19	1,20
19	154.17	1,35
20	126.11	1,04
21	109.13	0,83
22	119.12	1,40
23	136.15	1,12
24	150.22	3,21
25	138.17	2,03

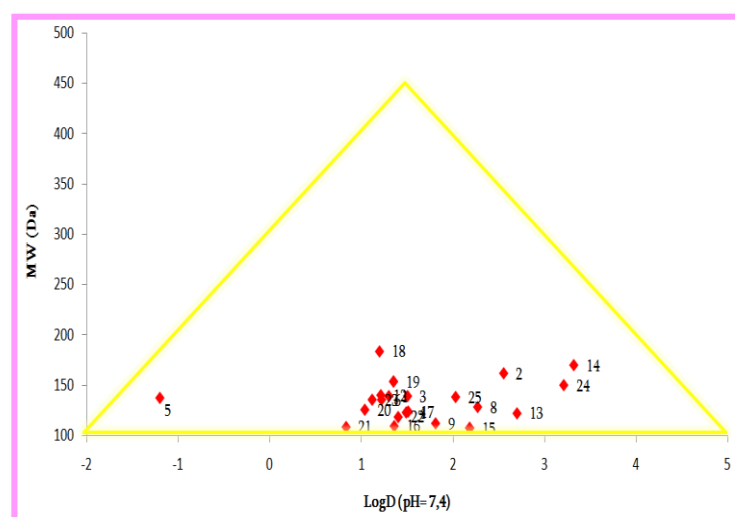


Figure.4: Triangle d'or (Golden triangle)

Johnson et ses collaborateurs ont indiqué que le poids moléculaire et la lipophilie (LogD à pH 7,4) remplacent de nombreux descripteurs moléculaires et ont été utilisés pour développer un outil de visualisation utile du triangle d'or.

Les composés qui résident dans le triangle d'or sont plus susceptibles d'être à la fois métaboliquement stables et de posséder une bonne perméabilité membranaire que ceux qui se trouvent à l'extérieur.

Le triangle d'or (Figure. 4) montre que tous les composés étudiés se situent à l'intérieur du triangle et que ces dérivés ont une bonne perméabilité et une bonne clairance.

III.5. Étude quantitative de la relation structure-activité d'une série de dérivés de phénol.

Dans cette partie de ce travail nous allons établir une relation linéaire structure-activité quantitative entre l'activité antioxydant d'une série dérivées de phénol et leurs descripteurs structuraux, par conséquent nous proposons un modèle quantitatif, et nous essayons d'interpréter l'activité de ces molécules en se basant sur la méthode statistique suivante : Régression linéaire multiple RLM.

III.5.1. Descripteurs

Dans ce travail les descripteurs ont été choisis pour décrire la structure des molécules constituant la série à étudier sont données dans le tableau 5.

Tableau.5 : Type des descripteurs moléculaires utilisés pour l'étude qualitative et quantitative des structure-activité.

<i>Descripteurs moléculaire</i>	<i>Type</i>
Réfractivité molaire (Réf)	<i>Propriétés physicochimique</i>
Polarisabilité (Pol)	
Lipophile (logP)	
Surface moléculaire (SAG)	
Energie totale (E _T)	<i>Electronique</i>
Moment dipolaire (MD)	
Charge (q _O , q _{C1})	
Energie de l'orbitale la plus haute occupée (E _{HOMO})	
Energie de l'orbitale la plus basse vacante (E _{LUMO})	
Gap énergétique (E _{gap}) (E _{gap} = E _{LUMO} - E _{HOMO})	
Potentiel d'ionisation (IP) (IP = -E _{HOMO})	
Affinité électronique (A) (A = -E _{LUMO})	
Dureté (η) (η = (IP-A)/2)	
Volume molaire (MV)	<i>Géométrique</i>
Masse moléculaire (WM)	<i>Constitutionnel</i>
Superficie de la surface polaire (PSA)	<i>Topologique</i>

Le tableau 6 regroupe tout les Descripteurs utilisés dans le développement de modèle QSAR

Tableau.6: Descripteurs utilisés dans le développement de modèle QSAR

Composés	Log P	Pol (Å ³)	MW (amu)	V (Å ³)	Réf (Å ³)	S (Å ²)	MD (D)	E _r (ev)	E _{romo} (ev)	E _{romo} (ev)	E _{sep} (ev)	IP	A	q _{ci}	q _o	η	PSA(Å ²)
1	0,26	12,91	139,11	413,60	37,29	268,11	7,513	-13935,248	-2,966	-7,157	4,191	7,157	2,966	0,481	-0,644	2,0955	66,05
2	2,57	12,63	162,11	437,13	33,47	280,02	4,553	-17542,777	-1,007	-6,912	5,904	6,912	1,007	0,463	-0,665	2,9525	20,23
3	0,26	12,91	139,11	414,59	37,29	266,42	8,005	-13935,166	-3,129	-7,021	3,891	7,021	3,129	0,467	-0,657	1,9460	66,05
4	0,79	14,26	123,15	435,13	33,18	244,40	7,707	-10879,565	-1,823	-6,912	5,089	6,912	1,823	0,485	-0,671	2,5445	44,02
5	0,99	13,63	138,12	415,66	35,42	252,40	4,594	-13501,847	-1,687	-6,857	5,170	6,857	1,687	0,492	-0,667	2,5850	57,53
6	1,12	14,83	136,15	445,77	38,66	274,93	5,049	-12523,675	-2,057	-6,694	4,653	6,694	2,041	0,475	-0,678	2,3265	37,30
7	1,45	11,07	94,11	345,12	28,46	213,91	1,837	-8368,871	-0,653	-6,504	5,850	6,504	0,653	0,469	-0,691	2,9255	20,23
8	2,07	13,00	128,56	388,85	33,48	247,73	3,024	-20875,185	-0,925	-6,504	5,578	6,504	0,925	0,465	-0,672	2,7895	20,23
9	1,90	10,98	112,10	353,42	28,42	223,29	2,790	11069,926	-0,925	-6,476	5,551	6,476	0,925	0,394	-0,668	2,7755	20,23
10	0,87	16,02	154,17	491,95	41,45	313,76	3,787	-14603,144	-0,027	-6,231	6,204	6,231	0,245	0,561	-0,646	2,9930	38,69
11	1,81	12,91	108,14	398,52	33,43	252,83	1,379	-9438,958	-0,626	-6,422	5,796	6,422	0,626	0,482	-0,691	2,8980	20,23
12	0,53	14,18	140,14	444,09	36,98	276,95	4,373	-13533,601	-0,272	-6,286	6,014	6,286	0,272	0,521	-0,667	3,0070	49,69
13	2,14	14,74	122,74	457,90	38,40	298,61	1,955	-10509,608	-0,517	-6,313	5,796	6,313	0,517	0,541	-0,697	2,8980	20,23
14	3,07	20,73	170,21	544,27	53,90	285,64	1,863	-14657,374	-1,333	-6,014	4,680	6,014	1,333	0,456	-0,685	2,3405	20,23
15	1,81	12,91	108,14	392,56	33,43	241,15	2,3899	-9439,431	-0,5412	-6,3835	5,842	6,383	0,541	0,374	-0,679	2,9212	20,23
16	0,79	11,71	110,11	365,45	30,49	222,91	3,0533	-10416,981	-0,5369	-6,3092	5,772	6,309	0,536	0,560	-0,695	2,8862	40,46
17	1,15	13,54	124,14	422,53	34,96	256,52	1,3635	-11486,467	-0,4356	-6,0877	5,652	6,087	0,435	0,303	-0,694	2,8260	29,46
18	0,60	18,49	184,19	569,22	47,94	335,73	1,1328	-17720,741	-0,4071	-6,0992	5,692	6,099	0,407	0,474	-0,667	2,8461	47,92
19	0,87	16,02	154,17	497,77	41,45	300,39	3,4484	-14603,746	-0,2310	-5,9878	5,756	5,987	0,231	0,063	-0,606	2,8785	38,69
20	0,18	12,35	126,11	378,97	32,51	212,09	4,0202	-12464,529	-0,3896	-6,1764	5,786	6,176	0,389	0,342	-0,619	2,8934	60,69
21	0,79	12,42	109,13	378,53	32,87	215,48	1,9775	-12556,598	-0,2668	-5,9135	5,546	5,913	0,366	0,255	-0,692	2,7734	46,25
22	0,79	12,92	119,12	398,32	33,18	264,48	6,8132	-10880,189	-1,5227	-6,8910	5,368	6,891	1,522	0,470	-0,656	2,6842	44,02
23	1,12	14,83	136,15	446,43	38,66	277,52	3,8345	-12524,402	1,85418	-6,7702	4,916	6,770	1,854	0,494	-0,670	2,4580	37,30
24	2,76	18,41	150,22	535,46	47,73	341,83	1,9228	-12 649,51	-0,5796	-6,2621	5,682	6,262	0,579	0,437	-0,684	2,8413	20,23
25	1,48	15,38	138,17	475,70	39,92	295,25	1,9775	-12 556,59	-0,3668	-5,9135	5,546	5,913	0,366	0,255	-0,692	2,7734	29,46

III.5.2. Corrélation entre différent descripteurs

D'entamer le développement effectif des équations de régression QASR, il est hautement recommandé d'examiner la qualité statistique des données de départ, à la fois les données à corrélérer (variable dépendante) et les descripteurs utilisés dans la corrélation (variables indépendantes).

Pour un ensemble de descripteurs différents, il est nécessaire d'effectuer une analyse des données bivariée, c'est-à-dire de calculer le coefficient de corrélation entre chacune des paires de l'ensemble des descripteurs. Si ce coefficient est statistiquement significatif ($R > 0,95$), ces deux descripteurs sont considérés comme fortement corrélés et ne peuvent être utilisés simultanément lors de l'analyse QSAR [23] et en pratique, ils seront alors enlevés dans le procédé de sélection.

Le tableau 7 représente la matrice de corrélation afin de vérifier les colinéarités des descripteurs.

Tableau.7: La matrice de corrélation entre différent descripteurs

	qC1	qO	E _{LUMO}	E _{HOMO}	MD	E _T	E _{gap}	Réf	Poi	MW	V	S	I	A	Logp	η	PSA
qC1	1,000																
qO	0,061	1,000															
E _{LUMO}	-0,301	-0,229	1,000														
E _{HOMO}	-0,501	-0,208	0,821	1,000													
MD	0,237	0,101	-0,765	-0,792	1,000												
E _T	-0,047	-0,045	0,047	-0,012	-0,104	1,000											
E _{gap}	-0,118	-0,200	0,931	0,556	-0,607	0,076	1,000										
Réf	-0,023	0,069	0,009	0,333	-0,198	-0,430	-0,199	1,000									
Poi	-0,028	0,061	0,108	0,385	-0,249	-0,398	-0,089	0,975	1,000								
MW	0,002	0,087	-0,058	0,102	0,038	-0,523	-0,149	0,816	0,778	1,000							
V	-0,046	0,060	0,123	0,328	-0,190	-0,438	-0,030	0,944	0,956	0,866	1,000						
S	0,058	0,080	0,078	0,156	-0,107	-0,404	0,015	0,817	0,800	0,787	0,911	1,000					
I	0,101	0,102	-0,265	-0,048	0,158	0,218	-0,355	0,015	0,015	-0,325	-0,118	-0,132	1,000				
A	0,230	0,106	-0,573	-0,691	0,522	-0,200	-0,393	-0,098	-0,168	0,270	-0,056	0,027	-0,626	1,000			
Logp	0,061	-0,122	0,223	0,185	-0,503	0,089	0,206	0,279	0,328	0,105	0,241	0,228	-0,408	0,143	1,000		
η	-0,139	-0,204	0,930	0,557	-0,619	0,085	0,997	-0,215	-0,104	-0,168	-0,047	-0,006	-0,362	-0,389	0,218	1,000	
PSA	0,030	0,097	-0,460	-0,321	0,647	-0,236	-0,464	-0,031	-0,123	0,158	-0,062	-0,091	0,331	0,087	-0,871	-0,475	1,000

D'après les résultats présents dans le tableau 7, il s'avère que l'analyse de la matrice de corrélation montre que :

- La polarisabilité est parfaitement corrélée avec la réfractivité molaire pour $r = 0,975$.
- La polarisabilité est parfaitement corrélée avec le volume $r = 0,956$.

Donc le descripteur hautement corrélé a été exclu

III.5.3. Equation du modèle QSAR

Le tableau 8 rassemble les descripteurs sélectionnés, coefficients et p-value (pour un seuil de signification inférieur à 0,05) pour un meilleur modèle.

Tableau.8: Les descripteurs sélectionnés par MLR

Modèle	coefficient	p-value
constant	-1.998	0.000
qC1	0.310	0.040
η	0.877	0.000
PSA	-0.003	0.017
A	0.446	0.000

Donc l'équation de la régression linéaire obtenue est la suivante (équation 1):

$$E_7 = -1,998 + 0,310qC1 + 0,877\eta - 0,003PSA + 0,446A \quad 1$$

$$N=25; R=0.955; R^2=0.911; S=0.0663; F=51,359$$

Interprétation du modèle:

- Les valeurs de la fraction de la variance peuvent varier entre 0 et 1. Le modèle QSAR doit considérer un $R^2 > 0.6$ pour qu'il soit valide [24]. Les valeurs $R = 0,955$ et $R^2 = 0,911$ nous permis d'indiquer fermement la corrélation entre les différents descripteurs utilisés tels que: qC1, η, PSA et A et l'activité biologique étudiée (antioxydant).
- La valeur de F calculée est supérieure à la valeur lue dans la table du Fischer (51,359 > 2,87) (Tableau. 5), et indique donc que la relation de régression était statistiquement significative [24].
- Dans ce modèle (équation 1), nous remarquons que toutes augmentations de qC1, A et η et que la diminution de PSA provoquent à une augmentation de potentiel redox E_7 (diminution de l'activité antioxydante).

- La matrice de corrélation pour l'activité antioxydant et les descripteurs sélectionnés pour construire le modèle QSAR est présentée dans le tableau 10. Les paramètres utilisés dans ce modèle sont presque indépendants, ce qui peut être vu à partir de la matrice de corrélation.
- Nous avons remarqué aussi que les descripteurs qui forment le modèle sont de type électronique, ils sont responsables d'améliorer l'activité antioxydant.
- Le tableau 10 indique l'importance des paramètres électronique dont l'affinité électronique est la plus important (76,6%) dans la description de l'activité spécifique de phénol.

Tableau.9: Table de Fischer

F limite à p 0,05

TABLE du F de FISCHER (comparaison de 2 variances)

degré de liberté du numérateur

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	30	40	50
1	161	199	216	225	230	234	237	239	241	242	243	244	245	245	246	248	250	251	252
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,66	8,62	8,59	8,58
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,80	5,75	5,72	5,70
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,56	4,50	4,46	4,44
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,87	3,81	3,77	3,75
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,44	3,38	3,34	3,32
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,15	3,08	3,04	3,02
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,94	2,86	2,83	2,80
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,77	2,70	2,66	2,64
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,65	2,57	2,53	2,51
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,54	2,47	2,43	2,40
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,46	2,38	2,34	2,31
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,39	2,31	2,27	2,24
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,33	2,25	2,20	2,18
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,28	2,19	2,15	2,12
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,23	2,15	2,10	2,08
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,19	2,11	2,06	2,04
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,16	2,07	2,03	2,00
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,12	2,04	1,99	1,97
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,10	2,01	1,96	1,94
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,07	1,98	1,94	1,91
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,05	1,96	1,91	1,88

degré de liberté du dénominateur

Tableau.10: La matrice de corrélation du modèle.

	E ₇	A	q _{c1}	η	PSA
E ₇	1				
A	0,766	1			
q _{c1}	0,593	0,318	1		
η	-0,519	-0,929	-0,139	1	
PSA	-0,295	-0,162	-0,074	0,152	1

III.5.4. Validation de modèle

Les puissances prédictives des équations ont été validées par la méthode de validation croisée "leave-one-out" (LOO) [25-27]. La validation croisée est une méthode pratique et fiable pour tester la signification d'un modèle.

Afin de tester la validité de pouvoir prédictif de modèle QSAR sélectionné, la technique de validation croisée LOO (leave-one-out) a été utilisée. Le modèle développé a été validé par le calcul des paramètres statistiques suivants: la somme des carrés résiduels (PRESS), la somme des carrés totaux (TSS ou SSY), qualité d'ajustement R^2_{adj} , coefficient de corrélation de validation croisée R^2_{cv} et la validation standard des erreurs de prédiction (SPRESS) (Tableau .11).

Tableau.11: Paramètres de validation croisée.

PRESS	SSY	PRESS/SSY	SPRESS	R^2_{cv}	R^2_{adj}
0,088	0,990	0,088	0,059	0,911	0,894

PRESS est un important paramètre de validation croisée car il s'agit d'une bonne approximation de l'erreur prédictive réelle du modèle. Sa valeur étant inférieure à SSY indique que ce modèle prédit mieux que le hasard et peut être considéré statistiquement significatif. La petite valeur de PRESS signifie la meilleure prévisibilité du modèle [26]. D'après les résultats présentés dans le tableau 11, cette valeur est égale à 0,088, le modèle est statistiquement significatif.

De plus, pour un modèle QSAR fiable, le rapport PRESS / SSY devrait être inférieur à 0,4 [24]. D'après les données présentées dans le tableau 7 indiquent que pour le modèle développé, ce rapport est de 0.088. Les valeurs élevées du R^2_{cv} et R^2_{adj} sont des critères essentiels pour une meilleure qualification des modèles QSAR [28]. Les résultats obtenus de R^2_{cv} et R^2_{adj} pour ce modèle QSAR est 0.911 et 0.894.

SPRESS (incertitude de prédiction) est un bon paramètre à utiliser pour décider de l'incertitude dans la prédiction. Plus la valeur de ce paramètre est basse, la capacité

prédictive du modèle sera mieux. Le modèle QSAR développé présente une faible valeur de S_{press} (<0.200) ce qui indique que ce modèle présente des petites valeurs résiduelles entre l'activité biologique observée et prédite. Dans notre cas, ce paramètre porte une petite valeur 0.059 ce qui explique que l'habileté de prédiction est la meilleure pour ce modèle.

L'activité antioxydant expérimentale, prédite et résiduelle de phénol et ses dérivés sont représentés dans le tableau 12 ont été déduit par logiciel SPSS [10]. Le résidu est dû d'une part à l'incertitude de la mesure expérimentale et d'autre part à l'imperfection du modèle, donc si l'activité prédite est plus proche de l'activité expérimentale, le modèle peut être applicable [29].

Tableau.12: Valeurs expérimentales, prédites et résiduelles de E_7 des dérivées de phénol

Composé	E_7_{EXP}	E_7_{PRED}	E_7_{RESID}
1	1.23	1,257	-0,027
2	1.13	1,130	0,000
3	1.13	1,194	-0,064
4	1.11	1,079	0,031
5	1.04	1,020	0,020
6	0.98	1,000	-0,020
7	0.97	0,950	0,020
8	0.94	0,951	-0,011
9	0.93	0,917	0,013
10	0.85	0,807	0,043
11	0.85	0,918	-0,068
12	0.84	0,789	0,051
13	0.84	0,888	-0,048
14	0.84	0,736	0,104
15	0.82	0,867	-0,047
16	0.81	0,838	-0,028
17	0.77	0,690	0,080
18	0.66	0,699	-0,039
19	0.58	0,546	0,034
20	0.58	0,657	-0,077
21	0.41	0,553	-0,143
22	1.17	1,063	0,107
23	1.06	1,038	0,022
24	0.80	0,834	-0,034
25	0.68	0,598	0,082

La figure ci-dessous (figure. 5) présente la corrélation entre les valeurs prédites et expérimentales pour l'activité biologique étudiée des dérivés de phénol. Du point de vue des performances statistiques, ce modèle présente de coefficient de corrélation R^2 significatif de valeur. La valeur de R^2 est égale à 0.911, pour ce modèle. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité étudiée.

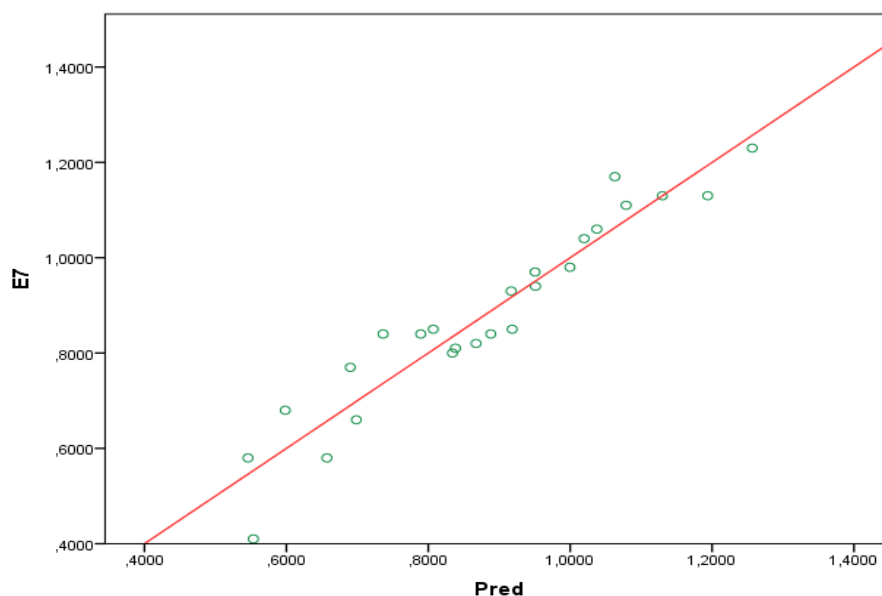


Figure.5: Corrélation entre l'activité biologique expérimentale et prédite.

Pour rechercher la présence d'une erreur systématique dans l'élaboration des modèles QSAR, les valeurs résiduelles de l'activité biologique ont été tracées en fonction des valeurs expérimentales, comme le montre la figure 6.

La répartition des résidus dans les deux côtés de l'axe des abscisses indique l'absence de l'erreur systémique [30]. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité biologique antioxydante pour l'ensemble des phénols utilisés dans le développement de modèle QSAR.

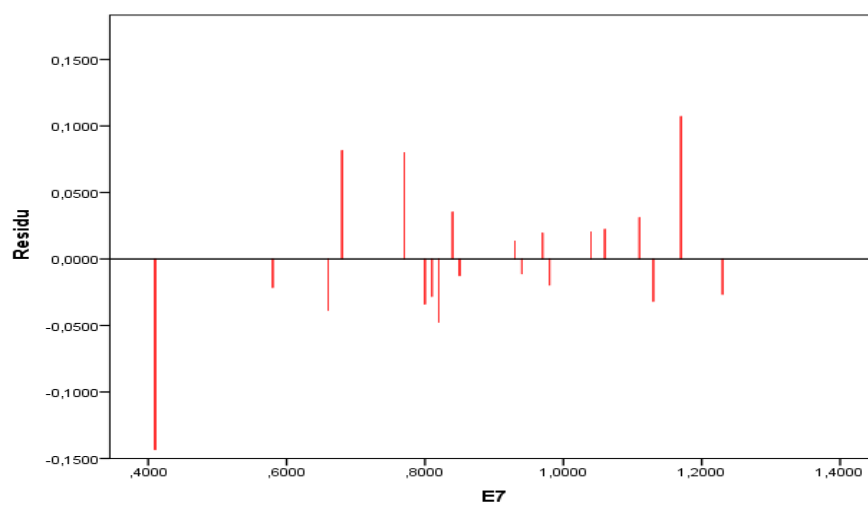


Figure.6: Corrélation entre l'activité biologique expérimentale et résiduelle.

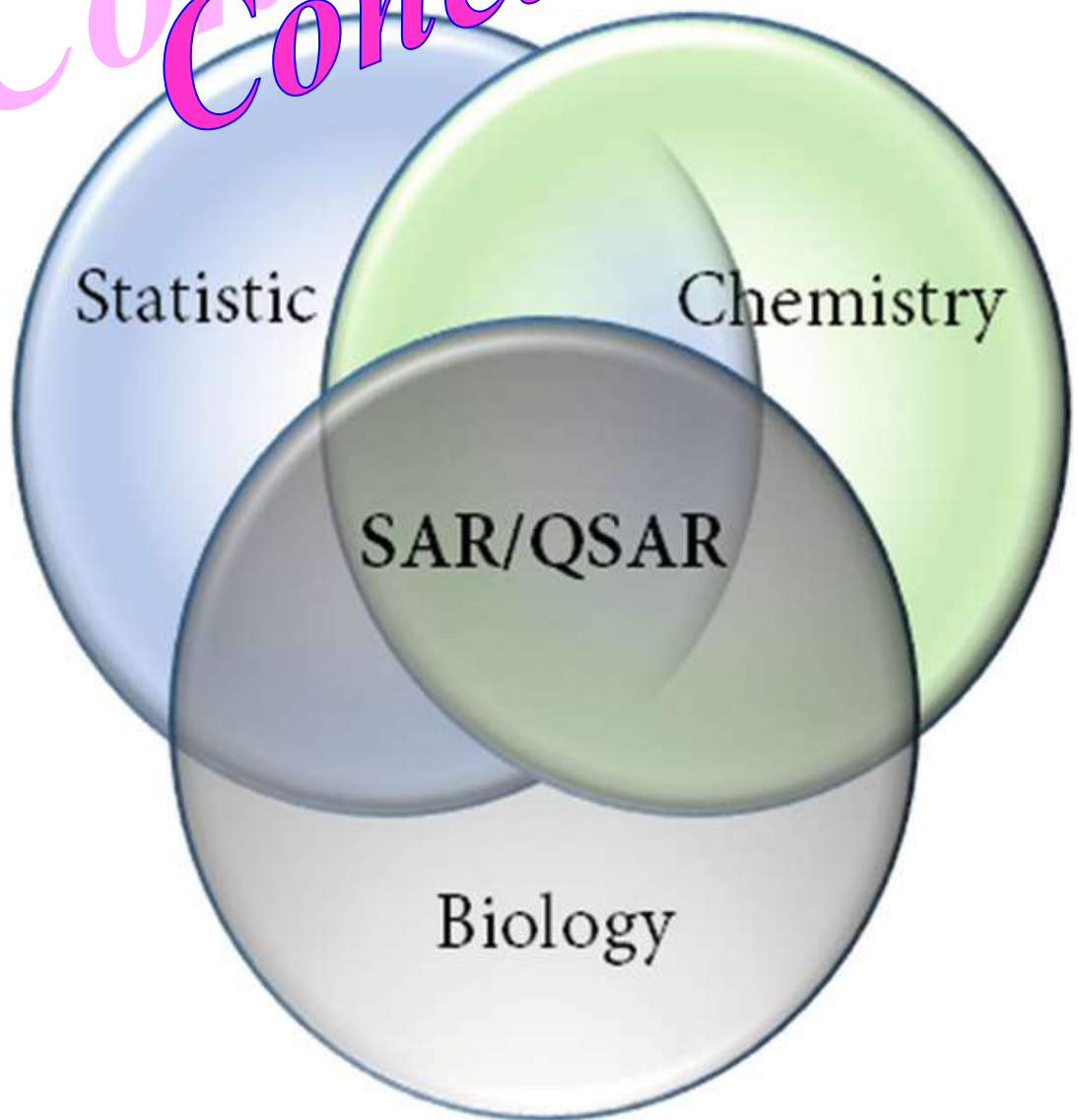
Références bibliographiques

- [1] M. Reis, B. Lobato, J. Lameira, A. S. Santos, C. N. Alves, A theoretical study of phenolic compounds with antioxidant properties, *Europ. J. Med. Chem*, 42, 6-440, 2007.
- [2] E. S. Lien, S. Ren, H. H. Bui, R. Wang, Quantitative structure-activity relationship analysis of phenolic antioxidants, *Free Radic. Biol. Med*, 26, 94-285, 1999.
- [3] S. Steenken, P. Neta, One-electron redox potentials of phenols. Hydroxy- and aminophenols and related compounds of biological interest, *J. Phys. Chem*, 86, 3661-3667, 1982.
- [4] S.V. Jovanovic, S. Steenken, M.G. Simic, Y. Hara, C.A. Rice-Evans, L. Packers, Iron complexes of galliccatechins. Antioxidant action or iron regulation, *J. Chem. Soc., Perkin Trans. 2*, Marcel Dekker, Inc., New York, 11, 2365-2370, 1998.
- [5] F.G. Bordwell, J.P. Cheng, Substituent effects on the stabilities of phenoxyl radicals and the acidities of phenoxyl radical cations, *J. Am. Chem. Soc*, 113, 1736-1743, 1991.
- [6] HyperChem v8. Molecular Modelling System, Hypercube Inc., 1115 NW 4th Street, Gainesville, FL 32601, USA, 2009.
- [7] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Gaussian 09*, Gaussian, Inc., Wallingford CT, 2010.
- [08] C. M. Breneman, K. B. Wiberg, Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis, *J. Compu. Chem*, 11, 361-373, 1990.
- [09] MarvinSketch, ChemAxon (<http://www.chemaxon.com>).
- [10] Spss 20 For Windows, SPSS software packages, SPSS Inc, 444 North Michigan Avenue, Suite 3000, Chicago, Illinois, 60611, USA. <http://www.spss.com>.
- [11] N. I. Zhokhova, I. I. Baskin, V. A. Palyulin, A. N. Zefirov, N. S. Zefirov, Fragmental descriptors in QSPR: Application to molecular polarizability calculations, *Russ. Chem*, 52, 1061-1065, 2003.

- [12] R. Mignani, A. Petrucci, F. Cardone, *Quantum Matter*, 3, 219, 2003.
- [13] A. K. Ghose, G. M. Crippen, Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions, *J. Chem. Inf. Comput. Sci*, 27, 21-35, 1987.
- [14] G. Burgot, N. Bona, M. V. Christen, J. L. Burgot, First relationships between detoxication properties of some 1,2-dithiole-3-thiones and their log *P*, *Int. J. Pharm*, 129, 295-299, 1996.
- [15] D. M. Segall, Multi-parameter optimization: identifying high quality compounds with a balance of properties, *J. Curr Pharm. Des*, 18, 1292-1310, 2012.
- [16] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *J. Adv. Drug Delivery Rev*, 64, 4-17, 2012.
- [17] D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, K. D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, *J. Med. Chem*, 45, 2615-2623, 2002.
- [18] T.W. Johnson et al. Using the Golden Triangle to optimize clearance and oral absorption, *Bioorg. Med. Chem. Lett*, 19, 5560-5564, 2009.
- [19] G. Vistoli, A. Pedretti, B. Testa, Assessing drug-likeness, *Drug Discov. Today*, 13, 285-294, 2008.
- [20] M. Aurélien, Création et utilisation de chimiothèques optimisées pour la recherche « in silico » de nouveaux composés bioactifs, Thèse de Doctorat, 2006.
- [21] M. Remko, M. Swart, F. M. Bickelhaupt, Theoretical study of structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some centrally acting antihypertensives, *J. Bioorg. Med. Chem*, 14, 1715-1728, 2006.
- [22] J. M. Beale, J. Block, R. Hill, *Organic Medicinal and Pharmaceutical Chemistry*, Philadelphia: Lippincott Williams and Wilkins, New York, 2010.
- [23] W.A. Wakeham, G.S. Cholakov, R.P. Stateva, Liquid density and critical properties of hydrocarbons estimated from molecular structure||, *J. Chem & Eng. Data*, 47, 559-570, 2002.
- [24] B.P. Ajeet, K. Bijander, Quantitative structure activity relationship (QSAR) modeling of 2-x-5,9-dimethoxy-1,4-naphthoquinones against 11210 cells, *Int. J. Pharm. Sci*, 4, 445-448, 2012.
- [25] M. Clark, R.D. Cramer. The probability of chance correlation using partial least squares (PLS). *Quant. Struct.-Act. Relat*, 12, 45-137, 1993.
- [26] S. Wold, L. Eriksson, Statistical validation of QSAR results, *Validation tools. Methods and Principles in Medicinal Chemistry*, 2, 18-309, 1995.

- [27] H. Van der Voet, Comparing the predictive accuracy of models using a simple randomization test, *Chemometr. Intell. sys. Lab*, 25, 23-313, 1994.
- [28] S.O. Podunavac-Kuzmanović, D.D. Cvetković, D.J. Barna, QSAR analysis of 2-amino or 2-methyl-1-substituted benzimidazoles against *Pseudomonas aeruginosa*, *Int. J. Mol. Sci*, 10, 1670-1682, 2009.
- [29] H. Gonzalez-Diaz, E. Olazabal, L. Santana, E. Uriarte, Y. Gonzalez-Diaz, N. Castanedo, QSAR study of anticoccidial activity for diverse chemical compounds: Prediction and experimental assay of trans-2-(2-nitrovinyl) furan, *Bioorg. Med. Chem*, 15, 962-968, 2007.
- [30] M. Jalali-Heravi, A. Kyani, Use of computer-assisted methods for the modeling of the retention time of a variety of volatile organic compounds, *J. Chem. Inf. Comput. Sci*, 44, 1328-1335, 2004.

Conclusion Conclusion



Dans ce travail nous avons effectué une étude qualitative de la relation structure / activité sur une série de vingt-cinq dérivés de phénols. Les molécules étudiées présentent un intérêt pharmacologique ou bien biologique. La nature des groupements qui se lient sur le noyau de base des molécules étudiées affecte sur leurs propriétés physicochimiques et par conséquence sur leurs propriétés pharmacologiques.

L'application des règles de Lipinski aux dérivés de phénol étudiés montre que les composés étudiés ne présentent théoriquement aucun problème de biodisponibilité orale.

Notre travail est ensuite porté l'étude la règle de Veber explique que tous les composés de la série ayant le nombre des liaisons inférieure à 10, et aussi que tous les composés de notre série avec des valeurs de la surface polaire sont en accord à le règle 2^{ème} de Veber, c'est-à dire tous les valeurs de la surface polaire sont inférieurs à 140 Å², et ces résultats montre une bonne corrélation avec le transport moléculaire passive à travers les membranes. Donc, il permet l'estimation des propriétés de transport des médicaments.

Tous les composés phénoliques étudiés se situent à l'intérieur du triangle d'or, de sorte que ces dérivés ont une bonne perméabilité et une bonne clairance.

En outre, l'étude de QSAR a permis de déterminer la relation quantitative entre la structure moléculaire des composés avec leurs activités étudiées.

L'analyse de régression MLR a été utilisée pour développer le modèle et prédire de l'activité biologique à partir de descripteurs moléculaires appartenant à la série dérivés de phénol. Notre modèle de QSAR développé est basé sur les descripteurs électroniques suivants: qC1, A, PSA, η . Ces derniers ont un rôle très important dans le mécanisme par lequel un antioxydant peut désactiver un radical libre par transfert d'électrons.

Le modèle QSAR indique que ces descripteurs ont des relations significatives avec l'activité observée. Nous avons observé une similitude élevée entre les valeurs expérimentales et les valeurs prédites de l'activité, qui indique l'excellente qualité du modèle QSAR.

Résumé

Ces dernières années, l'intérêt porté aux antioxydants, en relation avec leurs propriétés thérapeutiques, pharmacologiques... a augmenté considérablement.

Une étude qualitative de la relation structure-activité (SAR) a été effectuée pour une série de dérivés du phénol. La nature des groupements sur le noyau hétérocyclique des molécules étudiées affecte leurs propriétés physicochimiques et par conséquent sur leurs propriétés pharmacologiques. Une étude QSAR a été effectuée sur 25 dérivées de phénol. La régression linéaire multiple (MLR) a été utilisée pour quantifier la relation entre les descripteurs moléculaires et l'activité antioxydante des composés phénolique. Le pouvoir prédictif du modèle obtenu a été confirmé par $R = 0,955$, $R^2 = 0,911$, $S = 0,066$ et $F = 51,359$. La méthode de validation croisée LOO nous montre une forte corrélation qui a été observée entre les valeurs expérimentales et prédites des activités biologiques, ce qui indique la validité et la qualité des modèles QSAR obtenus.

Mot-clés: Activité antioxydante, Composés phénolique, Radical libre, SAR, QSAR, MLR.

Abstract

In recent years, the interest in antioxidants, in relation to their therapeutic and pharmacological properties, has increased considerably.

A qualitative study of the structure- / activity relationship (SAR) was performed for a series of phenol derivatives. The nature of the groups on the heterocyclic nucleus of the molecules studied affects their physicochemical properties and consequently on their pharmacological properties. A QSAR study was performed on 25 phenol derivatives. Multiple linear regression (MLR) was used to quantify the relationship between molecular descriptors and the antioxidant activity of phenolic compounds. The predictive power of the model obtained was confirmed by $R = 0.955$, $R^2 = 0.911$, $S = 0.066$ and $F = 51.359$. The LOO cross-validation method shows us that a strong correlation was observed between the experimental and predicted values of the biological activities, which indicates the validity and quality of the QSAR models obtained.

Keywords: Antioxidant activity; Phenolic compounds; Free radical, SAR, QSAR, MLR.