



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

MÉMOIRE DE MASTER

Domaine : Sciences de la Matière
Filière : chimie
Spécialité : chimie pharmaceutique

Réf. :

Présenté et soutenu par :

Djahara Imene

Gacem Yasmine

Le : .. / .. / 2021

Etude QSAR à l'aide de descripteurs moléculaires théoriques de dérivés xanthone.

Jury :

Mme Bounab Karima	M.C.A	Université Med Khider- Biskra	Président
Mme Youcef Oumhani	M.C.B	Université Med Khider- Biskra	Rapporteuse
Mme Aggoune Sihem	M.A.A	Université Med Khider- Biskra	Examinatrice

Année universitaire : 2020-2021

Remerciements

Ce travail a été réalisé à l'Université Med Khider Biskra,
Département des sciences de la matière, au Laboratoire de Chimie
Moléculaire et Environnement (LCME) au sein de l'équipe de Chimie
Informatique et Pharmaceutique.

Avant tout je remercie dieu pour tous puissant pour m'avoir donné le
courage, la foi et la force de continuer mes études.

Dieu merci.

Je tiens tout d'abord à exprimer ma sincère gratitude au pour la
confiance qu'il a bien voulu m'accorder en acceptant de diriger ce
mémoire pour la qualité de son encadrement, sa disponibilité, ses
conseils précieux et particulièrement ses encouragements

Youcef Oumhani.

Bounab Karima M.C.A à Université Med khider Biskra, qui a bien voulu accepter la
présidence du jury de soutenance.

Mes sincères remerciements vont aussi au **Aggoune Sihem M.A.A** à Université Med khider
Biskra pour avoir accepté d'examiner mon travail.

C'est un grand honneur pour moi de remercier le professeur
N.Melkemi pour son soutien et ses conseils, Vous avez tout
notre respect et notre appréciation.

Enfin, j'adresse mes remerciements à tous les collègues de l'équipe de
«chimie informatique et pharmaceutique» du laboratoire de recherche
LMCE et à tous ceux qui ont contribué de près ou de loin à la
réalisation de ce travail.

Dédicaces

A l'homme de ma vie, mon exemple éternel, mon soutien moral et source de joie et de bonheur, celui qui s'est toujours sacrifié pour me voir réussir, que dieu te préserver et t'accorder santé, longue vie à toi mon père

Mohamed.

A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur, maman **Fatima** que j'adore.

A mon fiancé : **Saad**

Toujours m'encourager et me soutenir, merci d'être dans ma vie.

A mon frère : **Youcef**

Mes sœurs que j'aime : **Nora, Samra, wahiba** et petite fils **Saif Eddine**

Randa et ses petites enfants **Nizar** et **Anes**

Mon binôme : **Yasmine Gacem**

Mes amis qui ont toujours été à mes côtés et avec eux j'ai partagé les meilleurs moments de ma vie.

Djahara Imen

Dédicaces

Je dédie ce modeste travail à

A mes parents : **Mohamed Saleh, Amina** ;Aucun hommage ne

pourrait être à la hauteur de l'amour Dont ils ne cessent

de me combler. Que dieu leur procure

bonne santé et longue vie.

A celui que j'aime beaucoup et qui m'a soutenue tout au

long de ce projet : mon homme : **Mokhter**

et bien sur mes frères : **Mohamed Read, Nourdin, Omer** et mes petites

enfants **Youssef ,Jihan, Halim** mes petites

enfants **Meriem, Adem**

A mon chère binôme : **Imene Djahara**

Et à tous ceux qui ont contribué de près ou de loin pour

que ce projet soit possible, je vous dis merci.

Gacem Yasmine

Sommaire

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES ABREVIATIONS

INTRODUCTION GENERALE.....1

Références bibliographiques.....3

CHAPITRE I : Les propriétés des dérivés xanthone et leurs applications médicinales (la maladie de diabète).

I.1. Introduction.....4

I.2. Alpha-glucosidase dans le traitement du diabète.....5

I.2.1. Le diabète.....5

I.2.1.1. Historique.....5

I.2.1.2. Présentation générale des différents types de diabète.....5

I.2.1.2.1. Le diabète de type 1.....5

I.2.1.2.2. Le diabète de type 2.....6

I.2.1.2.3. Les traitements actuels du diabète.....7

a- Les traitements du diabète de type 1.....7

b- Les traitements du diabète de type 2.....7

I.3. Caractérisation des α -glucosidases humaines et de leur mécanisme d'action...8

I.3.1. L'amidon.....8

I.3.1.1. Composition des granules d'amidon de maïs.....8

I.3.1.2. Organisation structurale des granules de l'amidon.....8

I.3.1.3. Réactivité des amidons et mode d'action des amylases9

I.3.1.3.1. Amylases.....9

a) Endoamylases.....9

b) Exoamylases.....	9
c) Enzymes de débranchement.....	10
d) Enzymes du quatrième groupe.....	10
I.3.1.3.2. Activité des amylases.....	10
I.3.1.3.3. Procédés d'hydrolyse enzymatique des amidons.....	11
I.3.1.3.4. Mode d'action des amylases.....	11
I.4. Le xanthone.....	13
I.4.1. Historique.....	13
I.4.2. Structure du xanthone.....	13
I.4.3. Applications.....	15
Références bibliographiques.....	17
CHAPITRE II : La modélisation moléculaire	
II.1. Introduction.....	22
II.2. Représentation des méthodes de calculs.....	22
II.2. 1. Mécanique quantique.....	22
II.2. 1.1. Base de la chimie quantique.....	23
II.2. 1.2. Les méthodes ab-initio.....	24
II.2.1.3. Théorie de la fonctionnelle de la densité (DFT)	24
II.2.1.4. Fondements de la théorie de la fonctionnelle de la densité.....	24
II.2.1.5. Méthodes semi-empiriques.....	25
II.2.2. Mécanique moléculaire.....	27
II.2.2.1. Champ de force.....	28
II.2.2.2. Quelques champs de force.....	30
II.2.2.3. Domaine d'application de la mécanique moléculaire.....	31
II.2.3. Dynamique moléculaire.....	32
II.2. 4. Méthodes de corrélation quantitative (QSAR)	32
II.2.4.1. Paramètres biologiques.....	32

II.2.4.2. Descripteurs moléculaires.....	33
II.2.4.3. Volume et surface moléculaire.....	33
II.2.4.4. Polarisabilité	35
II.2.4.5. Réfractivité molaire.....	35
II.2.4.6. Energie d'hydratation.....	35
II.2.4.7. Coefficient de partition (Log P)	37
II.2.4.8. Modèles statistiques.....	39
II.2.4.9. Les applications de l'étude QSAR/QSPR.....	39
II.3. La Régression Linéaire Multiple (MLR)	40
II.4. IC50 : Constante d'inhibition d'un complexe.....	41
Références bibliographiques.....	42
Chapitre III : Etude Quantitative De la Relation Structure/Activité Biologique	
III.1. Etude QSAR à l'aide de descripteurs moléculaires théoriques des dérivés xanthone	
III.1.1. Introduction.....	45
III.1.2. Méthodologie.....	45
III.1.2.1. Base de données.....	45
III.1.2.2 Sélection des descripteurs et méthodes de calcul.....	47
III.1.3. Développement des modèles QSAR.....	50
III.1.4. Validation de modèle QSAR.....	52
Références bibliographiques.....	56
Conclusion générale.....	58
Annexe.....	59

Liste des tableaux

Liste des tableaux	Page
<i>Chapitre I</i>	
Tableau I .1: Exemples de polysaccharides d'origine microbienne.	13
Tableau I.2: Applications alimentaires et non-alimentaires du xanthone.	16
<i>Chapitre II</i>	
Tableau II.1: Types de données biologiques utilisées dans l'analyse QSAR.	33
<i>Chapitre III</i>	
Tableau III.1 : Structures chimique et activité biologique expérimentale des molécules utilisées dans l'étude QSAR.	46
Tableau III.2 : Valeurs des descripteurs physicochimique utilisés dans l'étude QSAR.	47
Tableau III.3 : Valeurs des descripteurs quantique utilisés dans l'étude QSAR	49
Tableau III.4 : matrice de corrélation	52
Tableau III.5 : Paramètres de validation croisée	52
Tableau III.6: Les valeurs expérimentales, prédites et résiduelles.	53

Liste des figures

Liste des Figures	Page
<i>Chapitre I</i>	
Figure I .1: Schéma de dégradation de l'amidon par les différentes enzymes	12
Figure I .2: Structure (β/α) 8 (A) ou TIM barrel (B) contenant le site catalytique.	12
Figure I.3: Structure chimique du xanthone	14
Figure I.4: Structure chimique du xanthone selon Kool et al.21	14
<i>Chapitre II</i>	
Figure II.1: Représentation mécanique d'une structure moléculaire.	28
Figure II.2: Atome entouré par une grille.	34
Figure II.3: Atomes A, B et la zone de recouvrement X.	34
Figure II.4: La liaison hydrogène.	36
Figure II.5: Polarité des différents milieux cellulaires.	37
<i>Chapitre III</i>	
Figure III.1: Structures chimiques des dérivés de xanthone	45
Figure III.2 : corrélation entre les activités biologiques expérimentales et prédites pour le modèle.	55
Figure III.3: Représentations graphique des valeurs résiduelles en fonction des valeurs expérimentales pour le modèle.	55

Liste des abréviations

OMS	Organisation Mondiale de la Santé
3D	Trois dimensions.
2D	Deux dimensions
Ab-initio.	en anglais first principales.
OM-CLOA	Combinaison Linéaires des Orbitales Atomiques ou Moléculaires.
LCAO	Combinaison Linéaire d'Orbitale Atomique.
DFT	Density Functional Theory
CNDO	Complete Neglect of Differential Overlap.
PM3	Parametrization Model 3.
HOMO	Orbitale moléculaire la plus haute occupée.
LUMO	Orbitale moléculaire la plus basse inoccupée.
INDO	Intermediate Neglect of Differential Overlapping.
IC50	Inhibitory Concentration of 50%.
MINDO/3	modified intermediate neglect of differentail, overlap 3.
MNDO	Modified Neglect.
AM 1	Austrin Model 1.
SAM 1	Semi-ab-intio Model 1.
MM	La mécanique moléculaire.
OPLS	Optimised Potentials for Liquid Simulation.
GROMOS	Groningen Molecular Simulation Program Package.
CHARMM	Chemistry Harvard Macromolecular Mechanic.
QSAR.	Relation Structure Activité Quantitative.
MR.	Réfractivité Molaire.
Pol	Polarisabilité
MW	masse moléculaire.
HE	l'énergie d'hydratation
MLR	Multiple Linear Regression
Log P	Coefficient de Partage

V	Volume
B3LYP	Becke 3-Parameter Lee-Yang-Parr
SAG	Surface Area Grid
ΔE	Gap énergétique $E_{\text{HOMO}} - E_{\text{LUMO}}$.
HBD	Donneur de liaison hydrogène
HBA	Accepteur de liaison hydrogène
DM	Dipôle Moment
LOO	Validation croisée (leave-one-out).
SPRESS	SPRESS : Validation standard des erreurs de prédiction.
PE	PE : Erreur de prédiction.
R_{CV}^2	Coefficient de corrélation de validation croisée.
R_{adj}^2	Qualité d'ajustement.
HBA	Hydrogen Bond Acidity
HBB	Hydrogen Bond Basicity
\hat{H}	Opérateur hamiltonien
Ψ	fonction d'onde
SSY	Sum of Squares in Y.

Introduction

générale

Introduction générale

Vers la fin des années 1950, Allene R. Jeanes [1] du département américain de l'Agriculture découvre qu'en présence de sucre, une bactérie isolée des feuilles de chou, *Xanthomonas campestris*, est capable de produire un polysaccharide extracellulaire, la gomme de xanthone.

L'intérêt industriel de ce nouveau polysaccharide se dessine rapidement et repose notamment sur ses propriétés rhéologiques exceptionnelles. Le xanthone possède un important pouvoir épaississant même pour de faibles concentrations en polymère, ainsi qu'un caractère fortement rhéofluidifiant. Il est soluble en solution aqueuse à chaud et à froid. Ces solutions sont stables pour une large gamme de concentration (jusqu'à 150 g/L en NaCl), de température (jusqu'à 90°C) ainsi que de pH (entre 2 et 11) [2]. Une autre propriété intéressante du xanthone concerne sa capacité à interagir avec d'autres polysaccharides comme les galactomannanes (caroube, guar ...). En effet, leur addition à des solutions de xanthone entraîne une augmentation synergique de la viscosité. C'est-à-dire que les mélanges possèdent une viscosité bien supérieure à celles des solutions de polysaccharides seuls [3], [4].

Les glycosidases (glycoside hydrolases) sont des enzymes très répandues et impliquées dans un grand nombre de processus biologiques fondamentaux du métabolisme des sucres : la digestion des polysaccharides alimentaires dans la bouche et l'intestin grêle, la dégradation des polysaccharides au niveau du lysosome, ou encore la glycosylation des protéines au niveau du réticulum endoplasmique et de l'appareil de Golgi [5].

Les glycosidases catalysent la dégradation hydrolytique des polysaccharides en oligosaccharides et en hydrates de carbone monomères. De toutes les enzymes présentes dans la nature, ce sont les plus abondantes, les plus évoluées et comptent parmi les plus efficaces. Elles arrivent à augmenter jusqu'à 10¹⁷ fois la vitesse de la réaction de coupure de la très stable liaison glycosidique C-O (347 kJ.mol⁻¹) [6].

Le diabète sucré est un groupe de maladies métaboliques, d'étiologie diverse, caractérisées par une hyperglycémie chronique, accompagnées d'une perturbation des métabolismes glucidique, lipidique et protéique, résultant d'un défaut de la sécrétion de l'insuline ou de l'action de l'insuline ou de ces deux anomalies associées. Cette hyperglycémie provoque des complications métaboliques aiguës et chroniques dégénératives [7], [8].

La modélisation moléculaire est une application des méthodes théoriques et des méthodes de calcul pour résoudre des problèmes impliquant la structure moléculaire et la réactivité chimique [9]. Ces méthodes peuvent être relativement simples et utilisables rapidement ou au contraire elles peuvent être extrêmement complexes et demander des centaines d'heures de temps d'ordinateur, même sur un super-ordinateur. En plus, ces méthodes utilisent souvent des moyens infographiques très sophistiqués qui facilitent grandement la transformation de quantités impressionnantes de nombres en quelques représentations graphiques facilement interprétables [10].

L'expression "mécanique moléculaire" désigne actuellement une méthode de calcul largement utilisée qui permet, a priori, d'obtenir des résultats de géométries et d'énergies moléculaires en se basant sur la mécanique classique.

L'idée générale de cette méthode a été développée depuis 1930 par Andrews, où les forces régissant les interactions entre atomes liés et non liés ont une forme simple, facilement calculable. Ensuite, Westheimer en 1956 [11] grâce au développement des ordinateurs accomplit avec succès les premiers calculs par mécanique moléculaire, puis d'importants développements ont été apportés entre autres par Hendrikson, Schleyer et Allinger [12].

Le présent travail comporte trois chapitres :

- ✓ Chapitre 1 : Ce manuscrit sera consacré à une mise au point sur le noyau : Revue bibliographique sur la synthèse, la réactivité et l'activité biologique des xanthone et leurs applications médicales (la maladie de diabète).
- ✓ Chapitre 2 : Etude théorique des méthodes utilisées dans la modélisation moléculaire et les méthodes de sélection des candidats médicamenteux.
- ✓ Chapitre 3 : Étude qualitative de la relation structure-activité d'une série des dérivés de xanthone et l'application des méthodes (MM, PM3, DFT)

Références bibliographiques

- [1] Allene R Jeanes, A.; Pittsley, J. E.; Senti, F. R. Polysaccharide B-1459: A New Hydrocolloid Polyelectrolyte Produced from Glucose by Bacterial Fermentation. *J. Appl. Polym. Sci.* 1961, 5 (17), 519–526.
- [2] Rosalam, S.; England, R. Review of Xanthan Gum Production from Unmodified Starches by *Xanthomonas Comprestris* Sp. *Enzyme Microb. Technol.* 2006, 39 (2), 197–207.
- [3] Casas, J. A.; García-Ochoa, F. Viscosity of Solutions of Xanthan/Locust Bean Gum Mixtures. *J. Sci. Food Agric.* 1999, 79 (1), 25–31.
- [4] Secouard, S.; Grisel, M.; Malhiac, C. Flavour Release Study as a Way to Explain Xanthan–Galactomannan Interactions. *Food Hydrocoll.* 2007, 21 (8), 1237–1244.
- [5] L. EL BLIDI, “Synthèse chimio-enzymatique et évaluation d’inhibiteurs potentiels de glycosidases, analogues de la valioline” Thèse de doctorat, Université BLAISE PASCAL, France, 2006.
- [6] C. OUAIRY, « Synthèse de nouveaux glycoimidazoles, inhibiteurs potentiels de glycosidases préparation et utilisation en synthèse de glutaconaldéhydes oxydés en position 2 et de n-acylaminopentadiénals » Thèse de doctorat, Université Paris-sud, France, 2011.
- [7] Rodier M. (2001). Définition et classification du diabète. *Médecine Nucléaire – Imagerie fonctionnelle et métabolique* ; 25 (2) : 5-18.
- [8] (American Diabetes Association). (2008). *Diagnosis and Classification of Diabetes Mellitus*. *Diabetes Care*; 31(1): S55-S60.
- [9] D. Liotta, *Advances in Molecular Modeling*, 1, JAI Press, Greenwich, 1988.
- [10] C.S. Tsai, *An Introduction to Computational Biochemistry*, Wiley-Liss, New York, 2002.
- [11] F.H. Westheimer. “Steric Effects in Organic Chemistry”, Newman, M. S., Wiley. J.E.D, New York (1956).
- [12] N.L.Allinger. Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. *Am. J Chem. Soc.*, 99, 8127-8134 (1977).

Chapitre I

Les propriétés des dérivés xanthone et leurs applications médicales (la maladie de diabète).

1.1. Introduction

Le diabète est l'une des plus anciennes maladies connues de l'humanité dont l'effet dévastateur augmente de jour en jour et gravement à un niveau épidémique [1]. C'est une maladie du métabolisme désordonné des glucides, qui affecte également les protéines et les graisses qui sont causées par l'insuffisance totale ou relative de l'action de l'insuline [2].

Plusieurs données expliquent que les principaux facteurs de risques responsables de la propagation du diabète sucré, sont l'urbanisation, le changement de mode de vie, l'inactivité physique, l'obésité, et le stress psychosociologique qui menace les individus jour après jour [3].

Le diabète connaît une expansion très significative, selon les dernières estimations de l'OMS et de l'IDF (2016), 422 millions de personnes dans le monde sont diabétiques, ce chiffre atteindrait 552 millions d'ici 2030, soit une personne sur dix et la moitié des personnes atteintes de diabète ne sont pas diagnostiquées. Actuellement, Il y a plus de 199 millions de femmes diabétiques dans le monde soit 47% et ce total devrait augmenter à 313 millions. Le diabète se caractérise aussi par la gravité des complications qu'il engendre. En effet, il est la première cause d'insuffisance rénale terminale, de cécité et d'amputation des membres inférieurs, il est aussi la sixième cause de mortalité [4].

Le domaine de la recherche sur les plantes médicinales a gagné une importance considérable au cours des dernières décennies et la demande d'utiliser des produits naturels dans le traitement du diabète augmente dans le monde entier. La littérature disponible montre qu'il y a plus de 400 espèces végétales montrant l'activité antidiabétique [5]. Bien que certaines de ces plantes aient une grande réputation dans la médecine traditionnelle, beaucoup restent à démontrer scientifiquement [6]. Une étude scientifique des remèdes traditionnels à base de plantes pour le diabète peut fournir de précieuses pistes pour le développement de médicaments et de stratégies alternatives [7].

Les alpha-glucosidases sont des enzymes qui procèdent à la digestion des sucres. Dans un premier temps dégradés en disaccharides par l'amylase salivaire puis par l'amylase pancréatique, les sucres sont ensuite pris en charge par les alpha-glucosidases que sont la maltase, la lactase, la saccharase ou l'invertase. Ils sont alors dégradés en monosaccharides (eux seuls capables de franchir la barrière intestinale). Ainsi, les inhibiteurs des alpha-glucosidases bloquent le dernier stade de la digestion des sucres qui ne pourront pas être absorbés. Ils sont donc indiqués dans le traitement de certains diabètes [8].

Les propriétés du xanthone suggèrent des applications intéressantes en tant qu'agent de contrôle de la viscosité dans les systèmes aqueux ou en tant que

stabilisant pour les émulsions et les suspensions. De plus, issu de ressources renouvelables, il présente l'avantage d'être biodégradable et biocompatible. Ainsi, ses domaines d'application couvrent un large spectre. L'une de ses applications majeures est l'industrie alimentaire. Utilisé en faible quantité, le xanthone permet de contrôler la texture, la viscosité, l'aspect et les propriétés rhéologiques du produit final. Il est utilisé pour les assaisonnements de salade, les sauces, les boissons ou encore les plats surgelés [9] [10]. Le xanthone trouve également des applications en cosmétique (crème, dentifrice...), en agroalimentaire (amélioration de la capacité de pulvérisation des produits phytosanitaires et de leur adhérence [11]), pour les produits détergents (tolérance élevée vis-à-vis des acides et des produits alcalins) ou encore à la récupération assistée du pétrole [12].

I.2. Alpha-glucosidase dans le traitement du diabète

I.2.1. Le diabète

I.2.1.1. Historique

C'est 4000 ans avant J.C. en Chine que le diabète a été mentionné pour la première fois. On parlait alors d'urine sucrée ou d'urine de miel.

En 1500 avant J.C. un papyrus égyptien, appelé papyrus Ebers, décrit des symptômes similaires à ceux du diabète : soif intense et amaigrissement.

Au fil des siècles et des avancées de la recherche médicale, des expériences ont été réalisées pour comprendre et trouver des traitements au diabète. Voici les étapes clés de la découverte de cette maladie et des progrès réalisés par la recherche.

En parallèle les premières pompes à insuline sont commercialisées.

Depuis, la recherche continue ses avancées technologiques dont l'objectif est toujours d'améliorer le quotidien du patient diabétique et de guérir la maladie [13].

I.2.1.2. Présentation générale des différents types de diabète

Le diabète est une pathologie chronique qui se caractérise par une hyperglycémie, c'est-à-dire un excès de sucre dans le sang, dont le taux à jeun est égal ou supérieur à 1,26 g/L (7 mmol/L) de sang.

Pour diagnostiquer un diabète, une prise de sang réalisée à jeun (depuis au moins huit heures) est nécessaire. Il s'agit de mesurer la glycémie. On pourra poser un diagnostic si cette glycémie à jeun est égale ou supérieure à 1,26 g/L est constatée lors de deux mesures successives.

- Il existe deux principaux types de diabète [14] :

I.2.1.2.1. Le diabète de type 1

Il représente 10% environ de tous les cas de diabète et se déclare généralement à l'enfance suite à une destruction auto-immune des cellules insulino-sécrétrices dites cellules β des îlots de Langerhans pancréatiques. Survenant sur un terrain de

susceptibilité génétique à la cause de facteurs déclenchants (environnementaux), cette destruction est provoquée par les lymphocytes T (CD4+ et CD8+), aboutissant habituellement à une carence absolue en insuline « insulino-pénie ». L'insulinothérapie est importante chez les patients atteints de diabète type 1 [15].

I.2.1.2.2. Le diabète de type 2

C'est la forme de diabète la plus répandue représentant près de 90% des cas diagnostiqués. Il survient le plus souvent chez l'adulte. Le diabète de type 2 est caractérisé par une altération de l'insulinosécrétion et des anomalies des effets de l'insuline sur ses tissus cibles (insulinosensibilité). Il résulte de la conjonction de plusieurs gènes de susceptibilité, dont l'expression dépend de facteurs environnementaux comme la consommation excessive de graisses saturées, de sucres rapides, la sédentarité et l'obésité. L'histoire familiale et l'appartenance à un groupe ethnique à risque sont des facteurs de risques majeurs de développer un diabète de type 2 [16].

Des examens de différents types vont être réalisés [17] :

•Examens cliniques

- Le poids
- Le calcul de l'indice de masse corporelle (IMC)
- La mesure de la tension artérielle
- L'examen des vaisseaux sanguins
- L'examen neurologique

•Examens biologiques

- Le dosage de l'hémoglobine glyquée (reflet du taux de glycémie dans le sang au cours des mois précédents)
- Le taux de cholestérol sanguin
- Exploration de la fonction rénale.

•Des examens complémentaires

- Un électrocardiogramme
- Un examen ophtalmologique : le fond d'œil.

Les principales complications du diabète de type 2 qui sont aussi liées à l'obésité (athérosclérose, hypertriglycéridémie, hypercholestérolémie) et à des pathologies sous-jacentes sont :

•Les macroangiopathies [18]

- Accident vasculaire cérébral par atteinte des artères cérébrales.
- Formation d'un caillot au niveau des artères coronaires pouvant entraîner un angor ou encore un infarctus du myocarde.
- Artérite au niveau des membres inférieurs pouvant induire une gangrène et donc une amputation (notamment au niveau des pieds).

•Les microangiopathies [19]

- La rétinopathie pouvant engendrer une cécité (le diabète est la première cause de cécité en France) par atteinte des capillaires sanguins de la rétine.
- La néphropathie : insuffisance rénale par atteinte des vaisseaux irrigants les glomérules du rein, ce qui va diminuer progressivement la capacité du rein à filtrer (peut devenir terminale et nécessiter des dialyses).
- La neuropathie : la forme la plus fréquente étant la polyneuropathie distale et symétrique (atteinte des nerfs des membres inférieurs notamment les pieds) à l'origine de troubles sensitifs.

I.2.1.2.3. Les traitements actuels du diabète

Le diabète de type 1 étant d'origine génétique, il n'est pas possible de le prévenir. A l'inverse, le diabète de type 2 qui est majoritairement favorisé par le surpoids dépend beaucoup de l'hygiène de vie du malade et pourrait dans de nombreuses situations être évité [20].

a- Les traitements du diabète de type 1

résulte d'un déficit de sécrétion d'insuline par le pancréas. Le traitement repose donc sur des injections sous-cutanées d'insuline, pluriquotidiennes, pour compenser ce défaut de production endogène.

C'est un traitement lourd qui demande de la rigueur dans son suivi, c'est pourquoi il nécessite une éducation thérapeutique pour le patient. Ce dernier doit apprendre à mesurer son taux de glycémie afin d'adapter les doses d'insuline à s'injecter et limiter le risque d'hypoglycémie [14].

b- Les traitements du diabète de type 2

Avant tout traitement pour un diabète de type 2, la première chose à mettre en place c'est des mesures hygiéno-diététiques, qui peuvent dans certains cas éviter la prise d'un traitement aux patients ou dans les autres cas, en optimiser l'efficacité.

Par mesures hygiéno-diététiques on entend sur le plan alimentaire : une alimentation équilibrée, une réduction de l'apport calorique, une meilleure répartition des prises alimentaires ainsi que l'augmentation de la consommation de fibres, auxquelles on va associer une activité physique régulière et adaptée (au moins trois fois 45 min par semaine) ainsi que l'arrêt de la consommation de tabac et d'alcool. Ces mesures non exhaustives seront accompagnées d'une surveillance de l'équilibre glycémique [21].

L'équilibre glycémique cible correspond au taux sanguin d'hémoglobine glyquée (HbA1C) à atteindre. Il sera adapté au patient par le médecin, en fonction de son profil et pourra évoluer au cours du temps. Cependant, pour la plupart des diabétiques de type 2, l'objectif est un taux d'hémoglobine glyquée inférieur ou égal à 7% [22]. Si les mesures hygiéno-diététiques ne suffisent pas ou plus pour atteindre « l'objectif glycémique cible », le médecin, en concertation avec son patient,

pourra alors prescrire un traitement médicamenteux.

Ce type de traitement sera toujours débuté aux doses minimales afin d'en favoriser la tolérance. La dose sera progressivement augmentée en fonction des besoins du patient jusqu'à la dose maximale tolérée ou jusqu'à l'atteinte de l'objectif.

I.3. Caractérisation des α -glucosidases humaines et de leur mécanisme d'action

I.3.1. L'amidon

L'amidon est, après la cellulose, la principale substance glucidique synthétisée par les végétaux supérieurs à partir de l'énergie solaire. Il constitue une source énergétique indispensable à l'alimentation des êtres vivants et de l'homme en particulier. Les sources d'amidon les plus importantes sont représentées par les céréales, les tubercules et les légumineuses. Certains fruits peuvent également être riches en amidon.

L'étude de l'amidon peut être envisagée en le considérant soit comme une entité physique caractérisée par une diversité de forme et de taille [23], soit comme une entité chimique composée principalement de polymères de glucose ayant une structure cristalline typique [24] et présentant des comportements particuliers en fonction des conditions hydro-thermiques auxquelles ils ont été soumis et des interactions qu'ils peuvent établir avec d'autres constituants [25], soit encore comme entité de caractérisation des espèces en botanique [26].

À l'état natif, l'amidon est insoluble dans l'eau froide et est constitué de granules dont la taille, la composition et les propriétés physico-chimiques et fonctionnelles dépendent de l'origine botanique [27] [28] [29] et des conditions de culture [30] [31]. Ces propriétés peuvent cependant évoluer en fonction des traitements auxquels les granules sont soumis.

I.3.1.1. Composition des granules d'amidon de maïs

Les granules d'amidon sont composés d'une proportion d'amylose et d'amylopectine variable selon la variété de maïs. [28] rapportent des teneurs en amylose variant entre 16,1 et 23,3 % après analyse de 35 variétés différentes de maïs ; [32] ont quant à eux observé des teneurs en amylose variant entre 15,3 et 25,1 %. Les amidons de certaines variétés sont essentiellement composés d'amylopectine, tandis que d'autres sont beaucoup plus riches en amylose [32].

I.3.1.2. Organisation structurale des granules de l'amidon

Les granules d'amidon sont des entités semi-cristallines organisées sur une échelle à quatre niveaux : l'échelle moléculaire ($\sim \text{Å}$), l'échelle lamellaire ($\sim 90\text{Å}$), l'échelle dite de croissance radiale ($\sim 0,1 \mu\text{m}$) et l'échelle granulaire [34].

La structure granulaire est assurée par des liaisons glucosidiques qui forment les homopolymères (amylose et amylopectine) à la base de la formation des granules, ainsi que par des interactions de Van der Waals et des ponts hydrogène qui stabilisent l'organisation des polymères d'hydrates de carbone en doubles hélices (échelle moléculaire). L'empilement des doubles hélices en structures cristallines (échelle lamellaire) et la succession des phases amorphes et cristallines (échelle de croissance radiale) constituent le granule d'amidon [35] [36].

À l'état natif, le granule d'amidon comporte des zones cristallines et des zones amorphes disposées sous forme de structures lamellaires concentriques [37]. Les parties cristallines, dispersées dans une phase amorphe discontinue, sont constituées principalement de chaînes d'amylopectines organisées en doubles hélices, groupées densément (clusters), parallèles les unes aux autres, alors que l'amylose et les points de ramification des liaisons α -D-glycopyranoses (1-6) des amylopectines sont principalement localisés dans les parties amorphes du granule [36] [38].

I.3.1.3. Réactivité des amidons et mode d'action des amylases

I.3.1.3.1. Amylases

Les enzymes qui dégradent l'amidon ont été classées en quatre classes selon leur mode d'action [39] : les endoamylases, les exoamylases, les enzymes de débranchement et les enzymes qui servent à transférer un groupement d'atomes.

a) Endoamylases

Les endoamylases, connues aussi par le terme « enzymes liquéfiantes » sont des enzymes qui coupent les liaisons α -(1,4) glycosidiques présentes dans les parties internes des chaînes de l'amylose et de l'amylopectine. Les produits d'hydrolyse, qui sont moins visqueux, sont des dextrans et des oligosaccharides avec des chaînes variées ayant une configuration 36 α en C1 des unités réductrices du glucose d'où le nom « α -amylase » [40] [41].

b) Exoamylases

Les exoamylases coupent toutes les liaisons α -1,4 glycosidiques comme la β -amylase (EC.3.2.1.2), ou les liaisons α -(1,4) et α -(1,6) glycosidiques comme l'amyloglucosidase (la glucoamylase) (EC.3.2.1.3) et l' α -glucosidase (EC.3.2.1.20). Les exoamylases agissent sur les unités de glucoses externes de l'amylose et de l'amylopectine et produisent uniquement du glucose (glucoamylase et α -glucosidase), ou du maltose et des β -dextrans limités (β -amylase). La glucoamylase et l' α -glucosidase diffèrent selon leur préférence vis-à-vis des substrats ;

- ✓ La glucoamylase et l' α -glucosidase diffèrent selon leur préférence vis-à-vis des substrats ; l' α -glucosidase agit mieux sur les petits maltooligosaccharides et libère le glucose avec une configuration α .
- ✓ la glucoamylase hydrolyse mieux les chaînes longues de polysaccharide.

Les β -amylases se trouvent en abondance chez les végétaux (blé et l'orge) où elles sont constitutives alors que l' α -amylase est induite. Elles sont aussi produites à partir d'une variété de microorganismes [42]. La β -amylase se distingue de celle des endoamylases par une réduction beaucoup plus lente de la viscosité du substrat [41]. La glucoamylase est présente largement dans les microorganismes. Les glucoamylase du genre *Aspergillus* et *Rhizopus* sont disponibles dans le commerce sous forme de préparations enzymatiques industrielles.

La glucoamylase est présente largement dans les microorganismes. Les glucoamylase du genre *Aspergillus* et *Rhizopus* sont disponibles dans le commerce sous forme de préparations enzymatiques industrielles.

c) Enzymes de débranchement

Ces enzymes catalysent les réactions d'hydrolyse des liaisons glycosidiques α -1,6 qui participent à un débranchement des chaînes de glucose de l'amidon. Parmi les enzymes de cette classe, on peut citer l'isoamylase (EC.3.2.1.68) et la pullulanase (EC.3.2.2.48).

La première est spécifique aux liaisons de l'amylopectine alors que la deuxième est spécifique aux liaisons de l'amylopectine et du pullulane.

d) Enzymes du quatrième groupe

Ces enzymes catalysent la rupture d'une liaison α -(1,4) dans une molécule donneuse et le transfert d'une partie de la molécule à une autre molécule réceptrice capable de former une nouvelle liaison glycosidique de nature α -(1,4) ou α -(1,6). Parmi ces enzymes, on peut citer l'amyloamylase (EC.2.4.1.25) et la cyclodextrine glucotransférase (EC.2.4.1.19).

L'utilisation de l' α et de la β -amylase permet d'obtenir des fractions riches en maltose mais renfermant aussi des dextrans limités dues à la présence des ramifications. Pour l'industriel, il peut s'agir de freintes importantes et nécessite, donc, de dépolymériser complètement l'amidon. Pour obtenir une action maximale des enzymes, on associe généralement des enzymes de débranchement (pullulanase et isoamylase) permettant d'obtenir uniquement des enchaînements linéaires

I.3.1.3.2. Activité des amylases

L'activité enzymatique exprime l'efficacité de l'hydrolyse dans les conditions données. Elle peut être évaluée par plusieurs méthodes, telles que l'évolution de la

viscosité, de la turbidité, de l'amidon résiduel (méthode colorimétrique à l'iode) et des sucres réducteurs formés (DNS, Somogyi-Nelson) [43].

Dans les céréales, les α -amylases constituent 30 % des protéines totales synthétisées durant la germination. [44] ont constaté que les amylases du sorgho présentent la plus faible activité comparée à celles de l'éleusine et du maïs. Néanmoins, elles sont plus résistantes à l'inactivation thermique (50 °C). Cette résistance ainsi que le prix élevé des enzymes exogènes pour le brassage, encouragent les pays comme le Nigeria à orienter leur production en utilisant le malt de sorgho plutôt que le malt d'orge [44].

Les études faites sur le maltage des différentes céréales ont montré que le blé présente l'activité enzymatique la plus élevée. Elle est atteinte après 6 à 7 jours de germination. Pour le mil, par contre, l'activité maximale est obtenue après 72 heures. [45] (2000) rapportent que l'activité de la glucoamylase et de la pullulanase est 20 à 600 fois plus faible que celle de l' α -amylase.

I.3.1.3.3. Procédés d'hydrolyse enzymatique des amidons

Pour l'obtention de solutions de glucose et/ou maltose, le procédé de production de maltodextrines est suivi d'un traitement enzymatique par addition de glucosidase qui hydrolyse les oligosaccharides en monosaccharides (glucose). Les maltodextrines peuvent être convertis en sirop de maltose. Cette étape nécessite une catalyse enzymatique par la β -amylase. Un système à deux enzymes est utilisé dans le cas de la conversion des maltodextrines en sirop de glucose/maltose.

La composition du produit d'hydrolyse dépend de la quantité d'enzyme utilisée, du temps de réaction et du dextrose équivalent des maltodextrines [46]. La modélisation de la cinétique enzymatique avec des rappels sur les constantes cinétiques sont donnés ci-dessous.

I.3.1.3.4. Mode d'action des amylases

L'action catalytique des amylases dépend de la nature du substrat, des conditions opératoires ainsi que de la source et de la concentration de l'enzyme. La composition en acides aminés, leur disposition dans la chaîne peptidique ainsi que la structure secondaire et tertiaire de la protéine régissent le mécanisme catalytique.

L'action de l' α -amylase donnant lieu à la formation des oligosaccharides et des α -dextrines limites est schématisée à la figure I.1, [39]

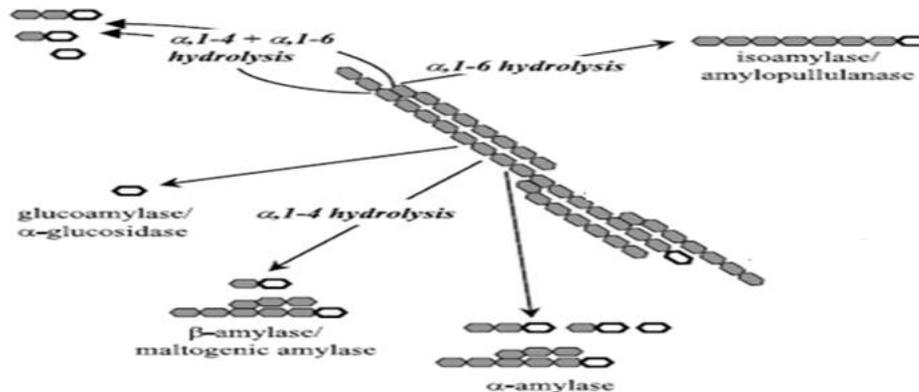


Figure I .1: Schéma de dégradation de l'amidon par les différentes enzymes

Le cycle ouvert symbolise l'extrémité réductrice de la molécule polyglucose.

Toutes les enzymes qui dégradent l'amidon possèdent des séquences d'acides aminés similaires : la famille des α -amylases est la 13^{ème} famille dans la classification de Henrissat [47] basée sur les séquences d'acides aminés. Ce groupe comprend les enzymes qui possèdent les propriétés suivantes :

- Elles agissent sur les liaisons α -glycosidiques α -(1,4) ou α -(1,6) ou les deux et hydrolysent ces liaisons pour produire des mono et/ ou oligosaccharides sous la forme α -anomères.
- Elles possèdent la structure $(\beta/\alpha)8$ ou TIM barrel (figure I .2) contenant le site catalytique.
- Elles ont quatre régions conservées dans leur structure primaire [41] qui contiennent les acides aminés qui forment le site catalytique, et d'autres acides aminés qui sont essentiels à la stabilité de la topologie de TIM Barrel.



Figure I .2: Structure $(\beta/\alpha)8$ (A) ou TIM barrel (B) contenant le site catalytique.

I.4. Le xanthone

I.4.1. Historique

La production et la commercialisation de polysaccharides microbiens à grande échelle a commencé dans les années 60 (Tableau I .1). Sous la dénomination de Kelzan® (Kelco Co. Inc., USA), le xanthone (ou gomme xanthone) fut le deuxième polysaccharide d'origine microbienne à être commercialisé après le dextrane [48] .Il s'agit d'un hétéropolysaccharide anionique ramifié, produit par une bactérie à Gram (-), Xanthomonas sp.

La gomme xanthone est répertoriée dans la liste des émulsifiants/stabilisants de la Communauté Economique Européenne sous le symbole E-415 [48]. Il est aujourd'hui le polysaccharide le plus vendu. Le marché mondial, estimé à plus de 30 000 tonnes par an, connaît une croissance soutenue de plus de 5% par an.

Polysaccharides	Souches productrices	Producteurs
Alginate	Azotobacter vinelandii Pseudomonas aeruginosa	Tate & Lyle
Cellulose	Acetobacter xylinum	Ajinomoto, Hayashibara, Kelco
Curdlan	Agrobacterium sp Alcaligenes faecalis	Takeda
Dextrane	Leuconostoc mesenteroides Leuconostoc dextranicum	Bayer, Henkel
Gellane	Sphingomonas paucimobilis	Kelco
Méthylane	Methylophilussp	Celgene, British Petroleum, Mitsubishi Kasei
Pullulane	Aureobasidium pullulans Azotobacterpullulans	Ajinomoto, British Petroleum, Mitsubishi Kasei, Hayashibara
Rhamsane	Alcaligenes sp	Kelco
Scléroglycane	Sclerotium glucanicum	SKW Biosystem, British Petroleum
Succinoglycane	Agrobacterium radiobacter Agrobacterium tumefaciens Erwinia herbicola	Aqualon, Kelco, Rhodia, Takeda
Wellane	Alcaligenes sp	Kelco
Xanthane	Xanthomonas campestris	Kelco, Rhodia, ADM, Jungbunzlauer, SKW Trotsberg,

Tableau I .1: Exemples de polysaccharides d'origine microbienne.

I.4.2. Structure du xanthone

Le xanthone est un hétéropolysaccharide anionique hydrophile. Comme la plupart des polysaccharides naturels, le xanthane possède une masse molaire élevée (≈ 106 g/mol).

La structure chimique du xanthone est illustrée en Figure I.3. Ce polymère est constitué d'un squelette cellulosique linéaire de β -1,4-D-glucose dont une unité sur deux est substituée en C3 par une chaîne triosidique comportant un α -D-mannose, un acide β D-glucuronique et un β -D-mannose terminal [49].

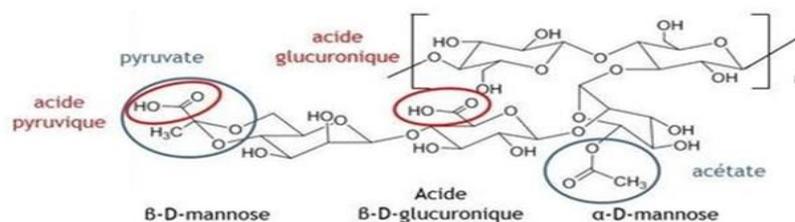


Figure I.3 : Structure chimique du xanthone

Le xanthone possède également des groupements acétate et pyruvate sur le carbone 6 du mannose interne et sur les carbones 4 et 6 du mannose terminal, respectivement, et dont les proportions varient en fonction de la nature de la souche bactérienne [50], des conditions de fermentation [51] [49][52] ou encore des traitements de post-fermentation [53]. De plus, la proportion de ces deux groupements a un impact sur les propriétés du polymère en solution.

Plus récemment, des travaux menés par Kool et al. [54] [55] ont abouti à l'identification de motifs de substitutions supplémentaires. Pour ces auteurs, la seule présence de l'acétate sur le α -D-mannose ne permet pas une stabilisation suffisante de la structure hélicoïdale. Ainsi, le mannose terminal peut également être substitué par un groupement acétate (voir la Figure I.4). Ces résultats ont été confirmés par Abbaszadeh et al.[56]

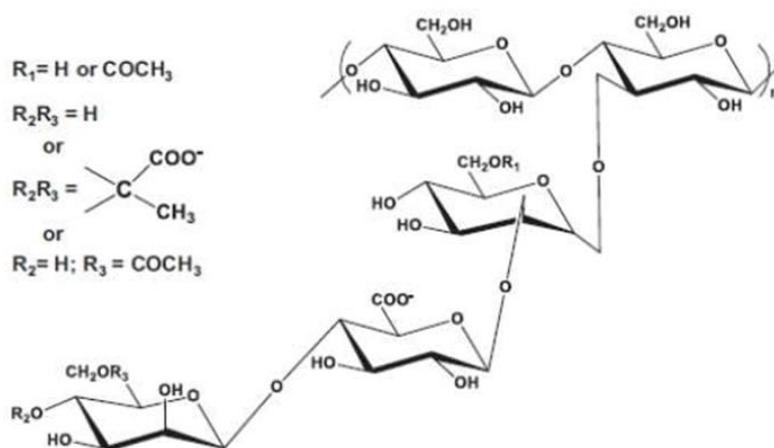


Figure I.4 : Structure chimique du xanthone selon Kool et al.21

Le xanthone doit son caractère anionique à la présence de deux acides carboxyliques le long de la chaîne latérale, à savoir un acide glucuronique et un acide pyruvique, situé sur le mannose terminal.

Leur proportion par unité de répétition, notamment pour l'acide pyruvique, est dépendante du degré de substitution en pyruvate. Le xanthone possède donc des capacités d'échange de protons en solution [57][58].

Dans les produits commerciaux, ces fonctions sont souvent sous forme carboxylate, et saponifiées par des ions Na⁺, K⁺ ou Ca²⁺. La forme neutralisée peut être ensuite convertie en forme acide par échange d'ions ou bien par dialyse contre une solution acide.

I.4.3. Applications

Au delà des caractéristiques économiques (prix de vente, qualité reproductible, etc.), les propriétés chimiques et physiques du xanthone en font un produit extrêmement utilisé, dont la principale caractéristique est de modifier les propriétés rhéologiques des solutions aqueuses [59] :

- bonne dissolution dans l'eau chaude et l'eau froide,
- viscosité élevée à faible cisaillement, même à faible concentration,
- viscosité constante et indépendante du pH (pour des pH compris entre 1 et 11) et de la température (jusqu'à 120°C),
- pseudoplasticité et thixotropie,
- récupération instantanée de la viscosité après l'application d'un cisaillement,
- compatibilité avec un grand nombre de substances,
- synergie avec des polysaccharides végétaux comme les galactomannanes de guar.
- Les industries chimiques, pétrochimiques, alimentaires et pharmaceutiques sont les principaux utilisateurs [60] (Tableau I.2).

Les qualités exceptionnelles, décrites ci-dessus, rendent le xanthane irremplaçable pour certaines applications alimentaires, sans dose limite. Il favorise la libération rapide des arômes, offre une bonne impression gustative, est neutre en goût, est non digestible, et est compatible avec d'autres ingrédients (protéines, lipides et polysaccharides) [61].

Dès son introduction sur le marché commercial, la gomme xanthone a présenté un vif intérêt dans des domaines aussi variés que l'alimentaire, la chimie et la récupération assistée du pétrole [62]. Mais, c'est dans ce dernier secteur que le potentiel du xanthone a été le plus exploité dans les premiers temps, correspondant aux périodes de crises pétrolières. Désormais, le domaine alimentaire connaît un essor important. En effet, environ 60% du xanthane produit est de qualité alimentaire [63],

et le marché continue de progresser, notamment aux Etats-Unis, où les habitudes alimentaires entraînent le développement de la gomme xanthone. De même, le marché de la cosmétique contribue fortement à l'essor de la gomme. En particulier, les gels douches, les crèmes et les pâtes dentifrices garantissent l'obtention de produits dont la texture satisfait aux exigences du consommateur.

Propriétés et domaine d'application		Exemples d'application
Agent stabilisant	Alimentaire Chimie	<ul style="list-style-type: none"> ➤ Sauces salades ➤ Desserts multiphasiques ou multicouches ➤ Boissons aux fruits ou chocolatées ➤ Potages ➤ Plats cuisinés ➤ Crèmes glacées ➤ Peintures ➤ Vernis ➤ Savons ➤ Teinture de textile ➤ Agrochimie (herbicides, fertilisants)
Agent viscosant	Alimentaire Pétrole et BTP Cosmétique Chimie	<ul style="list-style-type: none"> ➤ Préparation et acheminement de produits alimentaires ➤ (Potages, sauces salades, crèmes glacées, etc.) ➤ Substitution des graisses texturantes dans les aliments ➤ basses calories ➤ Récupération assistée du pétrole ➤ Mélange de ciments et mortiers ➤ Pâtes dentifrice ➤ Agents lavants (lave-glaces, gels WC)

Tableau I.2: Applications alimentaires et non-alimentaires du xanthane.

Références bibliographiques

- [1] Wild, S., Roglic, G., Green, A., Sicree, R., Hilary, K., 2004. Global Prevalence of Diabetes: Estimates for the year 2000 and projection for 2030. *Diabetes Care* 27, 1047– 1053. doi:10.2337/diacare.27.5.1047
- [2] Rodier, M., 2001. M . Rodier Définition et classification du diabète . Définition et classification du diabète. *Médecine Nucléaire - Imag. Fonct. métabolique* 25, 91–93.
- [3] Whiting, D., Guariguata, L., Weil, C., Shaw, J., 2011. IDF diabetes atlas: global estimates of the prevalence of diabetes for 2011 and 2030. *Diabetes Res Clin Pr.* 94, 311–21.
- [4] OMS, 2016. RAPPORT MONDIAL sur le diabete
- [5] Grover, J.K., Yadav, S., Vats, V., 2002. Medicinal plants of India with anti-diabetic potential. *J. Ethnopharmacol.* 81, 81–100. doi:10.1016/S0378-8741(02)00059-4
- [6] Kolling, M., Winkley, K., von Deden, M., 2010. “For someone who’s rich, it’s not a problem”. Insights from Tanzania on diabetes health-seeking and medical pluralism among Dar es Salaam’s urban poor. *Global. Health* 6, 8. doi:10.1186/1744-8603-6-8
- [7] Pan, S.Y., Zhou, S.F., Gao, S.H., Yu, Z.L., Zhang, S.F., Tang, M.K., Sun, J.N., Ma, D.L., Han, Y.F., Fong, W.F., Ko, K.M., 2013. New perspectives on how to discover drugs from herbal medicines: CAM’S outstanding contribution to modern therapeutics. *Evidencebased Complement. Altern. Med.* 2013, 25. doi:10.1155/2013/627375
- [8] <https://sante-medecine.journaldesfemmes.fr/faq/53136-alpha-glucosidase-definition>
- [9] Rosalam, S.; England, R. Review of Xanthan Gum Production from Unmodified Starches by *Xanthomonas Comprestris* Sp. *Enzyme Microb. Technol.* 2006, 39 (2), 197–207.
- [10] Sharma, B. R.; Naresh, L.; Dhuldhoya, N. C.; Merchant, S. U.; Merchant, U. C. Xanthan Gum - ABoon to Food Industry. *Food Promot. Chron.* 2006, 1(5), 27–30.
- [11] Palaniraj, A.; Jayaraman, V. Production, Recovery and Applications of Xanthan Gum by *Xanthomonas Campestris*. *J. Food Eng.* 2011, 106 (1), 1–12. [12] Jang, H. Y.; Zhang, K.; Chon, B. H.; Choi, H. J. Enhanced Oil Recovery Performance and Viscosity Characteristics of Polysaccharide Xanthan Gum Solution. *J. Ind. Eng. Chem.* 2015, 21, 741–745.
- [13] <https://www.dinnosante.fr/fre/17/histoire-du-diabete>

- [14] Le diabète. [ANSM] 2015. <http://ansm.sante.fr/Dossiers/Diabete/Le-diabete/%28offset%29/0>.
- [15] 121.Grimaldi A., (1999). Diabétologie. Questions d'internat. CHU-PS : 15-23 ; 99-129.
- [16] Haffner, S. M. (1998). Epidemiology of type 2 diabetes: risk factors. Diabetes care, 21(Supplement 3), C3-C6.
- [17] DELPECH Romain, Thèse de doctorat , université Biskra ,2015.
- [18] Pr. Altman, J.J. Diabète HEGP. [Assistance Publique Hôpitaux de Paris] juin 2015. disponible sur <http://www.hegp.fr/diabeto/index.html>.
- [19] Pr. Boitard, C. Diabète de type 1 (DID). [Inserm] avril 2014. disponible sur <http://www.inserm.fr/thematiques/physiopathologie-metabolisme-nutrition/dossiers-dinformation/diabete-de-type-1-did>.
- [20] Le diabète. [Institut Pasteur de Lille] consulté en juillet 2015. disponible sur <http://www.pasteurlille.fr/fr/etudier-et-comprendre/nos-avancees/diabete/>.
- [21] Traitement du diabète de type 2. [Faculté de médecine de Toulouse] consulté en juillet 2015. disponible sur http://www.medecine.upstlse.fr/dcem3/module14/diabetologie/Chap05_TRAITEMENT_DIABETE_TYPE%202.pdf.
- [22] Diabètes de type 2 : quand et quels médicaments prescrire pour le contrôle glycémique. [HAS] 13 février 2013. disponible sur http://www.has-sante.fr/portail/jcms/c_1359987/fr/diabete-de-type-2-quand-et-quels-medicaments-prescrire-pour-le-controle-glycemique.
- [23] Singh N. et al., 2003. Morphological, thermal and rheological properties of starches from different botanical sources. Food Chem., 81, 219-231.
- [24] Jenkins P.J. & Donald A.M., 1995. The influence of amylose on starch granule structure. Int. J. Biol. Macromol., 17(6), 315-321.
- [25] Tolstoguzov V., 2003. Thermodynamic considerations of starch functionality in foods. Carbohydr. Polym., 51, 99-111.
- [26] Reichter E.T., 1913. The differentiation and specificity of starches in relation to genera, species, etc. : the stereochemistry applied to protoplasmic processes and products, and as a strictly scientific basis for the classification of plants and animals. Part I. Washington, D.C., USA: Carnegie Institution of Washington.
- [27] Li J., Berke T.G. & Glover D.V., 1994. Variation of thermal properties of starch in tropical maize germ plasm. Cereal Chem., 71(1), 87-90.

- [28] Seetharaman K. et al., 2001. Thermal and functional characterization of starch from Argentinean corn. *Cereal Chem.*, 78, 379-386.
- [29] Sandhu K.S., Singh N. & Malhi N.S., 2005. Physicochemical and thermal properties of starches separated from corn produced from crosses of two germ pools. *Food Chem.*, 89, 541-548.
- [30] Paterson J.L., Hardacre A., Li P. & Rao M.A., 2001. Rheology and granule size distribution of corn starch dispersions from two genotypes and grown in four regions. *Food Hydrocolloids*, 15, 453-459.
- [31] Hardacre A.K. & Clark S.M., 2006. The effect of hybrid and growing environment on the rheological properties of starch and flour from maize (*Zea mays* L.) grain dried at four temperatures. *Int. J. Food Sci. Technol.*, 41, 144-150.
- [32] Sandhu K.S., Singh N. & Kaur M., 2004. Characteristics of the different corn types and their grain fractions: physicochemical, thermal, morphological, and rheological properties of starches. *J. Food Eng.*, 64, 119-127.
- [33] Rahman S. et al., 2007. Resistant starch in cereals: exploiting genetic engineering and genetic variation. *J. Cereal Sci.*, 46, 251-260.
- [34] Waigh T.A., Gidley M.J., Komanshek B.U. & Donald A.M., 2000. The phase transitions in starch during gelatinization: a liquid crystalline approach. *Carbohydr. Res.*, 328, 165-176
- [35] Van den Berg C., 1981. Vapour sorption equilibria and other water starch interactions: a physico-chemical approach. PhD thesis: Agricultural University Wageningen (The Netherlands).
- [36] Imberty A., Buléon A., Tran V. & Pérez S., 1991. Recent advances in knowledge of starch. *Starch*, 43(10), 375-384.
- [37] French D., 1984. Organisation of starch granules. In: Whistler R.L., Bemiller J.N. & Paschall E.F., eds. *Starch: chemistry and technology*. 2nd ed. New York, USA: Academic Press Inc.
- [38] Jenkins P.J. & Donald A.M., 1996. Application of small-angle neutron scattering to the study of the structure of starch granules. *Polymer*, 25, 5559-5568.
- [39] Marc, J.E.C., Maarel, V., Van der Veen, B., Uitdehaag, J.C.M., Leemhuis, H. & Dijkhuizen, L. 2002. Properties and applications of starch-converting enzymes of the α -amylase family. Review article. *Journal of Biotechnology*, 94, 137-155.
- [40] Muralikrishna, G., Nirmala, M. & Muralikrishna, G. 2005. Cereal α -amylases, an overview. *Carbohydrate Polymers*, 60, 163-173.

- [41] Kuriki, T. & Imanaka, T. 1999. The concept of α -amylase family: Structural similarity and common catalytic mechanism. *Journal of Bioscience and Bioengineering*, 87, 557- 565.
- [42] Pandey, A., Nigam, P., Soccol, C. R., Soccol, V.Y., Singh, D. & Mohan, R. 2000. Advances in microbial amylases. *Biotechnological Applied Biochemistry*, 31, 135-152.
- [43] Southgate, D.A.T. 1979. Determination of food carbohydrates. Applied science, LTD, London.178 p.
- [44] Adewale, I.O., Agumanu, E.N. & Oti-Okoronwo, F.I. 2006. Comparatives studies on α -amylases from malted maize (*Zea mays*), millet (*Eleusine corana*) and sorghum (*Sorghum bicolor*). *Carbohydrate Polymers*, 66 (1), 71-74.
- [45] Nirmala, M., Subba Rao, M.V.S.S.T. & Muralikrishna, G. 2000. Carbohydrates and their degrading enzymes from native and malted finger millet (*Ragi*, *Eleusine coracana*, Indaf-15). *Food Chemistry*, 69: 175-180.
- [46] Lurgi, AG. 2004. Starch saccharification technology. www.lurgi-lifescience.com. Consulté en Janvier 2005.
- [47] Henrissat, B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemistry Journal*, 280, 309-324.
- [48] Kang K.S. and Pettit D.J. (1993). Xanthan, gellan, wellan, and rhamosan. In: *Industrial Gums*, 3rd ed., Whistler R.L. and Bemiller J.N. Eds., USA, 341-397.
- [49] Jansson, P.; Kenne, L.; Lindberg, B. Structure of the Extracellular Polysaccharide from *Xanthomonas Campestris*. *Carbohydr. Res.* 1975, 45 (1), 275–282.
- [50] Torres, L. G.; Brito, E.; Galindo, E.; Choplin, L. Viscous Behaviour of Xanthan Aqueous Solutions from a Variant Strain of *Xanthomonas Campestris*. *J. Ferment. Bioeng.* 1993, 75 (1), 58–64.
- [51] Casas, J. A.; Santos, V. E.; Garcia-Ochoa, F. Xanthan Gum Production under Several Operationnal Conditions : Molecular Structure and Rheological Properties. *Enzyme Microb. Technol.* 2000, 26, 282–291.
- [52] Papagianni, M.; Psomas, S. K.; Batsilas, L.; Paras, S. V.; Kyriakidis, D. A. Xanthan Production by *Xanthomonas Campetris* in Batch Culture. *Process Biochem.* 2001, 37, 73–80.
- [53] Palaniraj, A.; Jayaraman, V. Production, Recovery and Applications of Xanthan Gum by *Xanthomonas Campestris*. *J. Food Eng.* 2011, 106 (1), 1–12.

- [54] Kool, M. M.; Gruppen, H.; Sworn, G.; Schols, H. A. Comparison of Xanthans by the Relative Abundance of Its Six Constituent Repeating Units. *Carbohydr. Polym.* 2013, 98 (1), 914–921.
- [55] Kool, M. M.; Gruppen, H.; Sworn, G.; Schols, H. A. The Influence of the Six Constituent Xanthan Repeating Units on the Order–Disorder Transition of Xanthan. *Carbohydr. Polym.* 2014, 104, 94–100.
- [56] Abbaszadeh, A.; Lad, M.; Janin, M.; Morris, G. A.; MacNaughtan, W.; Sworn, G.; Foster, T. J. A Novel Approach to the Determination of the Pyruvate and Acetate Distribution in Xanthan. *Food Hydrocoll.* 2015, 44, 162–171. [57] Young, S.-L.; Martino, M.; Kienzle-Sterzer, C.; Torres, J. A. Potentiometric Titration Studies on Xanthan Solutions. *J. Sci. Food Agric.* 1994, 64 (1), 121–127.
- [58] Richardson, R. K.; Ross-Murphy, S. B. Non-Linear Viscoelasticity of Polysaccharide Solutions. 2: Xanthan Polysaccharide Solutions. *Int. J. Biol. Macromol.* 1987, 9 (5), 257–264.
- [59] Sutherland I.W. and Ellwood D.C. (1979). Microbial exopolysaccharides : industrial polymers of CUITent and future potential. In: *Microbial technology: current state, future prospects*, Bull AT., Ellwood D.C. and Ratledge C. Eds, Cambridge, UK, 107-150.
- [60] Yalpani M. and Sandford P.A. (1987). Commercial polysaccharides: recent trends and developments. In: *Industrial polysaccharides: genetic engineering, structure/property relations and applications*. Yalpani M. Ed, Amsterdam, NL. 311-335.
- [61] Morris (1987). Organoleptic properties of food polysaccharides in thickened systems. In: *Industrial polysaccharides: genetic engineering, structure/property relations and applications*. Yalpani M. Ed, Amsterdam, NL. 225-238.
- [62] Linton J.D. (1991). Microbial polysaccharides. In: *Biomaterials*, Byron D. Ed, Basingstoke, U.K. 217-261.
- [63] Sutherland I.W. (1983). Extracellular polysaccharides. In : *Biotechnology : a comprehensive treatise*, Rehm R.I. and Reed G. Eds, Weinheim, NL, 531-574.

CHAPITRE II

La modélisation moléculaire

II.1. Introduction

La modélisation moléculaire peut être considéré comme un ensemble de techniques informatiques basées sur des méthodes de chimie théorique et les données expérimentales qui peuvent être utilisés soit pour analyser les molécules et les systèmes moléculaires ou à prédire les propriétés moléculaires, chimiques et biochimiques [1]. Il sert de pont entre la théorie et expérience pour:

- ✓ Extraire résultats pour un modèle particulier.
- ✓ Comparer les résultats expérimentaux du système.
- ✓ Comparer les prédictions théoriques du modèle.
- ✓ Aide comprendre et interpréter les observations expérimentales.
- ✓ Corrélation entre détails microscopiques au niveau atomique et moléculaire et les propriétés macroscopiques.
- ✓ Fournir de l'information non disponible à partir d'expériences réelles.

Grace au développement informatique de ces dernières années et à l'essor du calcul parallèle intensif en particulier, la modélisation moléculaire est devenue un véritable enjeu.

En effet les systèmes moléculaires qui sont amènes à être étudiés tendent à devenir de plus en plus complexes. Cette complexité est bien sûr liée à la taille des molécules envisagées (plusieurs centaines de milliers d'atomes pour les molécules biologiques par exemple) ainsi qu'à la structure intrinsèque des atomes eux-mêmes, mais aussi au degré de précision exigé pour le calcul de certaines grandeurs physiques [2].

La modélisation moléculaire consiste en la construction de modèles tridimensionnels à partir des données. Elle trouve sa raison d'être d'une part dans les limitations expérimentales des méthodes de détermination de structure de protéines et d'autre part dans l'incapacité actuelle de prédire la structure 3D à partir de la seule information de séquence. En outre, elle permet d'investiguer les changements de conformations liées à des mutations à partir de structures 3D expérimentales.

II.2. Représentation des méthodes de calculs

II.2. 1. Mécanique quantique

La mécanique quantique est le prolongement de la théorie des quanta, issue des travaux de Planck, de leur interprétation par Einstein et de leur application à la théorie atomique par Bohr et Sommerfeld. Elle explique la quantification de certaines grandeurs (énergie, moment cinétique) et fait émerger le principe d'exclusion de Pauli. La nouvelle conception des particules qui découle de la dualité onde-corpuscule, explicitée dans les travaux de De Broglie (1923) conduit à la mécanique ondulatoire.

Les méthodes de la mécanique quantique, qui font appel à la distribution des électrons répartis en orbitales autour de la molécule, impliquent des temps de calcul souvent élevés qui limitent leur usage à des petites molécules ou nécessitent le recours à de nombreuses approximations. Elles sont particulièrement adaptées au calcul des charges et des potentiels électrostatiques, à l'approche des mécanismes réactionnels ou à la polarisabilité. L'objectif de la mécanique quantique est principalement de déterminer l'énergie et la distribution électronique [3].

II.2. 1.1. Base de la chimie quantique

La chimie quantique définit la structure moléculaire comme un noyau autour duquel gravitent des électrons, qui sont décrits par leur probabilité de présence en un point être présentés par des orbitales. Les équations de la chimie quantique sont basées sur la résolution de l'équation de Schrödinger qui s'écrit pour les états stationnaires [4].

$$\hat{H}\Psi = E\Psi \quad (1)$$

Où E est l'énergie associée à la fonction d'onde $\Psi(r_1, r_2, \dots, r_n, t)$, obtenue en résolvant l'équation de Schrödinger indépendante du temps. \hat{H} : Hamiltonien totale d'une molécule comportant N noyaux et n électrons. Il est défini par la somme de cinq termes (terme cinétique des électrons, terme cinétique des noyaux, terme de répulsion entre électrons, terme de répulsion entre noyaux, et le terme d'attraction électrons - noyaux). Dans le cas d'un système composé de n électrons en interaction avec N noyaux et décrits par la fonction d'onde poly-électronique $\Psi = \Psi(r_1, r_2, \dots, r_n)$ d'énergie E, l'hamiltonien total du système est défini en unités atomique [4].

$$\hat{H} = \sum_{k=1}^{2n} \frac{\hbar^2}{2m_e} \nabla_k^2 - \sum_{k=1}^{2n} \sum_{A=1}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_A}{r_{kA}} + \sum_{k=1}^{2n} \sum_{k>1}^{2n} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{kl}} - \sum_{A=1}^N \frac{\hbar^2}{2M_A} \nabla_A^2 + \sum_{A=1}^N \sum_{B>A}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_A Z_B}{R_{AB}} \quad (2)$$

L'Hamiltonien contient les contributions suivantes :

$$\hat{H} = \hat{T}_N + \hat{V}_{NN} + \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne} \quad (3)$$

- ✓ énergie cinétique des noyaux : $\hat{T}_N = - \sum_{A=1}^N \frac{\hbar^2}{2M_A} \nabla_A^2 \quad (4)$
- ✓ énergie cinétique des électrons : $\hat{T}_e = - \sum_{k=1}^{2n} \frac{\hbar^2}{2m_e} \nabla_k^2 \quad (5)$
- ✓ interaction noyaux – noyaux : $\hat{V}_{NN} = \sum_{A=1}^N \sum_{B>A}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_A Z_B}{R_{AB}} \quad (6)$
- ✓ interaction noyaux –électrons : $\hat{V}_{Ne} = \sum_{k=1}^{2n} \sum_{A=1}^N \frac{e^2}{4\pi\epsilon_0} \frac{Z_A}{r_{kA}} \quad (7)$
- ✓ interaction électrons –électrons : $\hat{V}_{ee} = \sum_{k=1}^{2n} \sum_{l>k}^{2n} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{kl}} \quad (8)$

II.2. 1.2. Les méthodes ab-initio

Les méthodes ab-initio sont des méthodes non empiriques, toutes les intégrales sont rigoureusement calculées et il n'y a pas d'approximation à faire sauf celle de Born Oppenheimer et l'approximation OM-CLOA. Dans les méthodes ab-initio, toutes les particules (noyau et électrons) sont traitées explicitement. On n'utilise aucun paramètre empirique dans le calcul de l'énergie [5], lorsque ces méthodes sont poussées vers leurs limites, elles approchent de la solution exacte de l'équation de Schrödinger non-relativiste.

Si l'on veut obtenir un accord exact avec l'expérience, il est nécessaire d'inclure les termes relativistes et de spin-orbital, les deux n'étant importants uniquement pour les atomes lourds. Dans toutes ces approches, en plus du choix de la méthode, il est nécessaire de choisir une base adéquate. Cette base (au sens mathématique du terme) est un ensemble de fonctions, habituellement centrées sur les différents atomes de la molécule, qui sont utilisées pour étendre les orbitales moléculaires par le postulat de combinaison linéaire d'orbitales atomiques (LCAO). Les méthodes ab initio nécessitent donc d'établir un niveau d'application de la théorie (méthode) et une base [6].

II.2.1.3. Théorie de la fonctionnelle de la densité (DFT)

En mathématique, une fonction f fait correspondre à la variable x le nombre $f(x)$. Une fonctionnelle F est une fonction de fonctions qui fait correspondre à une fonction f un nombre noté $F[f]$.

Avant d'aborder la théorie de la fonctionnelle de la densité, il paraît évident de définir son élément central : la densité électronique.

La densité électronique notée ρ , est une fonction positive qui dépend de trois variables seulement, les coordonnées d'espace. Cette quantité, intégrée sur tout l'espace est égale au nombre total des électrons et s'annule à l'infini.

$$\int \rho(\mathbf{r}) = ne \quad (9)$$

$\rho(\mathbf{r})$ représente, par définition, la probabilité de présence d'un électron dans un élément de volume défini par \mathbf{r} . Elle permet la description d'un système chimique et la détermination complète de ses propriétés contrairement à la fonction d'onde ; c'est pour cette raison que plusieurs recherches ont été entreprises pour la mise en place d'un formalisme se basant sur cette quantité et qui ont abouti à la théorie de la fonctionnelle de la densité.

II.2.1.4. Fondements de la théorie de la fonctionnelle de la densité

Les méthodes de la théorie de la fonctionnelle de la densité (DFT de l'anglais Density Functional Theory) ont acquis une popularité grandissante pendant ces

dernières années. Elles constituent une alternative de choix aux méthodes ab-initio citées précédemment.

L'énergie est donc une fonctionnelle de la densité électronique et s'écrit $E[\rho]$. Par suite, l'énergie électronique totale s'écrit sous forme d'une somme de fonctionnelles : énergies cinétique et potentielle d'attraction électron-noyaux et de répulsion électron-électron.

$$E[\rho] = T[\rho] + E_{ne}[\rho] + E_{ee}[\rho] \quad (10)$$

La DFT a été introduite à l'origine par Thomas [7] et Fermi [8] en 1927 ; quand ils proposèrent de décrire les propriétés électroniques d'un système en terme de fonctionnelles de la densité. En considérant un gaz homogène à électrons libres non interagissant, ils ont pu donner l'expression analytique de l'énergie dans laquelle l'interaction électronique est limitée seulement à la répulsion coulombienne classique. En 1930 Dirac [9], compléta cette théorie, en introduisant une fonctionnelle d'échange, mais ces deux modèles étaient insuffisants pour traiter la liaison chimique. Ils furent complété par les travaux de Wigner (1932) [10] et de Von Weizsäcker (1935) [11] qui eurent l'idée de prendre en compte l'hétérogénéité de la distribution électronique dans la molécule et de rajouter un terme non local dépendant du gradient de la densité électronique à la fonctionnelle énergie cinétique, celle-ci étant très grande devant les autres composantes de l'énergie électronique du système. Il est en effet, très important de calculer le plus précisément possible l'énergie cinétique et toute approximation conduit à des erreurs sur l'énergie totale du système. Un peu plus tard, Hohenberg et Kohn en 1964 [12], ensuite Kohn et Sham en 1965 [13] ont pu définir les bases modernes de la DFT.

II.2.1.5. Méthodes semi-empiriques

La mécanique quantique est une technique mathématique rigoureuse basée sur l'équation de Schrödinger. La solution de cette équation permet d'obtenir des informations précises sur les propriétés géométriques et électroniques de la molécule. Les calculs peuvent être de type ab initio ou semi-empirique (ex : CNDO, PM3). En ab initio, on tient compte de tous les électrons de la molécule et on vise une solution rigoureuse de l'Hamiltonien.

Les calculs semi-empiriques traitent seulement les électrons de valence et utilisent un Hamiltonien plus simple ayant des facteurs de correction basés sur des données expérimentales. L'équation de Schrödinger d'un système moléculaire peut être résolue sans approximation (ab initio) ou en introduisant des approximations (semi-empirique).

En mécanique quantique, on se préoccupe de la distribution des électrons (orbitales) dans l'espace. Les meilleurs programmes comportent des processus d'optimisation de la géométrie [14].

L'objectif de la mécanique quantique est principalement de déterminer l'énergie et la distribution électronique. Ainsi les énergies moléculaires sont calculées en utilisant l'équation de Schrödinger avec le formalisme des orbitales moléculaire (MO). Par les méthodes MO, les énergies peuvent être extraites via le formalisme SCF ("Self Consistent Field") ou par une méthode de perturbation comme PCILO.

En mécanique quantique, on étudie la distribution électronique (orbitale) dans l'espace. Les propriétés moléculaires qui peuvent être calculées sont la géométrie, les spectres de vibration, les spectres d'absorption UV, les moments dipolaires, les spectres électroniques et les fonctions responsables de la réactivité chimique (Orbital HOMO, LUMO).

Les termes énergétiques les plus difficile à calculer sont estimés à partir des données expérimentales. Les temps de calculs sont considérablement raccourcis, mais la méthode est tributaire des composés qui ont servi à l'étalonner. Selon la nature des approximations utilisées [15], on distingue plusieurs variantes :

CNDO : (Complete Neglect of Differential Overlap) 1^{ere} méthode semi empirique, elle a été proposée par Pople, Segal et Santry en 1965. Méthode présentant certains défauts entre autres : elle ne tient pas compte de la règle de Hund.

INDO: (Intermediate Neglect of Differential Overlap) Propose par Pople, Beveridge et Dobosh en 1967. Elle permet de distinguer entre les états singulets et les états triplets d'un système en conservant les intégrales d'échange.

MINDO/3: Proposée par Bingham, Dewar et Lo en 1975. Paramétrisation effectuée en se référant aux résultats expérimentaux et non pas aux résultats ab-initio, de plus l'algorithme d'optimisation utilisé est très efficace (Davidon-Fletcher-Powell).

Cependant, elle surestime la chaleur de la formation des systèmes insaturés et sous-estime celle des molécules contenant des atomes voisins ayant des paires libres.

MNDO: (Modified Neglect of Diatomic Overlap) Proposée par Dewar et Thiel en 1977. Méthode basée sur l'approximation NDDO (Neglect of Diatomic Differential Overlap) qui consiste à négliger le recouvrement différentiel entre orbitales atomiques sur des atomes différents. Cette méthode ne traite pas les métaux de transition et présente des difficultés pour les systèmes conjugués.

AM 1 : (Austin Model 1) Proposée par Dewar en 1985. Il a tenté de corriger les défauts de MNDO.

PM 3 : (Parametric Method 3) Proposée par Stewart en 1989. Présente beaucoup de points en commun avec AM1. D'ailleurs, il existe toujours un débat concernant les mérites relatifs de paramétrisation de chacune d'elles. PM3 et AM1 généralement sont les méthodes les plus rigoureuses dans le logiciel HyperChem, PM3 a été paramétré pour beaucoup des éléments principaux des groupes et quelques métaux de transition. PM3 est différent d'AM1 seulement dans les valeurs des paramètres.

SAM 1 : (Semi-ab-initio Model 1) La méthode la plus récente proposée par Dewar en 1993. Elle inclut la corrélation électronique.

II.2.2. Mécanique moléculaire

Les techniques de modélisation basées sur la mécanique quantique souffrent d'un inconvénient majeur : elles sont très coûteuses en termes de temps de calcul et ne sont dès lors applicables qu'à des systèmes moléculaires de taille restreinte. Au final, le temps nécessaire au traitement d'un système par les méthodes ab-initio est environ proportionnel à la quatrième puissance du nombre d'électrons qu'il contient.

Par contre, la mécanique moléculaire [16] considère l'énergie d'un système uniquement en fonction de ses positions atomiques. Cette approximation repose elle aussi sur les travaux de Born et Oppenheimer. En effet, l'approximation de Born Oppenheimer, en découplant les mouvements des noyaux et des électrons d'une molécule, postule que ces derniers peuvent s'adapter de manière quasi instantanée à la position des noyaux qui sont donc implicitement traités. Le fait d'ignorer les mouvements des électrons épargne ainsi un temps de calcul considérable.

La mécanique moléculaire (MM), appelée parfois "calcul par champ de force empirique", qui est un outil informatique mis à la disposition du chimiste pour étudier la structure 3D des molécules et les propriétés physico-chimiques associées. C'est une méthode non quantique résultante de l'ajustement de résultats expérimentaux sur des fonctions mathématiques simples [17]. En particulier, la mécanique moléculaire permet d'étudier une gamme étendue de propriétés en décrivant l'énergie d'une somme d'une série de contributions rendant compte des interactions intra et intermoléculaires. Pour chacune des contributions, des pénalités énergétiques sont appliquées lorsqu'une variable (par exemple, une longueur de liaison ou un angle de valence), s'écarte de sa valeur de référence. Ces variables du calcul sont alors les coordonnées internes du système (longueur de liaison, angle de valence, angle dièdre et distance entre atomes non liés). Chaque coordonnée est définie par la disposition des atomes, et chaque fonction énergétique qui détermine la contribution d'un écart dépend d'une distance ou d'un angle. Un calcul MM aboutit à une disposition des atomes tels que la somme de toutes les contributions énergétiques est minimisée ; ses résultats concernent surtout la géométrie et l'énergie du système [17].

La MM est une méthode empirique où les atomes (noyaux) sont représentés par des masses ou des sphères, et les liaisons par des ressorts de différentes forces (figure II.1).

Les valeurs de ces forces proviennent de données expérimentales de spectroscopie vibrationnelle, de données de diffraction des RX ou de valeurs obtenues par les méthodes ab-initio. Le calcul de l'ensemble de ces forces, selon les règles établies pour les oscillateurs harmoniques, permet d'obtenir l'énergie moléculaire et de déterminer la structure tridimensionnelle. Aujourd'hui, la MM est utilisée pour interpréter les données de RMN [17].

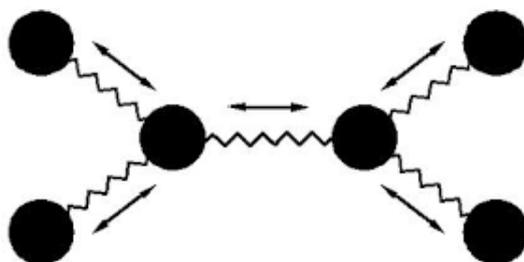


Figure II.1 : Représentation mécanique d'une structure moléculaire.

L'avantage principal de la MM par rapport aux autres méthodes est la rapidité des calculs. Ceci est dû à une simplification des équations du champ de force ; il est donc possible de traiter des systèmes moléculaires très grands. En revanche, cette méthode ne prend pas en compte la structure électronique moléculaire. Il est donc impossible d'étudier des systèmes dans lesquels les effets électroniques sont prédominants [17].

L'idée directrice de cette méthode est d'établir, par le choix des fonctions énergétiques et des paramètres qu'elles contiennent, un modèle mathématique, le «champ de force», qui représente aussi bien que possible les variations de l'énergie potentielle avec la géométrie moléculaire. Ce modèle dépend étroitement des propriétés expérimentales retenues, selon que l'on cherche à calculer la géométrie, la chaleur de formation, les fréquences vibrationnelles ou, même, la réactivité d'une entité moléculaire. Par conséquent, les solutions proposées par divers auteurs dont les objectifs diffèrent, ne sont pas convergentes bien que la méthodologie soit la même. Ainsi, il n'existe pas encore de modèle unique permettant de simuler tous les aspects du comportement moléculaire, mais un ensemble de modèles [17].

II.2.2.1. Champ de force

Le champ de force est une expression que la mécanique moléculaire à emprunter à la spectroscopie vibrationnelle, en modifiant légèrement sa signification

[18]. Sous ce terme sont en fait regroupés deux éléments : d'une part l'expression des différentes fonctions contribuant au calcul énergétique et d'autre part les valeurs des différentes constantes paramétrant ces fonctions. Ces paramètres sont identifiés à partir de données expérimentales (IR, micro - onde...) ou évalués théoriquement. L'énergie totale du système est donc une énergie potentielle multidimensionnelle décrivant les interactions intramoléculaires ou interactions liantes (impliquant des atomes reliés par des liaisons explicites) et les interactions intermoléculaires ou interactions non liantes (impliquant des atomes non liés directement par des liaisons explicites) :

$$E_{Totale} = E_{Intermoléculaire} + E_{Intramoléculaire} \quad (11)$$

Par conséquent, le champ de force est un modèle du système réel que l'on veut étudier. De sa qualité descriptive dépend la qualité du résultat des simulations numériques par comparaison avec les valeurs expérimentales. L'amélioration des champs de forces, modèles de système réel, représente un enjeu majeur pour le succès des techniques de la simulation moléculaire à prédire avec précision les propriétés et le comportement de systèmes toujours plus nombreux et plus complexes. La portée d'une interaction intermoléculaire qui concerne surtout les interactions non-liantes est déterminée par l'exposant de la distance entre les particules concernées : un potentiel en $1/r$ comme le potentiel de coulomb générera une force en $1/r^2$ qui sera ressentie beaucoup plus loin (portée jusqu'à 25 Å) qu'un potentiel en $1/r^6$ comme le terme dispersif du potentiel de Lennard-Jones (portée de 1 à 5 Å) [18].

En MM, l'expression "champ de force" recouvre l'ensemble de ces fonctions, ainsi que pour chaque type d'atome les valeurs des paramètres qu'elles contiennent. Ainsi par exemple comme dans un champ de force de valence du type harmonique en spectroscopie vibrationnelle, les fonctions représentant les contributions dues à une élongation de liaison ou une déformation d'angle de valence seront du type :

$$E(r) = 1/2Kr(r - r_0)^2 \quad (12)$$

$$E(\theta) = 1/2K_\theta(\theta - \theta_0)^2 \quad (13)$$

La plupart des champs de force en MM contiennent un terme croisé qui implique deux longueurs de liaison et l'angle contenu entre elles. Ce terme croisé rend compte d'un phénomène bien connu ; lorsqu'un angle de valence est anormalement faible on constate que les longueurs des liaisons associées sont anormalement élevées. Les champs de force, les plus utilisées actuellement en MM sont du type champ de force de valence généralisé. Par exemple dans le champ de force MM2 d'Allinger [17], l'énergie stérique" s'écrit :

$$E_{MM} = E_r + E_\theta + E_{r,\theta} + E_d + E_\phi + E_D \quad (14)$$

Les termes E_r et E_θ de l'équation (14) ont déjà été définis de même que le terme $E_{r,\theta}$ qui représente l'influence d'un angle de valence anormalement faible sur les longueurs des liaisons définissant cet angle (ce terme est désigné par "Stretch bend", par Allinger), E_ϕ représente les contributions des torsions autour des liaisons, E_d représente les contributions des interactions de Van der Waals entre atomes non liés et enfin E_D représente les contributions des interactions dipôle-dipôle, dipôle-charge, charge-charge. Bien que certaines des fonctions définissant un champ de force en MM soient analogues aux fonctions d'un champ de force en spectroscopie vibrationnelle, il est important de remarquer que les paramètres qu'elles contiennent ne sont pas caractéristiques d'une molécule mais d'un champ de force, ainsi K_r et K_θ sont des constantes de force, de même r_0 et θ_0 correspondent à des valeurs qui ont été fixées de façon à ce que le champ de force dans lequel elles sont incluses, donne les meilleurs résultats possibles dans le calcul des géométries d'un large ensemble de molécules [18].

II.2.2.2. Quelques champs de force

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. Type de composé : carbohydrate, complexe métallique.

Environnement : gaz, solution.

Type d'interaction à étudier : liaison hydrogène,

De tels champs de force sont apparus en début des années 1970 et continuent à évaluer aujourd'hui.

- **MM2 / MM3 / MM4** : est le premier champ de force développé par Allinger et col. Il a été conçu au début pour les molécules simples (alcane, alcène, alcynes non conjugués, amines...), mais ses versions améliorées MM3 (1989), MM4 (1996) [19] lui permet de traiter des molécules organiques de plus en plus complexes.

- **OPLS** : Le programme OPLS (Optimized Potentials for Liquid Simulations), comme l'indique son nom, est conçu pour optimiser le potentiel qui permet la description des propriétés de solvation. Il est écrit par W. L. Jorgensen et J. Tirado Rives [20].

- **GROMOS** : (Groningen Molecular Simulation Program Package), est écrit par Van Gunsteren [21] et 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.

- **CHARM (Bio+)** : Développe par Karplus et col [22], pour le calcul de biomolécules. Son concept est semblable à celui d'AMBER. Bien qu'au début, ce champ de force est conçu pour les aminoacides et les protéines, maintenant il traite d'autre biomolécules.

-**SPASIBA** : (Spectroscopic Potentiel Algorithm for Simulating biomolecular Conformational Adaptability), élaboré par Gérard Vergoten et col. (1995). Il combine le champ de force spectroscopique modifié de Urey-Bradly-Shimanouchi et le champ de force AMBER. Il permet de trouver les structures, les énergies de conformation et les fréquences vibrationnelle au minimum énergétique d'une molécule [23].

- **EMO** : Le programme EMO (Energy Of Molecule), est développé par B. Blaise, il est basé sur le champ de force MM2. Il est articulé autour de trois menus :

- ✓ Menu 1 : entrée de la molécule à l'aide d'un clavier.
- ✓ Menu 2 : manipulation géométrique sur la molécule.
- ✓ Menu 3 : minimisation de l'énergie par mécanique moléculaire

-**AMBER**: (Assisted Model Building with Energy Refinement), a été écrit par Kollman [24]. Le champ est paramétré pour les protéines et les acides nucléiques (UCSF, 1994). Il a été utilisé pour les polymères et pour d'autres petites molécules [24].

II.2.2.3. Domaine d'application de la mécanique moléculaire

On peut diviser l'application de la MM en trois catégories :

- Soit pour obtenir une géométrie à laquelle on attache de l'intérêt. Cette situation se présente lorsque la modélisation guide l'interprétation des résultats provenant des études de structure par rayons X ou par diffraction électronique, ou lorsqu'il s'agit de modéliser une molécule pour les besoins de l'infographie.
- Dans l'interprétation des effets stériques sur la réactivité ou bien de la stabilité relative des isomères en tant qu'énergie stérique ou de tension.
- Quand aucune liaison n'est rompue, ni formée et qu'aucun intermédiaire chargé n'intervient, l'interconversion conformationnelle se prête particulièrement bien à une description par la MM. On peut obtenir grâce à cette analyse des informations structurales sous forme d'un profil énergétique (en fonction d'un angle dièdre par exemple) ou des cartes énergétiques 3D.

En conclusion, on peut dire que la mécanique moléculaire aujourd'hui est à la porte de tous les chercheurs. La mécanique moléculaire ne peut pas encore rivaliser

avec la mécanique quantique dans beaucoup de domaine qui lui sont propre, mais elle reste une méthode de choix dans l'interprétation de phénomènes sous contrôle stérique et dans le calcul de structure.

La mécanique moléculaire permet de passer en revue de grosses molécules (Produits pharmaceutiques, colorants, etc.) Pour établir des relations entre structure et réactivité et ainsi faire un tri avant de passer au stade expérimental.

La différence du temps de calculs par la mécanique moléculaire par rapport aux autres méthodes quantiques est d'environ de quelque puissance de dix, cette différence augmente en fonction de la taille de la molécule [17].

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 [23] 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 [23].

II.2. 4. Méthodes de corrélation quantitative (QSAR)

II.2.4.1. Paramètres biologiques

Les données biologiques sont habituellement exprimées sur une échelle logarithmique en raison de la relation linéaire entre la réponse et le logarithme de dose dans la région centrale de la courbe de log dose-réponse. Les logarithmes inverses de l'activité ($\log 1/C$) sont également utilisés pour obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces. Des exemples de données biochimiques ou biologiques, utilisés dans l'analyse de QSAR, sont décrits dans le tableau II.1 [25].

Source d'activité	Paramètres biologiques
1. Récepteurs isolés	
Constante de vitesse	Log k
Constante de Michaelis-Menten	Log 1/K _m
Constante d'inhibition	Log 1/K _i
2. Systèmes cellulaires	
Constante d'inhibition	Log 1/IC ₅₀
Résistance croisée	Log CR
Données biologiques in vitro	Log 1/C
Mutation de gène	Log TA ₉₈
3. Systèmes in vivo	
Facteur de bioconcentration	Log BCF
Vitesses de la réaction in vivo	Log I (induction)
Vitesses pharmacodynamiques	Log T (clairance totale)

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

II.2.4.2. Descripteurs moléculaires

Les descripteurs moléculaires sont des propriétés numériques qui peuvent être calculées à partir de la table de connectivité d'une molécule ou d'une représentation planaire (2D) de la structure. Ils sont basés sur les éléments présents, les charges partielles, la nature des liaisons, mais n'exploitent pas les coordonnées atomiques spatiales [26]. On peut calculer ainsi les paramètres caractérisant la structure moléculaire:

II.2.4.3. Volume et surface moléculaire

Le volume moléculaire est souvent calculé par la méthode des grilles [27]. Le volume d'un atome est déterminé à partir de rayon de Van der Waals R donné par

$$: \frac{4}{3} \pi R^3$$

La figure II.2 donne une représentation d'un atome dont le rayon de Van der Waals est R, entouré d'une grille tridimensionnelle avec même distance entre les points. Le centre de cette grille est le centre de l'atome et les bords de la grille correspondent au rayon de Van der Waals

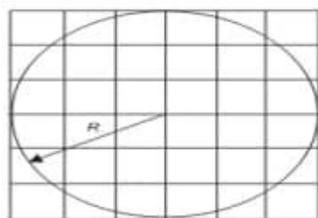


Figure II.2: Atome entouré par une grille.

Pour chaque point de la grille, on calcule sa distance par rapport au centre et on détermine si le point de la grille se trouve à l'intérieur ou à l'extérieur de l'atome. Si n est le nombre total de points de grille et n_a le nombre de points qui se trouvent à l'intérieur de l'atome, le volume est donnée par :

$$\frac{V}{8R^3} = \frac{n_a}{n} \quad (15)$$

Pour un système polyatomique, nous devons donner une attention particulière aux points de la grille qui se trouvent dans la zone de recouvrement. Figure II.3 montre deux atomes A et B, avec des rayons R_A et R_B . La zone de recouvrement est marquée par X.

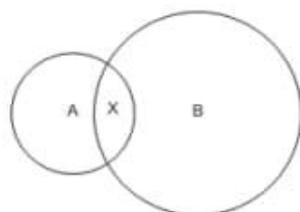


Figure II.3: Atomes A, B et la zone de recouvrement X.

Nous savons que le volume de l'atome A est donné par $\frac{4}{3}\pi R^3$. Nous entourons maintenant l'atome B par une grille comme décrit ci-dessus, et nous testons chaque point de la grille à son tour. Si le point de la grille se trouve dans la sphère B, nous testons pour voir si elle se trouve dans la région X et donc déjà a été considérée comme faisant partie du volume d'atome A. L'algorithme se produit jusqu'à ce que tous les atomes aient été testés. Le volume moléculaire est obtenu en additionnant toutes les contributions atomiques. Il existe des méthodes similaires pour l'estimation de la surface moléculaire.

II.2.4.4. Polarisabilité

La polarisabilité d'une molécule est définie par la capacité de son système électronique de se déformer par un champ électrique externe, il joue un rôle important dans la modélisation de nombreuses propriétés moléculaires et des activités biologiques [28].

Une molécule très polarisée peut donner des attractions fortes avec d'autres molécules. La polarisabilité d'une molécule peut également améliorer sa solubilité aqueuse [29].

II.2.4.5. Réfractivité molaire

La réfractivité molaire (MR) est un critère important dans la mesure des facteurs stériques. Elle est généralement désignée comme une simple mesure du volume occupé soit par un atome individuel ou par un groupe d'atomes [30]. La réfractivité molaire est particulièrement importante dans une situation où le substituant possède les électrons π ou les doublets électroniques libres. La réfractivité molaire est donnée par la relation suivante :

$$MR = \frac{(n^2-1)}{(n^2+2)} \cdot \frac{MW}{d} \quad (16)$$

Où

n : indice de réfraction.

MW : masse moléculaire.

d: densité.

MW/d : volume.

$(n^2 - 1) / (n^2 + 2)$: facteur correctif.

La détermination expérimentale de la polarisabilité est basée sur l'équation de Lorentz-Lorentz [31]. Cette équation établit une relation entre l'indice de réfraction d'une molécule et sa polarisabilité (α). La forme générale de cette équation est donnée par :

$$MR = \frac{(n^2-1)}{(n^2+2)} \cdot \frac{MW}{d} \cdot \frac{4\pi}{3} N \alpha \quad (17)$$

II.2.4.6. Energie d'hydratation

La liaison hydrogène (ou liaison H) est de type électrostatique (charge partielle, dipôle) et stérique entre deux groupements d'une même molécule ou de deux molécules voisines. La liaison hydrogène joue un rôle primordial dans la solubilité des molécules médicamenteuses et leurs interactions avec les récepteurs

biologiques [32].

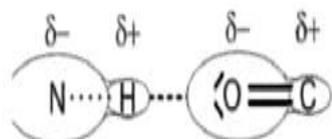


Figure II.4: La liaison hydrogène.

Dans le figure II.4, l'oxygène (O) est appelé l'accepteur (accepteur de proton H⁺) et l'azote (NH) est appelé le donneur (donneur de proton H⁺).

Les liaisons hydrogènes sont à la fois plus longues et plus faibles que les liaisons covalentes établies entre les mêmes atomes. La solubilité des substances dans un milieu aqueux dépend fortement de leur capacité de former des liaisons hydrogènes avec l'eau.

Hine et Mookerjee ont proposé un modèle qui considère que l'énergie libre des interactions entre les petites molécules et l'eau peut être estimée par une fonction additive de leurs groupes constitutifs [33]. Pour les petites molécules, dont tous les fragments accessibles au solvant, l'énergie libre d'hydratation est:

$$\Delta G_{Hyd} = \sum_{i=1}^N \Delta G_{fr_i} \quad (18)$$

Où ΔG_{fr_i} est l'énergie libre d'hydratation du fragment i et N est le nombre total de fragments dans la molécule.

Dans le modèle proposé, les systèmes aromatiques sont traités dans une représentation de Kekulé pour l'analyse fragmentaire, mais l'énergie obtenue avec les groupes C=C doit être corrigée comme suit:

$$\Delta G_{Hyd_{aro}} = n \times \Delta G_{C=C} - m \times 1.8 \quad (19)$$

Où n est le nombre de doubles liaisons dans la représentation de Kekulé, et m le nombre des centres aromatiques du système, par exemple:

Dans le benzène n = 3 et m = 1.

Dans le naphthalène n = 5 et m = 2.

Ce qui donne les valeurs -1.05 et -2.1 kcal/mole de l'énergie d'hydratation pour ces molécules respectivement.

II.2.4.7. Coefficient de partition (Log P)

Le partage d'une molécule entre une phase aqueuse et une phase lipidique conditionne en partie ses propriétés biologiques telles que le transport, le passage à travers les membranes, la biodisponibilité (distribution et accumulation), l'affinité pour un récepteur et la fixation par une protéine, l'activité pharmacologique ou encore la toxicité. Depuis les travaux de Collander à la fin des années 1950, puis ceux du groupe de Hansch quelques années plus tard, le coefficient de partage P d'une molécule dans un système bi-phasique constitué de deux solvants non-miscibles (le plus souvent le système n-octanol/eau), est reconnu pour sa faculté à mimer le passage de cette molécule à travers les membranes biologiques. Pour des solutions diluées, ce coefficient de partage n-octanol/eau est le rapport de la concentration d'une molécule de soluté dans le n-octanol sur sa concentration dans l'eau lorsque le système biphasique est en équilibre [34].

Log P, est l'un des paramètres largement utilisé dans des études de relations structure-activité quantitatives (QSARs) dans les sciences pharmaceutiques, biochimiques, toxicologiques et dans les sciences de l'environnement. La lipophilie intéresse donc tout autant la communauté qui étudie les problèmes de santé humaine que celle qui est impliquée dans les problèmes de l'environnement.

Un médicament doit être caractérisé par une balance correcte entre ses propriétés hydrophiles et hydrophobes. Pendant leur parcours à travers du corps, les molécules médicamenteuses doivent se dissoudre dans les milieux aqueux de cytoplasme et des liquides extracellulaires ainsi que dans le milieu apolaire de la membrane cellulaire (Figure II.5). Ainsi, la solubilité dans des milieux à la fois polaires et apolaires joue un rôle important dans la disponibilité et de la distribution des médicaments.

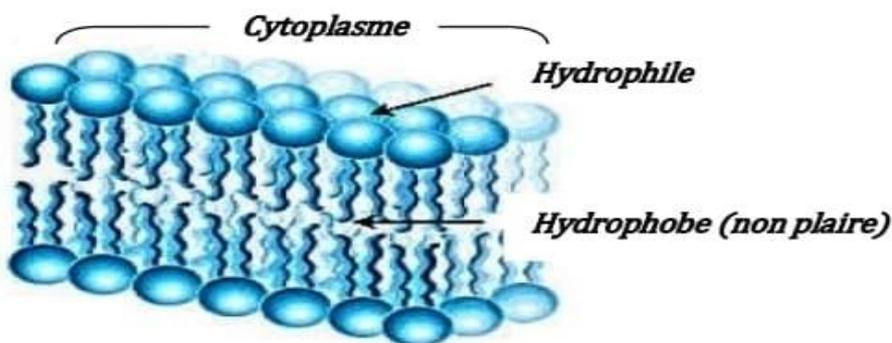


Figure II.5: Polarité des différents milieux cellulaires.

Le coefficient de partage d'une substance chimique entre deux phases non miscibles est le rapport des concentrations à l'équilibre du composé dans la phase organique apolaire et la phase aqueuse.

Le solvant organique couramment utilisés pour mesurer expérimentalement le coefficient de partage des molécules médicamenteuses est le 1-octanol. Le coefficient de partage P se calcule de la manière suivante:

$$\text{Log P} = \frac{\text{Concentration du composé en solution dans l'octanol}}{\text{Concentration du composé en solution dans l'eau}} \quad (20)$$

La phase aqueuse est habituellement tamponnée à pH 7,4. Log P, est souvent appelée la constante de lipophilie d'une molécule. Une valeur positive pour log P, ($P > 1$), indique que le composé est lipophile. Une valeur négative pour log P, ($P < 1$) indique que le composé est hydrophile.

La plupart des méthodes expérimentales de détermination de log P souffrent du même inconvénient, à savoir que leur domaine d'application est relativement étroit. D'autre part, du fait de la nature intrinsèque de certaines molécules, leurs log P sont inaccessibles à l'expérience. C'est le cas en particulier des surfactants qui ont tendance à s'accumuler à l'interface du système bi-phasique au lieu de se disperser dans les deux phases [34]. Enfin, dans le domaine de la conception assistée par ordinateur ou dans le domaine de la chimie combinatoire, les chercheurs travaillent sur des modèles moléculaires avant même que les molécules aient été synthétisées. Ceci explique le succès des nombreuses méthodes d'estimation de log P qui ont été décrites dans la littérature depuis plus de trente ans. Les plus anciennes sont des méthodes fragmentables dans lesquelles une molécule est divisée en fragments prédéfinis et les contributions correspondantes sont sommées pour conduire à une valeur estimée du log P [35].

Parmi les méthodes utilisées dans l'estimation de log P, on trouve la méthode Hansch. Cette méthode repose sur l'utilisation de paramètre d'hydrophobicité d'un substituant qui mesure l'intensité hydrophobe du substituant en question par rapport à celle de l'hydrogène.

On peut mesurer log P expérimentalement pour un composé donné sans et avec le substituant X en utilisant l'équation suivante:

$$\Pi_X = \log P_X - \log P_H \quad (21)$$

Une valeur positive de π indique que le substituant est plus hydrophobe que l'hydrogène. Une valeur négative signifie que moins hydrophobe que l'hydrogène.

La méthode de Rekker permet de calculer le coefficient de partage octanol/eau. Elle consiste à décomposer la molécule en fragments à chacun desquels on attribue une lipophilie f appelée constante fragmentale hydrophobe. La formule générale de Rekker est donnée par :

$$\log P = \sum f + k.CM \quad (22)$$

Avec k entier et $CM = 0, 219$ (l'entier k dépend du type d'interaction).

II.2.4.8. Modèles statistiques

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

Dans la terminologie des statistiques, la propriété étudiée est appelée variable cible ou variable dépendante ou encore variable Y . C'est la caractéristique que l'on cherche à expliquer et à optimiser. Les descripteurs, en l'occurrence chimiques, sont appelés variables explicatives, variables indépendantes ou encore variables X .

L'analyse statistique détermine et quantifie les corrélations entre les descripteurs et la variable cible. Elle indique également la contribution relative de chaque descripteur dans l'explication globale de l'activité. Le modèle statistique est une équation donnant la valeur de la variable cible en fonction de la somme des valeurs pondérées des descripteurs. Les principaux outils statistiques pour obtenir un modèle sont :

- ✓ la régression linéaire multivariée (Multivariate Linear Regression - MLR) ;
- ✓ la régression en composante principale (Principale Component Regression - PCR) ;
- ✓ la régression des moindres carrés partiels (Partial Least Squares - PLS) [36] ;
- ✓ les réseaux de neurones artificiels (Artificial Neural Network - ANN).

II.2.4.9. Les applications de l'étude QSAR/QSPR

Certaines études QSAR semblent être un peu plus d'études universitaires, il existe un grand nombre d'applications de ces modèles tels que [37]:

- ✓ 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 chimiques 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 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 du devenir des molécules qui sont libérées dans l'environnement.
- ✓ La prédiction des effets conjugués de molécules, que ce soit dans des mélanges ou des formulations.

II.3. La Régression Linéaire Multiple (MLR)

La régression linéaire multiple MLR est l'une des méthodes de modélisation les plus en vogue grâce à sa simplicité d'utilisation et son interprétation facile. 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 [38].

La méthode MLR se base sur l'hypothèse que la propriété « y » dépend linéairement des différentes variables (les descripteurs), selon la relation :

$$y = a_0 + \sum_{i=1}^n a_i x_i \quad (23)$$

Où:

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

Les **x_i** : sont les variables indépendantes (explicatives)

n : est le nombre de variables explicatives

a₀ : est la constante de l'équation du modèle les **a_i** représente les coefficients de descripteurs dans l'équation du modèle

Les variables indépendantes **x_i**, comme leur nom l'indique, sont supposées indépendantes entre elles.

La régression linéaire multiple a certains désavantages. Le principal découle de sa linéarité. Elle est défailante pour la mise en évidence de dépendances non-linéaires. Cela dit, elle n'en reste pas moins une méthode simple et efficace dans la plupart des cas. De plus, pour peu que les variables indépendantes soient choisies de manière raisonnée, les équations obtenues peuvent être interprétées d'un point de vue phénoménologique.

II.4. IC₅₀ : Constante d'inhibition d'un complexe

L'IC₅₀ traduit la concentration en inhibiteur (I) nécessaire pour inhiber 50% de l'interaction entre le récepteur (L= Lectine) et son ligand (G= Glycocluster). Ainsi, plus un glycocluster possède une forte affinité avec une lectine, plus son IC₅₀ est grand.



Il existe plusieurs types de réaction d'inhibition (compétitive, incompétitive, noncompétitive, etc.) et pour chacune, une équation de Michaelis et Menton associée. Cheng et Prusoff ont proposé des formules pour relier l'IC₅₀ aux autres paramètres cinétiques, dans le cas d'un inhibiteur compétitif [39].

$$IC_{50} = K_d \left(1 + \frac{[G]}{K_m} \right) \quad (26)$$

Avec [G] = concentration en Glycocluster, K_d = coefficient de dissociation du complexe Glycocluster/Lectine et K_m = constante de Michaelis du système associé. L'industrie pharmaceutique utilise très souvent le paramètre IC₅₀ pour des raisons pratiques lors de la comparaison de molécules actives contre un germe ou un virus. En effet, la valeur de l'IC₅₀ permet d'obtenir une indication sur l'efficacité d'une molécule sans connaître précisément les autres paramètres pouvant influencer la réaction. Il est facile de comparer l'affinité pour un récepteur donné de plusieurs ligands synthétiques par rapport au même substrat. Néanmoins,

Les mesures de compétition (IC₅₀) sont très dépendantes de la technique utilisée et des conditions expérimentales [40]. Dans l'idéal, il faudrait que la mesure ne dépende pas des conditions expérimentales. Il est donc très important de comparer les résultats obtenus par d'autres méthodes de mesures d'affinités comme la mesure de la constante de dissociation par exemple.

Références bibliographiques

- [1] H.-D. Höltje, G. Folkeis, *Molecular Modeling: Basic Principles and Applications*. VCH, New-York, 1997.
- [2] MELKEMI Nadjib , Thèse de doctorat , université Biskra ,2013.
- [3] R.G.Parr , W.Wang, oxford university press :Oxford ,1989.
- [4] D.Harkati , Thèse de doctorat , université Biskra , 2015.
- [5] PD.Leeson , and M.Andrew , Davis J. of med. chem ., 47, (25) , 6338-6348 ,(2004).
- [6] Zergui Amira , mémoire de master, université Biskra, 2019
- [7] Thomas L. H. Cambridge Phil. Soc., 23 , 713(1927).
- [8] Fermi E., Statistical method to determine some properties of atoms ,Rendiconti, 6 ,602(1927).
- [9] Dirac P. A. M., Note on Exchange Phenomena in the Thomas Atom, Proc. Cambridge Phil. Soc, 26 ,376(1930).
- [10] Wigner E., On the Quantum Correction For Thermodynamic Equilibrium, Phys. Rev, 40 ,749(1932).
- [11] von Weisacker C. F., Z. Phys, 96,431(1935).
- [12] Hohenberg P., Kohn W., Inhomogeneous Electron Gas,Phys. Rev. B, 136 ,864(1964).
- [13] Kohn W., Sham L. J., Self-Consistent Equations Including Exchange and Correlation Effects, Phys. Rev. A ,140 ,1133(1965).
- [14] Sordo J. A. , Probst M. , Corongiu G. , Chin S. and Clementi E., Ab initio pair potentials for the interactions between aliphatic amino acids, J. Am.Chem.Soc, 109,1708 (1987).
- [15] Liotta D., *Advances in Molecular Modeling*, JAI Press Inc, London (1988). [16] U.Bukert. U, N.L.Allinger. *Molecular Mechanics*, American Chemical Society monograph,Washington, DC, 177 (1982).
- [17] N.L.Allinger. Conformational analysis. 130. MM2. A hydrocarbon force field utilizing V1 and V2 torsional terms. Am. J Chem. Soc., 99, 8127-8134 (1977). [18] J.S.Lomas. L'actualité chimique. La mécanique moléculaire, une méthode non quantique pour le calcul de la structure et de l'énergie d'entité moléculaire, Mai, 7

- (1986). [19] L.N.Allinger, K. Chen, J-H. Lii. An improved force field (MM4) for saturated hydrocarbons. *J. Comp. Chem.*, 17, 642-668 (1996).
- [20] 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).
- [21] W.F .Van Gusten, M.Karplus. Effect of constraints on the dynamics of macromolecules. *Macromolecules*, 15, 1528-1544 (1982).
- [22] B.R. Brooks, R.E .Brucoleri, B. D. Ofalson, D. J. States, S. Swaminathan, M. Karplus. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* , 4, 187-217 (1983).
- [23] 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).
- [24] D.F .Mierke , O.E .Said-Nejad , P.W .Schiller , M.Goodman , *Biopolymers*, 29, 179-96 (1990).
- [25] Selassie C.D., History of quantitative Structure-Activity relationships", in Book "History of quantitative Structure-Activity relationships", D. J. Abraham, Editor., John Wiley & Sons, Inc., California(2003).
- [26] Fortuné A., Thèse de doctorat, Techniques de Modélisation Moléculaire Appliquées à l'Etude et à l'Optimisation de Molécules Immunogènes et de Modulateurs de la Chimiorésistance, Université Joseph Fourier – Grenoble I (2006).
- [27] Hinchliffe A., *Molecular Modeling for Beginners*, JohnWiley & Sons Ltd, United Kingdom (2008).
- [28] Wang J., Xie X.Q., Hou T., Xu X., *J. Phys. Chem.*, 111, 4443 (2007).
- [29] Qaneinasab S., Bayat Z., *J. Chem. Pharm. Res.*, 3, 561 (2011). [30] Patrick G.L., Winter H.D., Langenaeker W., Tollenaere J.P., *Computational Medicinal Chemistry for Drug Discovery*, Marcel Dekker, New York (1995).
- [31] Autin Ludovic M., Thèse de Doctorat, Analyse des systèmes tenase et prothrombinase par bioinformatique structurale: prédiction de complexes macromoléculaires et proposition d'agents anti-coagulants, Université Paris 5(2005).
- [32] Deleu M., thèse de doctorat es sciences, FUSAGx, Belgique(2000). [33] Pépe G., Guiliani G., Loustalet S., Halfon P., *Eur. J. Med. Chem.*, 37, 865 (2002).
- [34] Carpy A., Importance de la lipophilie en modélisation moléculaire, *Analisis*, 27 (1999)

[35] Mannhold R., Kubinyi H., Timmerman H., *Methods and Principles in Med Chem*, VCH Publishers, New York (1996). [36] Khadikar P., Jaiswal M., Gupta M., Mandloi D. and Sisodia R. S., *Bioorg. Med. Chem. Lett.*, 15, 1249 (2005).

[37] J. Leszczynski, *Challenges and advances in computational chemistry and phys*, université Jackson state, U.S.A, 2010.

[38] Roy K, Kar S, Narayan Das R., *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, Chapter 6-Selected Statistical Methods in QSAR, Academic Press, Boston, 2015, 191

[39]: Cheng, Y., Prusoff, W.H., 1973. *Biochem. Pharmacol.* 22(23), 3099-3108. [40]: Kitov, P.I., Bundle, D.R., 2003. *J. Am. Chem. Soc.* 125(52), 16271-16284.

CHAPITRE III

***Etude Quantitative De la
Relation Structure/Activité
Biologique.***

III.1. Etude QSAR à l'aide de descripteurs moléculaires théoriques des dérivés xanthone

III.1.1. Introduction

La Relation Quantitative Structure et Activité (QSAR) est un outil permet d'expliquer l'interaction des composés chimiques avec les sujets actuels, le principe des méthodes QSAR est comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative des propriétés moléculaires, appelées descripteurs, avec une observable macroscopique (activité biologique, toxicité, propriété physico-chimique, etc.), pour une série des molécules biologiquement actives à l'aide des méthodes d'analyse de données [1].

Dans ce travail, nous nous intéressons à l'étude des propriétés physico-chimiques et électroniques des dérivés de xanthone qui permettent de prévoir ces paramètres, qui influent sur l'activité biologique, afin de prédire l'activité biologique de nouvelles molécules. C'est pour cette raison, nous avons choisis une série de trente-neuf dérivés des différents descripteurs physico-chimiques et électroniques.

III.1.2. Méthodologie

III.1.2.1. Base de données

Une série de 39 dérivés de xanthone a été utilisée pour développer des modèles QSAR. Les structures et l'activité biologique expérimentale des dérivés de xanthone sont rapportées dans le tableau III.1. L'activité a été exprimée par pIC50.

Avant de procéder l'étude QSAR, nous avons sélectionné l'ensemble de molécules en deux classes. La première classe comporte 26 molécules nommées « training set » utilisés pour développer des modèles QSAR. La seconde classe comporte 13 molécules nommées « test set » utilisés dans la validation des modèles QSAR obtenus des molécules de training set.

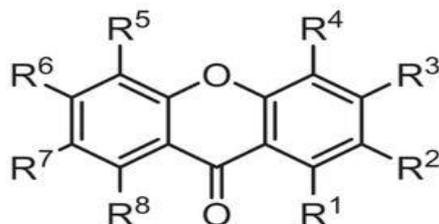


Figure III.1: Structures chimiques des dérivés de xanthone

a : test set

Co mpo se	R1	R2	R3	R4	R5	R6	R7	R8	pIC ₅₀ (exp)	- Log Ic 50
1 ^a	OH	glc	OH	H	H	OH	OH	H	85.35	-1.931203525
2 ^a	OH	glc	OH	H	H	OH	O-glc	OH	158	-2.198657087
3 ^a	OH	H	OH	Glc	H	OH	OH	H	76.98	-1.886377907
4 ^a	OH	H	OH	H	OH	H	H	O-glc	679	-2.831869774
5 ^a	OH	H	OH	H	H	H	H	H	160.8	-2.206286044
6 ^a	OH	H	OH	H	H	OH	H	OH	17.1	-1.23299611
7 ^a	OH	H	OH	H	H	H	OH	OH	7.09	-0.850646235
8 ^a	OH	H	OMe	H	H	OH	OH	OMe	44.5	-1.648360011
9 ^a	O- glc	H	OMe	H	OMe	OH	H	H	87.6	-1.942504106
10 ^a	OH	isoprenyl	OH	H	H	OH	OMe	isoprenyl	5	-0.698970004
11 ^a	OH	isoprenyl	OMe	H	H	OH	OMe	isoprenyl	14.4	-1.158362492
12 ^a	OH	isoprenyl	OH	H	H	OH	OH	isoprenyl	1.5	-0.176091259
13 ^a	OH	isoprenyl	OMe	H	OH	H	H	H	58.5. 4	-1.767155866
14	OH	isoprenyl	OMe	H	OH	H	H	OH	50.5	-1.703291378
15	OH	isoprenyl	OH	Isopr enyl	OH	H	H	H	21.5	-1.33243846
16	OH	H	H	H	H	H	H	H	177.4	-2.248953615
17	OH	H	H	H	H	H	H	OH	91.5	-1.961421094
18	OH	H	H	H	H	OH	H	H	131.4	-2.118595365
19	OH	H	OH	H	H	H	H	OH	81.8	-1.912753304
20	OH	H	OMe	OMe	H	H	O-glc	OH	133.7	-2.126131407
21	OH	H	OMe	H	H	H	O-glc	OH	394.9	-2.596487134
22	OH	H	OMe	OMe	H	H	O-glc	H	115.0	-2.06069784
23	O- glc	H	OMe	H	H	H	OMe	OMe	83.8	-1.923244019
24	OH	H	OMe	OMe	H	H	OH	OH	5.42	-0.733999287
25	OH	H	OMe	H	H	H	OMe	OH	107.8	-2.032618761
26	OH	H	OMe	H	H	H	OMe	OMe	142.1	-2.152594078
27	OH	H	OMe	H	H	H	OH	OH	30.6	-1.485721426
28	OH	H	OH	H	OH	H	H	OH	5.3	-0.72427587
29	OH	H	OMe	OMe	H	H	OH	OMe	77.4	-1.888740961
30	OH	H	OMe	OMe	H	H	OMe	OMe	65.8	-1.818225894
31	OH	H	OH	H	H	H	OMe	OH	71.2	-1.852479994
32	OH	H	OMe	H	H	H	OH	H	17.6	-1.245512668
33	OH	H	OH	H	H	H	OH	H	31.8	-1.50242712
34	OH	H	OH	H	H	H	OMe	OMe	18.0	-1.255272505
35	OH	H	OMe	OMe	OH	H	H	H	82.2	-1.914871818
36	OH	H	OMe	OMe	OMe	H	H	OMe	40.1	-1.603144373
37	OH	H	OMe	H	OMe	H	H	OMe	75.8	-1.879669206
38	O- glc	H	OH	OMe	OMe	H	H	OH	578	-2.761927838
39	OH	geranyl	OH	H	OH	H	H	OH	6.9	-1.255272505

Tableau III.1 : Structures chimique et activité biologique expérimentale des molécules utilisées dans l'étude QSAR.

III.1.2.2 Sélection des descripteurs et méthodes de calcul

Un descripteur moléculaire peut être considéré comme la conséquence d'un processus logique et mathématique, appliqué à l'information chimique codifiée à travers la représentation d'une molécule [2].

Le choix des descripteurs dépend des outils dont on dispose, de la nature des composés décrits et de la propriété ciblée.

Les molécules de la série de xanthone ont été pré-optimisées au moyen de la mécanique moléculaire, champ de force (MM+) en utilisant le logiciel HyperChem 8.07 [3].

Ensuite, ces dérivées ont été réoptimisées en utilisant la méthode PM3 et DFT avec une fonctionnelle hybride B3LYP [4] de logiciel Gaussien 09[5].

Le module "propriétés QSAR" de l'HyperChem 8,07 [3] a été utilisé pour calculer les paramètres suivants : le coefficient de partage octanol/eau (logP), le volume moléculaire (V), la surface moléculaire(SAG) et mass, réfractivité molaire (MR), (tableau III.2).

Les descripteurs électroniques: le moment dipolaire (MD), l'énergie des orbitales frontières (E_{HOMO} , E_{LUMO}), les charges (qO), sont calculés par la méthode DFT (B3LYP /3-21G (d, p) en utilisant logiciel Gaussien 09 et GaussView (5.0) et en utilise le molinspiration [6] pour calculé la liaison hydrogène donneurs (HBD) et la liaison hydrogène accepteurs (HBA). (tableau III.3).

Composé	Surface area grid (Å ²)	Volume (Å ³)	Hydration energy (Kcal/mol)	Log p	Refractivity (Å ³)	Polarizability (Å ³)	Mass (uma)
1 ^a	558.83	971.83	-27.61	-5.69	103.60	37.85	422.35
2 ^a	719.32	1295.03	-30.79	-7.15	136.04	51.27	584.49
3 ^a	535.56	947.36	-32.01	-5.69	103.60	37.85	422.35
4 ^a	569.95	981.75	-30.79	-4.98	103.68	37.85	422.35
5 ^a	380.20	610.72	-17.09	-1.47	68.03	23.15	228.20
6 ^a	399.41	649.66	-27.35	-3.52	71.24	24.43	260.20
7 ^a	394.72	643.66	-27.09	-3.52	71.24	24.43	260.20
8 ^a	455.25	751.07	-23.61	-4.48	82.39	28.73	304.26
9 ^a	570.67	994.51	-17.45	-4.92	113.21	41.52	450.40
10 ^a	642.37	1124.6	-15.02	-1.02	122.98	44.23	410.47

		5					
11 ^a	669.14	1173.2	-11.12	-0.99	127.75	46.06	424.49
		2					
12 ^a	631.25	1095.3	-18.84	-1.05	118.21	42.39	396.44
		5					
13 ^a	543.89	912.49	-10.92	-1.23	97.89	34.61	326.35
14	551.67	930.23	-14.67	-2.25	99.50	35.24	342.35
15	634.54	1089.5	-10.62	-0.02	116.61	41.76	380.44
		5					
16	369.84	590.14	-10.59	-0.45	66.43	22.51	212.20
17	377.73	607.91	-14.30	-1.47	68.03	23.15	228.20
18	381.37	611.36	-17.29	-1.47	68.03	23.15	228.20
19	387.98	628.47	-20.79	-2.50	69.64	23.79	244.20
20	597.35	1052.5	-20.97	-5.95	114.82	42.15	466.40
		2					
21	578.93	1004.5	-22.29	-4.95	108.45	39.68	436.37
		3					
22	597.27	1052.4	-20.59	-4.92	113.21	41.52	450.40
		2					
23	544.14	966.42	-14.81	-4.89	117.98	43.35	464.43
24	444.97	750.24	-20.91	-4.48	82.39	28.73	304.26
25	444.22	732.29	-17.01	-3.46	80.78	28.10	288.26
26	451.15	746.75	-16.55	-3.42	85.55	29.93	302.28
27	421.59	692.23	-21.99	-3.49	76.01	26.26	274.23
28	387.54	633.08	-25.05	-3.52	71.24	24.43	260.20
29	464.01	780.06	-16.89	-4.45	87.16	30.57	318.28
30	469.56	794.79	-15.25	-4.42	91.93	32.40	332.31
31	417.46	683.72	-22.11	-3.49	76.01	26.26	274.23
32	417.54	680.44	-18.72	-2.46	74.41	25.62	258.23
33	390.68	631.87	-23.83	-2.50	69.64	23.79	244.20
34	424.38	698.18	-21.65	-3.46	80.78	28.10	288.26
35	439.94	742.13	-18.08	-4.48	82.39	28.73	304.26
36	493.96	839.01	-11.22	-4.42	91.93	32.40	332.31
37	479.42	794.24	-12.72	-3.42	85.55	29.93	302.28
38	557.31	986.63	-17.55	-5.95	114.82	42.15	466.40
39	638.92	1103.6	-16.42	-0.81	118.53	42.39	396.44
		3					

Tableau III.2 : Valeurs des descripteurs physicochimique utilisés dans l'étude QSAR.

Comp ose	Momen t dipolair e (Debyes)	La charge O	E _{HOMO}	E _{LUMO}	ΔE	Energie stérique	H BA	H BD
1 ^a	1.0438	- 0.639485	- 0.08709	- 0.05090	0.03619	30.882992	11	8
2 ^a	4.7293	- 0.641519	- 0.08607	- 0.03922	0.04685	51.288765	16	11
3 ^a	6.3944	- 0.635897	- 0.10918	- 0.05143	0.05775	29.464569	11	8
4 ^a	10.8590	- 0.630025	- 0.11807	- 0.03742	0.08065	33.879459	11	7
5 ^a	4.7473	- 0.632543	- 0.22077	- 0.06629	0.15448	15.447466	4	2
6 ^a	6.0991	- 0.637150	- 0.21436	- 0.04660	0.16776	17.564993	6	4
7 ^a	4.4754	- 0.635739	- 0.21249	- 0.05722	0.15527	16.821163	6	4
8 ^a	4.8543	- 0.638276	- 0.20508	- 0.04947	0.15561	27.577343	7	3
9 ^a	8.8823	- 0.645497	- 0.06463	- 0.00604	0.05859	41.028309	11	5
10 ^a	3.1395	- 0.641261	- 0.19882	- 0.05375	0.14507	27.753252	6	3
11 ^a	2.9184	- 0.630115	- 0.20827	- 0.05401	0.15426	31.726051	6	2
12 ^a	2.4335	- 0.640285	- 0.20405	- 0.05591	0.14814	31.752697	6	4
13 ^a	5.7175	- 0.624961	- 0.20433	- 0.05973	0.1446	19.584146	5	2
14	6.8601	- 0.627206	- 0.19763	- 0.04929	0.14834	21.961304	6	3
15	6.5313	- 0.633964	- 0.19984	- 0.05782	0.14202	16.594479	5	3
16	3.4788	- 0.628319	- 0.21832	- 0.07136	0.14696	15.143076	3	1
17	4.6979	- 0.630723	- 0.21304	- 0.05838	0.15466	17.033369	4	2
18	3.7073	- 0.630981	- 0.21665	- 0.06558	0.15107	15.366761	4	2
19	5.7673	- 0.634915	- 0.21554	- 0.05334	0.1622	18.485500	5	3
20	7.1668	- 0.676956	- 0.11782	- 0.04724	0.07058	46.112103	12	6
21	4.2844	- 0.627679	- 0.20530	- 0.04614	0.15916	41.758034	11	3
22	6.2046	- 0.623679	- 0.20295	- 0.05856	0.14439	38.006397	11	5
23	6.9199	- 0.607353	- 0.22312	- 0.04892	0.1742	47.855537	11	4
24	3.9124	- 0.625714	- 0.19794	- 0.05270	0.14524	26.480654	7	3
25	4.6681	-	-	-	0.15269	28.699512	6	2

		0.626671	0.20399	0.05130				
26	5.9586	-	-	-	0.1495	27.158718	6	1
		0.626110	0.20752	0.05802				
27	4.0323	-	-	-	0.15218	21.691162	6	3
		0.626364	0.20730	0.05512				
28	6.6427	-	-	-	0.15214	18.886747	6	4
		0.616369	0.20432	0.05218				
29	0.2996	-	-	-	0.14087	34.553406	7	2
		0.622667	0.20228	0.06141				
30	4.8554	-	-	-	0.14087	32.348942	7	1
		0.627192	0.20954	0.06435				
31	4.9835	-	-	-	0.15356	24.712667	6	3
		0.627666	0.20568	0.05212				
32	3.2320	-	-	-	0.14977	19.646912	5	2
		0.623767	0.21219	0.06242				
33	3.2883	-	-	-	0.1509	15.644538	5	3
		0.624648	0.21420	0.06330				
34	4.4381	-	-	-	0.15671	25.859095	6	2
		0.626537	0.21786	0.06115				
35	6.5673	-	-	-	0.1485	25.235357	7	3
		0.616051	0.20035	0.05185				
36	4.7554	-	-	-	0.15239	33.718628	7	1
		0.626531	0.20604	0.05365				
37	6.2200	-	-	-	0.153	28.431934	6	1
		0.614544	0.20186	0.04886				
38	7.0227	-	-	-	0.15388	40.766979	12	6
		0.594282	0.20889	0.05501				
39	6.2398	-	-	-	0.15897	28.658985	6	4
		0.665428	0.20190	0.04293				

Tableau III.3 : Valeurs des descripteurs quantique utilisés dans l'étude QSAR

III.1.3. Développement des modèles QSAR

Le développement d'un modèle QSAR exige un ensemble des diverses données, pour cet effet, un nombre considérable de descripteurs doivent être considérés dans l'étude QSAR. Les descripteurs sont des valeurs numériques qui codent les différentes caractéristiques structurales des molécules. La matrice de corrélation de Pearson a été effectuée sur tous les descripteurs en utilisant le logiciel SPSS19 [7]. L'analyse de la matrice révéla 15 descripteurs pour le développement du modèle QSAR. Les valeurs des descripteurs sélectionnés pour le modèle QSAR sont présentées dans le tableau III.2 et tableau III.3.

Les corrélations entre l'activité biologique et les descripteurs utilisés sont exprimées par les relations mathématiques suivantes :

$$\text{pIC50} = 1.137 - 0.053\text{SAG} + 0.032 \text{V} - 0.063 \text{HE} - 0.299 \log \text{P} - 0.078 \text{MD} + 23.260 \Delta\text{E} - 0.58\text{HBA}$$

N= 26

R= 0.863

R²= 0.744

S=0.363102

F= 8.148

N : nombre de composés.

R : coefficient de corrélation.

R²: coefficient de détermination.

S : erreur standard de l'estimation.

F: statistique Fischer.

- 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. [8]
- La valeur $R = 0.863$ ($R^2 = 0.744$) (modèle), nous permis d'indiquer fermement la corrélation entre les différents descripteurs utilisés et l'activité biologique dans (PIC50).
- La valeur F est jugé statistiquement significative au niveau de 95%, étant donné que la valeur F calculée est supérieure par rapport aux valeurs du tableau critiques [9, 10].
- Dans l'équation, le coefficient positif de V montrer que toute augmentation dans les valeurs de ce paramètre entraine une augmentation de l'activité xanthone.
- Dans l'équation, les coefficients positifs ΔE , montrent que toute augmentation dans les valeurs de ces paramètres entraine une augmentation de l'activité xanthone.
- Le coefficient de SAG a un signe négatif dans le modèle, ce qui suggère diminuant d'activité, En effet, toute de diminuant paramètre SAG conduit à diminuant de surface du contact des molécules avec le récepteur biologique.
- Le coefficient du La moment dipolaire a un signe négatif dans le modèle, ce qui suggère qu'une activité accrue peut être obtenue en diminuant la polarité des dérivés xanthone.
- Les paramètres HBA sont responsables des interactions électrostatiques des dérivés de xanthone dans les sites actifs de récepteur biologique.
- le coefficient négatif de l'énergie d'hydratation (HE) et log P explique que toute diminution de lipophile des molécules provoque une diminution de l'activité biologique.

Les matrices de corrélation entre les descripteurs obtenus par l'analyse MLR et l'activité biologique.III.4.

	hba	MD	HE	ΔE	SAG	Log P	Vol
HBA	1						
MD	-0.318	1					
HE	0.315	-0.132	1				
ΔE	0.005	-0.467	-0.046	1			
SAG	0.521	-0.121	0.189	-0.179	1		
Log P	0.852	-0.256	0.099	0.096	0.142	1	
Vol	-0.587	0.147	-0.221	0.163	-0.996	-0.214	1

Tableau III.4 : matrice de corrélation

Le tableau III.4 montre que les descripteurs obtenus sont corrélés entre eux et avec l'activité biologique. Il et montre que le descripteur (Log P) est le paramètre le plus important (85.2%) dans la corrélation entre les descripteurs sélectionnés et l'activité biologique des dérivés xanthone.

III.1.4.Validation de modèle QSAR

Les puissances prédictives des équations ont été validées par la méthode de validation croisée "leave-one-out" (LOO) [11 ,12]. 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_{adj}^2 , coefficient de corrélation de validation croisée R_{CV}^2 et la validation standard des erreurs de prédiction (**S_{PRESS}**) et l'erreur de prédiction (**PE**) (Tableau III.5)

Modèle	PRESS	SSY	PRESS/SSY	S_{PRESS}	R_{CV}^2	R_{adj}^2	6PE
1	3.955	12.549	0.315	0.101	0.685	0.601	0.203

Tableau III.5 : Paramètres de validation croisée

- La somme des carrée résiduels (PRESS, Prédictive Residual Sum of Squares) est une mesure des capacités prédictives d'un modèle. La petite valeur de PRESS signifie la meilleure prévisibilité du modèle [13].
- En outre, pour un modèle raisonnable de QSAR, le rapport PRESS/TSS doit être inférieur à 0.4 [8].
- Les résultats présents dans le tableau III.5 indiquent que ce rapport est 0.315 pour le modèle. Les valeurs élevées du R_{CV}^2 et R_{adj}^2 sont des critères essentiels pour une meilleure qualification des modèle QSAR. Les résultats obtenus de R_{CV}^2 et R_{adj}^2 pour le modèle QSAR [14] est 0.685 et 0.601 Respectivement.
- On peut également utilise le paramètre S_{PRESS} qui rend compte des variations d'erreurs des prédictions .Le modèle QSAR développé présente une faible valeur de S_{PRESS} (0.101) ce qui indique que le modèle présente des petites valeurs résiduelles entre les activités biologiques observées et prédites.
- L'erreur de prédiction du coefficient de corrélation (PE) est un autre paramètre utilise pour déterminer le pouvoir prédictif du modèle propose. Nous avons calculé la valeur de $6PE$ du modèle obtenue et il est présenté dans le tableau III.6. Pour le modèle, Le condition $r > 6PE$ est satisfaite et donc ils peuvent être considères comme ayant un bon pouvoir prédictif [15].

Les activités biologique expérimentales, prédites et résiduelles des dérive xanthone sont représentés dans le tableau III.6. 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 [16].

Composé	$PIC_{50}(exp)$	$PIC_{50}Perd$	$PIC_{50}Resd$
1	-1.931	-1.71960	0.217173
2	-2.199	-2.27215	0.073491
3	-1.886	-1.36743	-0.518952
4	-2.832	-2.39496	-0.436915
5	-2.206	-1.91348	-0.292803
6	-1.233	-1.26000	0.027000
7	-0.851	-1.34171	0.491067
8	-1.648	-1.52609	-0.122270
9	-1.943	-2.11878	0.176276
10	-0.699	-0.58789	-0.111079

11	-1.158	-.69187	-0.466489
12	-0.176	-.66021	0.484121
13	-1.767	-2.03221	0.265055
14	-1.703	-1.78230	0.079009
15	-1.332	-1.63917	0.306731
16	-2.249	-2.17499	-0.073968
17	-1.961	-1.87633	-0.085092
18	-2.119	-1.85353	-0.265064
19	-1.913	-1.59227	-0.320487
20	-2.126	-2.45498	0.328847
21	-2.596	-2.19441	-0.402081
22	-2.061	-2.21033	0.149629
23	-1.923	-2.27971	0.356462
24	-.734	-1.19453	0.460535
25	-2.033	-1.70785	-0.324768
26	-2.153	-1.90795	-0.244643
27	-1.486	-1.48596	0.000236
28	-0.724	-1.49500	0.770723
29	-1.889	-1.43211	-0.456627
30	-1.818	-1.88220	0.063973
31	-1.852	-1.53721	-0.315273
32	-1.246	-1.67553	0.430013
33	-1.502	-1.47637	-0.026052
34	-1.255	-1.62509	0.369814
35	-1.915	-1.54622	-0.368656
36	-1.603	-1.67955	0.076404
37	-1.880	-1.96351	0.083843
38	-2.762	-2.60828	-0.153647
39	-1.255	-1.02974	-0.225537

Tableau III.6: Les valeurs expérimentales, prédites et résiduelles.

La figure III.2 ci-dessous montre la courbe de corrélation obtenue par MLR entre les valeurs prédites par rapport aux les valeurs expérimentales de l'activité inhibitrice PIC50 des dérivées xanthone. La courbe obtenue pour le modèle présente un coefficient de corrélation $R^2(0.744)$ de valeur significative. Il indique que ce modèle, peut être appliqué avec succès pour prédire l'activité inhibitrice de cette série des molécules.

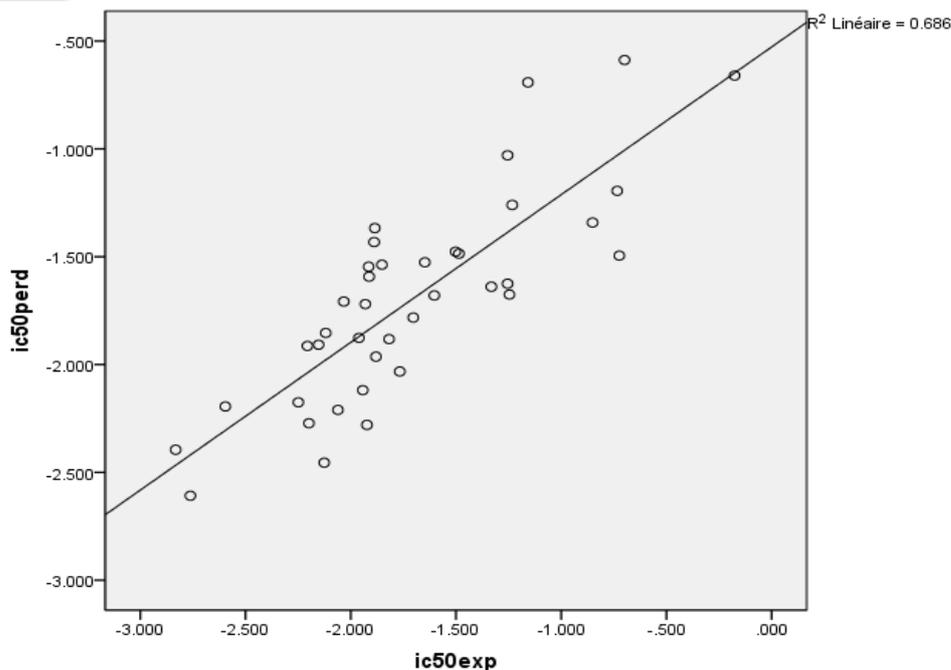


Figure III.2 : corrélation entre les activités biologiques expérimentales et prédites pour le modèle.

Dans l'élaboration de modèle QSAR, les valeurs résiduelles de l'activité biologique ont été tracées en fonction des valeurs expérimentales pour trouver d'une erreur systématique comme montre la Figure (III.3). La répartition des résidus dans les deux côtés de l'axe des abscisses indique l'absence de l'erreur systématique [17]. Il indique que ce modèle peut être appliqué avec succès pour prédire l'activité biologique des dérivés xanthone pour la série de molécules utilisée dans le développement du modèle QSAR.

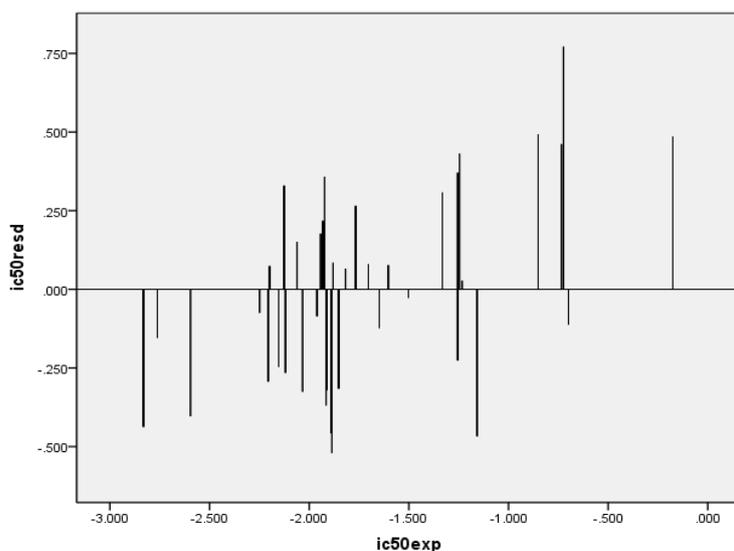


Figure III.3: Représentations graphique des valeurs résiduelles en fonction des valeurs expérimentales pour le modèle.

Références bibliographiques

- [1] G. Fayet, Développement de modèles QSAR pour la prédiction des propriétés d'explosibilité des composés nitroaromatiques, Thèse de doctorat de l'université pierre et marie curie, Paris VI, 2010
- [2] H. Gonzalez-Diaz, E. Olazabal, L. Santana, E. Uriarte, Y. Gonzalez-Diaz, N. Castanedo, *Bioorg. Med. Chem.*, 15, 962 (2007)
- [3] HyperChem (Molecular Modeling System) Hypercube, Inc., 1115 NW, 4th Street, Gainesville, FL 32601, USA, (2008).
- [4] F. Jensen; J. Wiley, *Introduction to Computational Chemistry*, Chichester, Angleterre, (1999).
- [5] 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; J.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. D. J. Fox, Gaussian, Inc., Wallingford CT, (2009).
- [6] www.molinspiration.com/cgi-bin/properties.
- [7] H.A. Carlson. *Curr. Op. Chem. Biol. S*, 447-452, 2002.
- [8] K. Ajeet, Bijander, *Int. J. Pharm. & Pharm. Sci.*, 4, 445, (2012).
- [9] Z. Almi, et al. "Structure Activity Relationships, QSAR Modeling and Drug-like calculations of TP inhibition of 1, 3, 4-oxadiazoline-2-thione Derivatives." *International Letters of Chemistry, Physics and Astronomy*, vol.18, pp. 113-124, 2014.
- [10] K. Dermeche, et al. "Qualitative Structure-Activity Relationships and 2D-QSAR Modeling of TNF- α Inhibition by Thalidomide Derivatives." *Journal of Bionanoscience*, vol.9.no.5, pp.395-400, 2015.
- [11] M. Clark, R.D. Cramer. The probability of chance correlation using partial least squares (PLS). *Quant. Struct.-Act. Relat*, 12, 45-137, 1993.
- [12] H. Van der Voet, Comparing the predictive accuracy of models using a simple

randomization test, *Chemometr. Intell.sys.Lab*, 25, 23-313, 1994.

[13] P.Kuzmanović,O.Sanja,D.Dragoljub , Cvetković, and D. Barna. "QSAR analysis of 2- amino or 2-methyl-1-substituted benzimidazoles against *Pseudomonas aeruginosa*." *International journal of molecular sciences* vol.10, pp.1670-1682,2009.

[14]. S.O. Podunavac-Kuzmanović, D.D. Cvetković, D.J. Barna, *Int. J. Mol. Sci.*, 10, 1670, (2009).

[15] A. K.Srivastava,and N. Shukla. "Quantitative structure activity relationship (QSAR) studies on a series of imidazole derivatives as novel ORL1 receptor antagonists."*Journal of Saudi Chemical Society* ,vol.17 , pp.321-328,2013.

[16] C.Matthew , and Richard D. Cramer,"The probability of chance correlation using partial least squares (PLS)." *Molecular Informatics* , vol.12 , pp. 137-145 ,1993.

[17] R.Tabet , *Etude des propriétés QSAR d'une série des isoindoles à activité thérapeutique* , mémoire master , université Biskra , 2013.

***Conclusion
générale***

Conclusion générale

La modélisation moléculaire et l'analyse QSAR ont été effectuées pour déterminer quantitativement l'effet de la structure moléculaire des composés étudiés sur leur activité biologique. Des modèles mathématiques ont été développés pour prédire l'activité de dérivés xanthone.

La validité des modèles obtenus ont été établie par la détermination des paramètres statistiques appropriés. Le modèle mis en place a été utilisé pour prédire l'activité biologique étudiée. Les valeurs prédites sont plus proches que les valeurs expérimentales pour l'activité biologique.

Réalisé une étude quantitative de la relation structure-activité (QSAR) sur une série des dérivés de xanthone. Plusieurs descripteurs ont été utilisés dans le développement des modèles QSAR à savoir : log P, SAG, V, HBA, DM et les charges atomiques q O.

Les descripteurs utilisés dans cette étude ont été calculés par la méthode DFT avec la fonction B3LYP associée à un ensemble de base 3-21G.

Modèles QSAR ont été établis en utilisant la méthode de régression multilinéaire (MLR). Les modèles obtenus peuvent être utilisés pour prédire l'activité biologique des dérivés xanthone.

Le pouvoir prédictif du modèle QSAR obtenu a été confirmé par la méthode de validation croisée LOO (Leave-one-out). Les résultats obtenus de validation externe confirme la fiabilité de modèle prédictif obtenue.

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.

Annexe

Table du F (de Fischer)

Table de Fisher-Snedecor, $\alpha = 5\%$ (95^e centile)

v_2 (dén.)	v_1 (numérateur)																			
	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60	80	100	200	500	1 000
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	248.02	250.10	251.14	251.77	252.20	252.72	253.04	253.68	254.06	254.19
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.45	19.46	19.47	19.48	19.48	19.49	19.49	19.49	19.49	19.49
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.66	8.62	8.59	8.58	8.57	8.56	8.55	8.54	8.53	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.80	5.75	5.72	5.70	5.69	5.67	5.66	5.65	5.64	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.56	4.50	4.46	4.44	4.43	4.41	4.41	4.39	4.37	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.87	3.81	3.77	3.75	3.74	3.72	3.71	3.69	3.68	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.44	3.38	3.34	3.32	3.30	3.29	3.27	3.25	3.24	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.15	3.08	3.04	3.02	3.01	2.99	2.97	2.95	2.94	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	2.94	2.86	2.83	2.80	2.79	2.77	2.76	2.73	2.72	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.77	2.70	2.66	2.64	2.62	2.60	2.59	2.56	2.55	2.54
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.12	2.04	1.99	1.97	1.95	1.92	1.91	1.88	1.86	1.85
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	1.93	1.84	1.79	1.76	1.74	1.71	1.70	1.66	1.64	1.63
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	1.84	1.74	1.69	1.66	1.64	1.61	1.59	1.55	1.53	1.52
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.78	1.69	1.63	1.60	1.58	1.54	1.52	1.48	1.46	1.45
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.75	1.65	1.59	1.56	1.53	1.50	1.48	1.44	1.41	1.40
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.72	1.62	1.57	1.53	1.50	1.47	1.45	1.40	1.37	1.36
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.70	1.60	1.54	1.51	1.48	1.45	1.43	1.38	1.35	1.34
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.69	1.59	1.53	1.49	1.46	1.43	1.41	1.36	1.33	1.31
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.68	1.57	1.52	1.48	1.45	1.41	1.39	1.34	1.31	1.30
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.62	1.52	1.46	1.41	1.39	1.35	1.32	1.26	1.22	1.21
300	3.87	3.03	2.63	2.40	2.24	2.13	2.04	1.97	1.91	1.86	1.61	1.50	1.43	1.39	1.36	1.32	1.30	1.23	1.19	1.17
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.59	1.48	1.42	1.38	1.35	1.30	1.28	1.21	1.16	1.14
1 000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.84	1.58	1.47	1.41	1.36	1.33	1.29	1.26	1.19	1.13	1.11
2 000	3.85	3.00	2.61	2.38	2.22	2.10	2.01	1.94	1.88	1.84	1.58	1.46	1.40	1.36	1.32	1.28	1.25	1.18	1.12	1.09

Résumé :

Une étude QSAR a été effectuée sur trent-neuf (39) molécules de la régression multilinéaire (MLR) a été utilisée pour quantifier relations entre les descripteurs moléculaires et les propriétés de chimioprévention des dérivés xanthone. Les descripteurs utilisés ont été calculés par la méthode DFT en utilisant la base 6.31G. Le pouvoir prédictif de modèle obtenus a été confirmé par la méthode de Validation -Croisée (LOO). Une forte corrélation a été observée entre les valeurs expérimentales et prédites pour l'activité biologique.

Mot Clés: xanthone , DFT, LOO, QSAR, MLR.

Abstract:

A QSAR study was carried out on 39 similar molecules of XANTHONE .The multiple linear regression (MLR) was used to quantify the relations between the molecular descriptors and the properties of chemo-prevention of . The descriptors used were calculated by xanthone method by using the base 6.31G. The predictive capacity of the model obtained was confirmed by the method of Crosse Validation (LOO). A strong correlation was observed between the experimental and predicted values for biological activity inhibitors.

Keywords: xanthone, DFT, LOO, QSAR, MLR.