UNIVERSITE YAHIAFARES DE MEDEA
FACULTE DES SCIENCES
DEPARTEMENT DES SCIENCES DE LA MATIERE

Mémoire

Présenté pour obtenir le diplôme de **:**

Master En Chimie

Spécialité : CHIMIE MACROMOLECULAIRE

Par :
MAAMOUR Amina
HACENI Sakina

Thème :

Développement d'un modèle QSAR pour prédire l'indice de réfraction des polymères (IR)

Soutenu publiquement/ le :23 /06 /2022, devant le jury composé de :

GHARDAOUI Dahbia	Grade :MCB	Université de Médéa	Président
EULDJI Amel	Grade :MAA	Université de Médéa	Examinateur
BITAM Said	Grade :MCA	Université de Médéa	Encadreur

Année universitaire : 2021/2022

Remerciements

Avant tout je remercie Allah qui m'aide ma donné la patience et le courage durant ces langues années d'étude.

Mes vifs remerciements vont également à mon encadreur Dr. Bitam Said m'avoir confié ce sujet et dirigé ces travaux pour sa disponibilité sa gentillesse, ces conseils précieux, son apport scientifique et la confiance totale qu'il m'a accordée durant ces années.

Je remercie toute ma famille et toutes mes amies qui m'ont apporté leur aide au cours de mes études.

Je ne serais oublié dans ces remerciements la direction de la Faculté de Chimie, le personnel technique et administratif, ingénieur de labo informatique ainsi que tous ceux qui, de près ou de loin ont contribué à la réalisation de ce travail.

Dédicaces

Je dédie ce Modeste travail à :

A mes parents

A mes cher frères Walid et Mohamed et mes soeurs Nour el houda et son mari Aziz et Assia.

A tous ma famille Maamour et Houari.

A tous mes meilleurs amis : Hiba. Houda. Fati. Sakina. Mohamed

A tous mes amis.



Amina

Remerciements

Dieu merci pour m'avoir donné la santé, la volonté et le courage sans lesquels ce travail n'aurait pas été réalisé.

Mes remerciement à mon encadreur Dr. Bitam Said pour sa gentillesse, pour ses conseils et de m'avoir guidée pas à pas dans mon travail.

Merci a tout ma famille, mes parents, mes soeurs.

Merci à mon cher Mari et tout les personnes qui m'ont encouragé et soutenu de prés ou de loin durant la réalisation de ce travail.

Dédicaces

Je remercie DIEU de m'avoir donné le courage pour accomplir ce modeste travail.

Je dite se modeste travail a deux personnes qui ont sacrifiées pour que je grandisse avec un savoir faire et qui ont fait de moi ce que je suis aujourd'hui, sans lesquelles je n'y serais jamais parvenue Mon Père et Ma Mère.

A ma grand-mère : Asma

A mes chère soeurs : Nabila. Amira. Rania. Linda. Assia je souhaite son succès dans le BEM

A mes frères : Amir. Mohamed je lui souhaite succès dans la Cinquième année.

A mon cher Mari ; Oussama, et sa mère, et sa soeur Rima.

Aux petites beaux enfants : Nour. Djawad. Bilal.

A tous ma famille Haceni et Ben hecin et Boutalba.

Je souhaite un prompt rétablissement à la femme de mon oncle et de ma tante

A mes amis : Faiza. Amira. Hayet. Sabrina. Yassmine. Houda. Nessrine .fariha . Bahdja .Om alkhair .Ahlam.

A tout le groupe de cette Promotion.

A tous mes amis.

Sakina

Rrésumé

Dans cette étude, deux modèles QSAR ; linéaires et non linéaires ont été développés pour prédire les valeurs de l'indice de réfraction de 53 polymères en fonction de leur structure moléculaire, représentée par cinq descripteurs pertinents. Les deux modèles développés ont montré une bonne capacité prédictive. Pour sélectionner le meilleur modèle, le paramètre R^2 a été utilisé. Le meilleur modèle a été celui établi avec le réseau de neurones artificiel (RNA). Les deux modèles développés dans ce travail ont montré l'importance de la forme, le nombre d'hydrogènes et la réfraction molaire dans la prédiction de l'indice de réfraction.Les modèles développes sont utiles pour améliorer les propriétés des polymères.

Mots clés : QSAR ; RLM ; RNA ; Indice de réfraction ; Polymères

Abstract

In thisstudy, two QSAR models;Linear and nonlinearmodels have been developed to predict the refractive index values of 53 polymers based on their molecular structure, represented by five relevant descriptors. The two developed models showed a good predictive capacity. To select the best model, the R2 parameterwasused. The best model was the one est a blishedwith the artificial neural network (ANN). The twomodelsdeveloped in this work showed the importance of the shape, the number of hydrogens and the molar refraction in the prediction of the refractive index. The models developed are useful for improving the properties of polymers.

Keywords :QSAR ; MLR ; RNA ; Refractive index ; Polymers

الملخص

في هذه الدراسة، تم تطوير نموذجان منQSAR ؛ خطية وغير خطية للتنبؤ بقيم معامل الانكسار لـ 53 بوليمر بناءً على هيكلها الجزيئي، ممثلة بخمسة واصفات ذات صلة. أظهر النموذجان المطوران قدرة تنبؤية جيدة. لتحديد أفضل نموذج، تم استخدام المعامل .R2 كان أفضل نموذج هو ذلك الذي تم إنشاؤه باستخدام الشبكة العصبية الاصطناعية .(ANN) أظهر النموذجان اللذان تم تطويرهما في هذا العمل أهمية الشكل وعدد الهيدروجين والانكسار المولي في التنبؤ بمعامل الانكسار. .النماذج المطورة مفيدة في تحسين خصائص البوليمرات.

الكلمات المفتاحية: QSAR؛MLR. ANN. معامل الانكسار؛ البوليمرات

SOMMAIRE

Introduction Générale

Chapitre I :Rappel bibliographique

Introduction 1
I.1 Historique des polymères 1
I.2 Définition des polymères 2
I.3 Différentes classes des polymères 2
I. 3.1 Classification selon l'origine 3
I.3.2 Classification selon l'architecture 3
I.3.3 Classification selon la structure chimique 4
I.3.3.1 Un homopolymère 4
I.3.3.2 Un copolymère 5
I.3.4 Classification selon le comportement thermique 6
I.3.5 Classification selon le comportement physique 7
I.4 Propriétés des polymères 9
I.4.1 Les propriétés électriques et électro-optiques 9
I.4.2 Propriétés thermiques 10
I.4.3 Propriétés mécaniques 11
I.4.4 Propriétés physiques 11
I.4.5 Propriétés chimiques des polymères 11
I.5 Les applications des polymères 11
I.6 Les méthodes de détermination de leurs propriétés 13
I.6.1 Spectroscopie infrarouge IR 13
I.6.2 Etude thermique ATD/ATG 15
I.6.2.2 L'analyse thermique différentielle ATD 15
Références 17

Chapitre II : Méthodes de Calcul

II.1 La modélisation moléculaire 19
II.2 Les méthodes de calcul 20
II.2.1 Les méthodes empiriques 20
II.2.2 Les méthodes semi-empiriques 20
II.2.3 Les méthodes quantiques 21
II.2.3.1 Théorie de la fonctionnelle de la densité (DFT) 21
II.2.3.2 Ab-initio 21
II.2.4 Les méthodes QSAR (relation quantitative structure-activité) 22
II.3 Modèle QSAR / QSPR 22
II.3.1 Introduction 22
II.3.2 Principe de QSAR/QSPR 23
II.3.3 Outils et Méthodologie de QSAR/QSPR 23
II.3.3.1 Activité biologiques /propriété physique 23
II.3.3.2 Les descripteurs moléculaires 24

II.3.4 Les méthodes statistiques 25
II.3.4.1 Régression linéaire simple (RLS) 25
II.3.4.2 Régression linéaire multiple (RLM) 25
II.3.4.3 L'analyse des moindre carrée partiels (PLS 26
II.3.4.5 Réseaux de Neurones Artificielle (RNA) 26
II.3.5 Test de la signification totale de la régression 28
II.3.5.1 Coefficient de détermination (R^2) 28
II.4.5.2 Coefficient de corrélation (r) 28
II.3.5.3 Ecart type (s) 29
II.3.5.4 Le coefficient de détermination ajusté (R_{adj}^2) 29
II.3.5.5. Coefficient de Prédiction (Q^2) 29
II.3.6 Validation du modèle 29
II.3.6.1 Validation Interne 29
II.3.6.2 Validation externe 30
II.3.7 Les applications de l'étude QSAR 31
Références 32
Chapitre III : Application et interprétation des résultats
III.1 Préparation du modèle 34
III.1.1 Construction de la base de données 34
III.1.2 Calcul des descripteurs 38
III.2 Le modèle RLM 38
III.3 Le modèle RNA 42
III .4. Conclusion 45
Références 46
Conclusion Générale 48

LISTE DES FIGURES

Chapitre I

Figure I.1. Modèle de la synthèse d'un polymère 2
Figure I.2. Classification des architectures macromoléculaires 4
Figure I.3. Homopolymère linéaire 4
Figure I.4. Homopolymère branché 5
Figure I.5. Homopolymère étoilé 5
Figure I.6. Copolymère statistique 5
Figure I.7. Copolymère alterné 6
Figure I.8. Copolymère séquence 6
Figure I.9. Copolymère greffé 6
Figure I.10. Principaux groupes des polymères 7
Figure I.11. Formation d'un polymère 8
Figure I.12. Quelque Applications des polymères 13
Figure I.13. Appareil infrarouge 14
Figure I.14. Appareil Couplage ATD/.ATG 15

Chapitre III

Figure II.1. Mise en correspondance neurone biologique / neurone artificiel 27

Figure II.2. Représentation d'un réseau de neurone 27

Chapitre III

Figure III.1. Graphique des valeurs de RI observées par rapport aux valeurs prédites du modèle RLM pour l'ensemble d'apprentissage 41

Figure III.2. L'architecture du réseau de neurones sélectionnés (MLP 5-5-1) 43

Figure III.3. Graphique des valeurs de RI observées par rapport aux valeurs prédites du modèle RNA pour l'ensemble d'apprentissage 44

Figure III.4. Graphique des valeurs de RI observées par rapport aux valeurs prédites du modèle RNA pour l'ensemble du test 45

LISTE DES TABLEAUX

Chapitre I

Tableau I.1. Quelques indices de réfraction pour les grandes familles de plastiques optiques 10
Tableau I.2. Les domaines d'application des polymères 12

Chapitre III

Tableau III.1. Structures chimique de polymères étudiées 34
Tableau III.2. Les meilleurs descripteurs sélectionnés 38
Tableau III.3. La signification des descripteurs utilisés 39
Tableau III.4. Les valeurs de VIF 39
Tableau III.5. Matrice des coefficients de corrélation des descripteurs sélectionnés 39
Tableau III.6. Les valeurs de RI obtenues par le modèle RLM 40
Tableau III.7. Les paramètres du meilleur réseau de neurones sélectionnés 42
Tableau III.8. Les valeurs RI (expérimentales et prédites) pour l'ensemble d'apprentissage et du test avec RNA 43

ABREVIATIONS

QSAR/QSPR	Relation quantitative structure-activité
QSPR	Relation quantitative structure-propriété
RI	Indice de réfraction
RLS	La régression linéaire simple
RLM	La régression linéaire multiple
RNA	Réseaux de neurones artificiels
R^2_{adj}	Le coefficient de détermination ajusté
R^2	Coefficient de détermination
VIF	Facteur de variation d'inflation

La chimie-informatique, également dénommée chemoinformatique.est le domaine de la science qui consiste en l'application de l'informatique aux problèmes relatifs à la chimie. Elle a pour objectif de fournir des outils et des méthodes pour l'analyse et le traitement des données issues des différents domaines de la chimie.

Le développement important des moyens informatiques a permis à la chimie de s'enrichir d'outils informatiques spécialement dédiés à la représentation des molécules et de mieux appréhender les problèmes liés au domaine moléculaire : la modélisation moléculaire.

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.

La modélisation QSAR (Quantitative Structure Activity Relationship) constitue actuellement un outil fiable et performant. C'est une méthode qui tente à identifier une relation quantitative entre l'activité mesurée expérimentalement et des descripteurs théoriques ou empiriques lies à la structure moléculaire.

L'indice de réfraction (RI) d'un matériau particulier est défini comme le rapport de la vitesse de la lumière dans le vide à la vitesse de la lumière dans cette substance. La fréquence de la lumière incidente et la polarisabilité des molécules éclairées sont les deux principaux facteurs qui déterminent comment la lumière sera déviée lors du passage à travers une interface entre deux milieux. En tant que tel, RI est une propriété fondamentale de la matière.

Dans ce contexte, l'objectif de ce travail est d'élaborer des modèles QSAR robustes et fiables, en respectant toutes les étapes d'une étude QSAR pour expliquer et prédire l'indice de réfraction d'une série de 53 polymères.

Le manuscrit de ce mémoire est articulé sur trois chapitres :

- Le premier chapitre présente un aperçu général sur les polymères.

- Le deuxième chapitre est consacré à une étude bibliographique sur les approches de base, les méthodologies, le développement, les techniques de validation et les applications des méthodes QSAR.

- Le troisième chapitre a été consacré à la génération des modèles QSAR par un ensemble de 53 polymères.

Enfin nous terminons ce travail par une conclusion générale récapitulative de notre travail.

Chapitre I :

Rappels Bibliographiques



Introduction

Les polymères sont utilisés depuis longtemps dans de nombreux domaines, notamment industriels, qu'ils soient d'origine naturelle ou préparés synthétiquement, en raison de leurs propriétés mécaniques (polyester) et isolantes (polyamide)

I.1Historique des polymères

Le terme polymère a été utilisé pour la première fois en 1832 par **Jöns Jacob Berzelius**(1779–1848). En 1835, le Français **Henri Victor Regnoult** (1810–1878) prépare du chlorure de vinyle à partir de chlore et d'éthylène. Il a remarqué que ce produit en ampoule scellée se transforme progressivement en PVC.

En 1863, l'imprimeur américain **John Wesley Hyatt** (1837–1920)et son frère ont été attirés par un concours de 10 000 $ pour trouver une substance pour remplacer l'ivoire pour la fabrication de billards. Après plusieurs années d'essais, les frères Hyatt inventent en 1870 le celluloïd, une matière plastique fabriquée à partir d'un mélange à chaud de nitrocellulose et de camphre (un plastifiant de la nitrocellulose). En 1897, apparaît la corne artificielle ou galalithe, un mélange caséine-formaldéhyde.
Puis le verre plastique ou PMMA en 1928 (une idée allemande d'**Otto Röhner** (1883-1960)), le PVC et le polystyrène (PS) en 1931, le polyéthylène (PE) en 1935 et le nylon en 1937.

Le développement des matériaux polymères n'est devenu considérable qu'après la fin de la Seconde Guerre mondiale, grâce aux travaux de :**Herman Standinger**(1881–1965), **Wallace Hume Carothers**(1896–1937), **Karl Waldemar Ziegler** (1898–1973) et **Giulio Natta** (1903–1979).

Vers le milieu des années 1960, le PVC a commencé à remplacer le verre dans les emballages alimentaires. C'est également à cette époque que la pétrochimie a remplacé la chimie du carbone dans la production de polymères.

En 1977, les professeurs **Alan Graham Mac Diarmid** (1927–2007), **Alan Jay Heeger** (né le 22 Janvier 1936) et **Hideki Shirakawa**(né le 20 Aout 1936) ont découvert que la conductivité électrique du polyacétylène pouvait être réglée par l'introduction contrôlée de groupes donneurs ou accepteurs d'électrons, transformant ainsi le polymère d'un état isolant à un état conducteur.

Aujourd'hui les matériaux polymères se trouvent dans tous les domaines, où ils se substituent aux matériaux usuels des céramistes et des métallurgistes**[1].**

I.2 Définition des polymères

Un polymère est une macromolécule obtenue par la répétition d'une unité constitutive, encore appelée unité de répétition, comportant un groupe d'atomes liés par des liaisons covalentes **[2].** Le terme polymère (du grec *polus* qui signifie « nombreux, plusieurs» et *méros* qui signifie « unité, partie») ; regroupe tout matériau formé par la répétition d'un très grand nombre n de petites molécules de faible masse moléculaire appelées monomères qui sont liées entre elles par des liaisons primaires (liaisons lentes) sachent que le monomère est une molécule de base (pouvant être par exemple non saturée ou cyclique ou encore comportant des fonctions réactives à ses extrémités)**[3].**

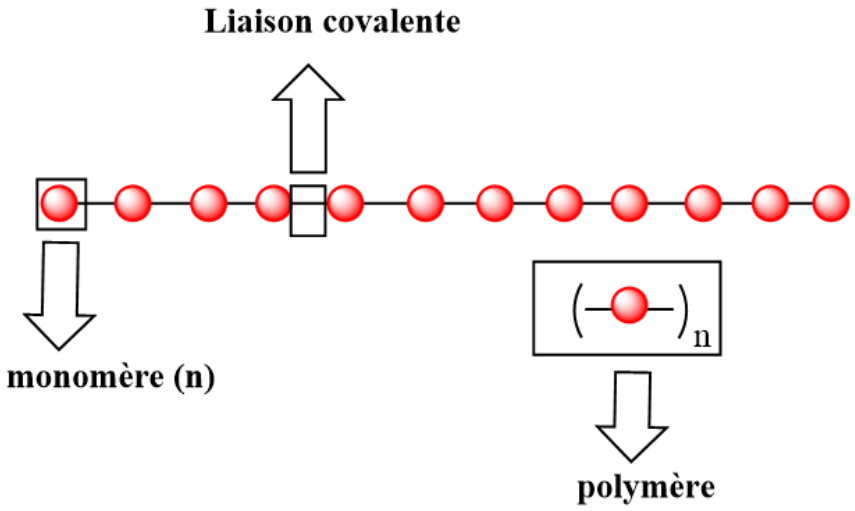


Figure I.1. Modèle de la synthèse d'un polymère **[4].**

La lettre *n* est appelée **« degré de polymérisation»** ou **« indice de polymérisation»**, pour des valeurs élevées de *n* (*n*> 100) **[5].**

I.3Différentes classes des polymères

Plusieurs classifications des composés macromoléculaires peuvent être proposées selon qu'on choisit l'origine, le type d'architecture, la structure chimique des motifs ou le types de polymérisation, comme base de la classification.

I. 3.1Classification selon l'origine

Les polymères classés selon l'origine peuvent être :

- **Des polymères naturels :** ce sont des composés organiques formant la matière vivante, comme les protéines, les acides nucléiques, la cellulose, la chitine…
- **Des polymères obtenus par modification chimique d'un polymère naturel :** exemple : méthylcellulose.
- **Des polymères synthétiques :** ce sont les matières plastiques, les élastomères, les fibres, les adhésifs.

I.3.2 Classification selon l'architecture

Selon Tomalia**[6],** les architectures macromoléculaires peuvent être divisées en quatre grandes familles : les polymères linéaires, ramifiés, réticulés et plus récemment, les polymères dendritiques.

- **Les polymères d'architecture linéaire :**Obtenu principalement par des techniques de polymérisation en chaîne (existantes depuis de nombreuses années, ou plus récemment contrôlées par radicaux libres) ou polymérisation par étapes ou polycondensation (entre monomères strictement bifonctionnels).
- **Les architectures ramifiées :**Produit par des réactions de transfert de chaîne ou par des réactions de greffage. Dans tous les cas, ces structures linéaires ou ramifiées définissent le domaine des thermoplastiques. Le nylon, le plexiglas (polymères linéaires), dont la production a débuté dans les années 1930, ou le polyéthylène basse densité (polymères ramifiés), dont la production a débuté dans les années 1960, sont autant d'exemples de thermoplastiques. L'introduction de liaisons covalentes entre chaînes de polymères linéaires ou ramifiés Aboutit à la troisième famille.
- **Les polymères réticulés :Paul Florey** (1910–1985) a été le premier à étudier ces nouvelles structures au début des années 1940. Elles définissent aujourd'hui le domaine communément appelé thermodurcissable. Les

époxydes et les caoutchoucs sont des exemples de systèmes polymères réticulés.

- **Les polymères dendritiques :** constituent la quatrième famille d'architecture macromoléculaire. La possibilité de synthétiser des dendrimères à partir de monomères multifonctionnels a été évoquée pour la première fois en 1953 par Paul Flory dans son célèbre livre « Principles of Polymer Chemistry »**[7]**. Mais ce n'est qu'à la fin des années 1970 que Vögtle décrit les premiers dendrimères synthétiques (qu'il appelle « arborols ») **[8]**. suivis de Tomalia**[6]** et de Newkome**[9-10].**

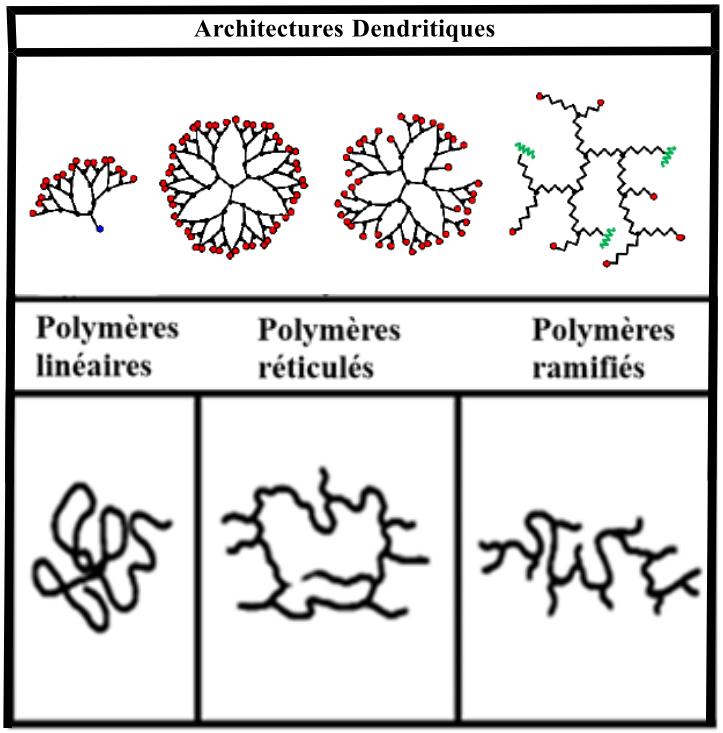


Figure I.2. Classification des architectures macromoléculaires

I.3.3 Classification selon la structure chimique

La structure chimique des motifs permet une classification des composés macromoléculaire sen homopolymères et copolymères.

I.3.3.1 Un homopolymère

Un polymère qui comporte des motifs monomères (fabriqué à partir d'un seul type de monomère) tous identiques, par exemples : Polyéthylène, polystyrène **[1,11].**

Il existe au sein des homopolymères différents familles, on trouve :

- **Les homopolymères linéaires :**

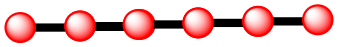


Figure I.3. Homopolymère linéaire **[12].**

- **Les homopolymères branchés :**

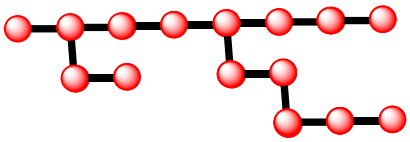


Figure I.4. Homopolymère branché **[12].**

- **Les homopolymères étoilés :**

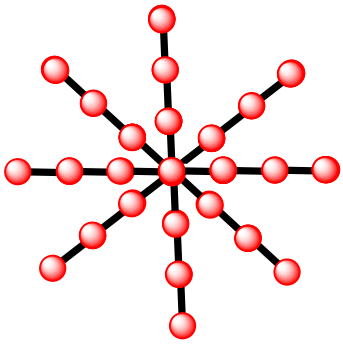


Figure I.5. Homopolymère étoilé **[12].**

I.3.3.2Un copolymère

Un polymère qui comporte des motifs monomères de deux ou plus sortes différentes (dès qu'au moins deux types de monomère participent à la formation des macromolécules) **[1].**

Il y'a quatre différents types de copolymère dans la figure suivant :

- **Le mode statistique :** les motifs de répétition sont répétais de manière aléatoire dans la chaine.

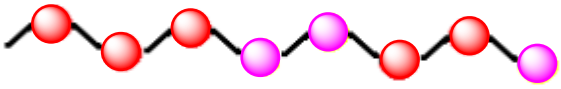


Figure I.6. Copolymère statistique **[12].**

- **Alterné :** les motifs de répétition sont toujours alternés.

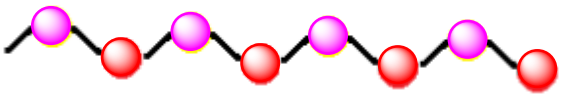


Figure I.7. Copolymère alterné **[12].**

- **Séquence (à blocs) :** les motifs différents sont regroupés en séquences. On a habituellement 2 ou 3 séquences de nature chimique différente.

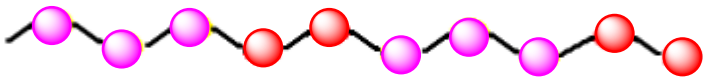


Figure I.8. Copolymère séquence**[12].**

- **Greffé :** fixation latérale de séquences (voilette) appelées « greffons » fixées sur une chaine principale formée de chainons (rouge) ou l'inverse **[12].**

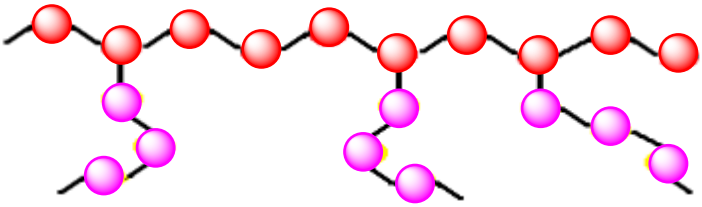


Figure I.9. Copolymère greffé **[12].**

I.3.4 Classification selon le comportement thermique

Les polymères sont souvent classés d'après leurs propriétés thermodynamiques en trois types :

- **Les thermoplastiques :**Adoucis par la chaleur, ils redeviennent souples, malléables et durcissent à nouveau en refroidissant, ces matériaux conservent

leurs propriétés et sont facilement recyclables. Leurs polymères de base sont composés de macromolécules linéaires reliées par des liaisons faibles qui se rompent lorsqu'elles sont exposées à la chaleur ou à de fortes contraintes. Elles peuvent alors glisser les unes par rapport aux autres pour prendre différentes formes. Lorsque le matériau refroidit, les liaisons se reforment et le thermoplastique conserve une nouvelle forme.

- **Les thermodurcissables :** les matières thermodurcissables sont les produits dont la transformation conduit, par une réaction chimique, à des composés macromoléculaires tridimensionnels qui sont des matières thermo durcies ou thermo rigides. Dans ce cas les chaînes de départs sont beaucoup plus courte set plus réactives, ces chaînes vont se lier ensemble chimiquement, cette réaction conduit à des liens chimiques rigides et met en jeu toutes les molécules présentes pour former un réseau tridimensionnel.
- **Les élastomères :** ces polymères présentent les mêmes qualités élastiques que le caoutchouc, un élastomère au repos est constitué de longues chaînes moléculaires repliées sur elles même, sous l'action d'une contrainte, les molécules peuvent glisser les unes par rapport aux autres et se déformer **[13].**

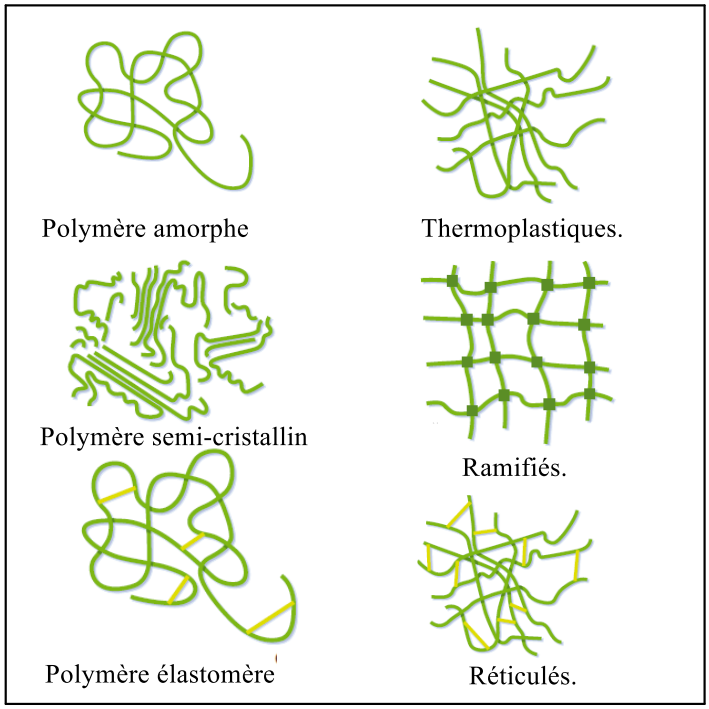


Figure I.10. Principaux groupes des polymères

I.3.5 Classification selon le comportement physique

En fonction de la structure physique des polymères, on peut les classer en quatre types d'état :

- **L'état amorphe,** L'arrangement moléculaire n'a pas d'ordre donné. Les polymères sont entrelacés pour former une boule. Ce manque d'ordre produit une absence de point de fusion, la présence d'une température de transition vitreuse qui indique un passage de l'état vitreux (le matériau se comporte comme du verre, c'est-à-dire dur et cassant) à l'état caoutchouteux (les chaînes peuvent glisser).) Plus clair) et transparence visible comme dans le cas du polystyrène "cristallin" ou poly (chlorure de vinyle) "cristallin". Les deux sont amorphes.
- **L'état amorphe orienté,** où les arrangements moléculaires sont tous orientés dans la même direction.
- **L'état cristallin,** caractérisé par un arrangement organisé des molécules. Les chaînes s'organisent de façon ordonnée et compacte. Le cristal polymère est décrit par son réseau et son motif. Le réseau caractérise la périodicité dans les trois directions de l'espace. Les principales caractéristiques de l'état cristallin sont une compacité supérieure à celle de la phase amorphe, l'existence d'un point de fusion et l'absence de transition vitreuse, une rigidité supérieure à celle de la phase amorphe.
- **L'état semi cristallin,** qui associe à la fois un arrangement organisé et des zones amorphes **[13].**

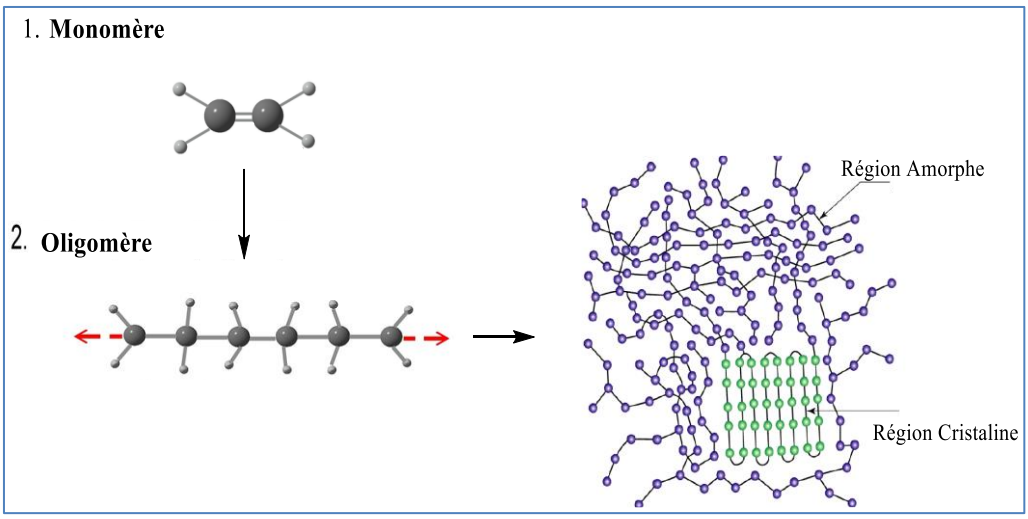


Figure I.11. Formation d'un polymère

I.4 Propriétés des polymères

Les polymères présentent plusieurs propriétés intéressantes pour cela leur domaine d'utilisation est très large et très varié.

I.4.1 Les propriétés optiques

Les polymères sont largement utilisés comme isolants électriques, en particulier dans les circuits électroniques et les gaines de câbles électriques. Il existe aussi des polymères conducteurs, soit à l'état intrinsèque, soit parce qu'ils sont chargés de particules de carbone conductrices. Certaines molécules de polymères, rigides et allongées, sont susceptibles de s'orienter sous l'effet d'un champ électrique ; cet effet est utilisé dans des dispositifs d'affichage comportant des polymères de structure cristal-liquide : lorsque le champ n'est pas appliqué, les molécules sont en désordre local, le matériau est opaque et d'apparence laiteuse, lorsque le champ est appliqué, les molécules s'orientent dans la même direction et laissent passer la lumière le matériau devient transparent**[14].**

- **Indice de réfraction**

L'indice de réfraction est une propriété importante lorsque des élastomères sont utilisés pour des applications optiques.

La faiblesse des plastiques « optiques » vient de la littérature inégale les concernant, souvent moins fournie que celle des verres optiques : les données sont de ce fait moins fournies, moins complètes et peuvent varier d'un fabricant à l'autre, voire entre deux polymères d'un même groupe. **[15]**

Comparativement aux verres, dont les indices sont compris dans une fourchette 1,28—1,95, les plastiques ont des indices plus restreints, de 1,31 à 1,65, leur constringence est par contre comparable, 91 à 20 pour les verres, 92 à 20 pour les plastiques.**[15]**

Tableau I.1. Quelques indices de réfraction pour les grandes familles de plastiques optiques **[15].**

Polymère	**Abréviation**	**λ (nm)**	**Indice**
Polytétrafluoroéthylène	PTFE	589,3 nm	1,31—1,32
Polycarbonate	PC	589,3 nm	1,5853—1,586
Polycarbonate	PC	365,0 nm	1,6432
Polycarbonate	PC	1014,0 nm	1,5672
Styrène-méthacrylate de méthyle	NAS	486,1 nm	1,574
Polystyrène co-méthacrylate de méthyle	SMMA	589,3 nm	1,564
Styrène-acrylonitrile	SAN	365,0 nm	1,6125
Styrène-acrylonitrile	SAN	589,3 nm	1,5673
Styrène-acrylonitrile	SAN	1014,0 nm	1,5519
Styrène-méthacrylate de méthyle	NAS	656,3 nm	1,558
Poly (méthacrylate de méthyle)	PMMA	589,3 nm	1,4917
Polyester	PETG	589,3 nm	1,567
Polyétherimide	PEI	589,3 nm	1,658
Polystyrène	PS	365,0 nm	1,6431
Polystyrène	PS	1014,0 nm	1,5726

I.4.2 Propriétés thermiques

La température de transition vitreuse **"Tg"** et la température de fusion **"Tf"** sont les deux températures fondamentales nécessaires dans l'étude des matériaux polymères. La température de transition vitreuse **"Tg"** est partiellement importante pour les polymères amorphes, notamment les thermoplastiques amorphes, pour lesquels, il n'existe aucune force de cohésion importante autre que l'enchevêtrement.

Les zones cristallines ne fondent que bien au-delà de la température de transition vitreuse **"Tg"**. Les températures caractéristiques d'un seul et même matériau peuvent alors être classées de la façon suivante **[16]**: Température de transition vitreuse < Température de cristallisation < Température de fusion < Température de décomposition thermique.

I.4.3 Propriétés mécaniques

Les propriétés mécaniques des polymères sont fortement influencées par la température et la vitesse de déformation. Pour un même matériau, une augmentation de la température conduit à un comportement évoluant de fragile à ductile, de même qu'une décroissance de la vitesse de déformation **[17].**

I.4.4 Propriétés physiques

- **La masse volumique :** La masse volumique des matières plastiques est peu élevée. La légèreté des polymères est sans aucun doute une des qualités qui a le plus largement contribué à leur diffusion. **[16].**

I.4.5 Propriétés chimiques des polymères

La teneur en eau affecte les propriétés mécaniques des polyamides. La présence d'eau peut favoriser l'hydrolyse des polyesters (ainsi que la rupture des fibres de verres qu'ils contiennent éventuellement). Mais généralement les polymères sont peu affectés par les solutions salines, les acides faiblement dosés. En revanche ils sont pour la plupart sensibles aux solvants organiques (dérivés des pétroles…). Compte-tenu de la diversité des polymères, il faut toujours spécifier au fabriquant (via le fournisseur éventuellement) les conditions d'utilisation de son produit et l'associer au choix du meilleur produit ou compromis **[18].**

I.5 Les applications des polymères

Les polymères sont utilisés pour un nombre extraordinaires d'applications à tous les échelons de la vie **[5].**

Les différents domaines d'applications sont la construction mécanique, la robotique, L'énergétique ou encore l'aéronautique **[19].** Voir tableau suivant :

Tableau I.2. Les domaines d'application des polymères **[5,19].**

Domaines	**Exemples d'applications**
Automobile	Mousse de pare choc, finition intérieure, sièges, cache du moteur, garde boue, pneu, poignée de porte, volant, cadrans, réservoir,…
Avion	Sièges, bacs de rangements, roue d'atterrissage, poignées, aménagement du sol, tuyaux de ventilation, gilet de sauvetage,…
Eolienne	Pâles en composite, isolants, gaine des câbles
Chaine de montage	Tapis roulant, guides, tuyaux, pièces isolantes, écrans, manettes,
Le secteur de la construction	Les portes, conduites d'eau et l'assainissement, peinture des murs des plan chaires et des plafonds, revêtement du sol.
Industrie de l'emballage	Bouteilles, pots de yaourt, boites aux lettres, problème thermique, gainage films vidéo.
Médecine et santé	Poches de sang, gants, lentilles, verres de lunettes, les organes artificiels, Seringues, industrie dentaire, prothèses, outils de chirurgie
Articles ménagers	Tupperware, poubelles, seaux, vaisselle.
Matériel électrique et électronique et les communications	Tension d'isolement, isolation, laveuses, ordinateurs et les caméras, Radio et télévision, téléphone.
Industrie au Tom mobile	Système vitre claire, Tapie, carrosserie, optiques, planches de bord, habillage intérieur, Façades de téléviseurs, gainage de câbles.
Industrie textile	Vêtement, fibres textiles naturelles et synthétiques, sacs, skaï, similicuir, non tissés.
Sports-loisirs	Certaines parties des bâtiments, des avions et des bateaux, DVD, bandes magnétiques, Files de pêche et les cordons utilisés, piscines, coques de bateaux.
Industrie chimique	Tuyauterie, cuves, revêtements.
Le domaine alimentaire	Procédés de fabrication d'aliments, emballages (bouteilles, pots de yaourt, briques de lait, boites à oeufs) ou industriel (flacons de détergent, sachets et sacs, casiers)…

Le matériel de maison	Meuble, vaisselle, accessoires… ; Utilisés en industrie pour maintenir le matériel d'une violente collision.



Figure I.12. Quelque Applications des polymères **[3].**

I.6 Les méthodes de détermination de leurs propriétés

Les différentes méthodes physico-chimiques permettant l'analyse des polymères évoluent rapidement dans le sens d'une simplicité cob issante. La rapidité avec laquelle on peut obtenir un grand nombre d'informations et de possibilités de leur utilisation empirique fait que l'ingénieur (utilisateur) ne peut plus ignorer un certain nombre de ces méthodes.

Cette analyse physico-chimique des polymères faits l'objet de plusieurs fascicules ou les principales méthodes est :

- Spectroscopie infrarouge IR.
- Spectrométrie électronique SM.
- Spectroscopie d'absorption UV-Visible.
- Chromatographie en phase gazeuse (CPG).
- Analyse thermique ATD/ATG
- Chromatographie liquide (C L.)**[20]**

I.6.1Spectroscopie infrarouge IR

La spectroscopie infrarouge est une technique simple et fiable utilisée aussi bien en chimie organique qu'inorganique, en recherche et dans l'industrie. Cette

technique exploite le fait que les molécules absorbent des fréquences spécifiques caractéristiques de leur structure. Les absorptions correspondent aux fréquences de résonance, à savoir les fréquences telles que le rayonnement absorbé correspond à la fréquence de la liaison ou du groupe en vibration. La partie infrarouge du spectre électromagnétique est généralement divisée en trois régions : l'infrarouge proche, moyen et lointain, cette appellation se rapportant au spectre visible. L'infrarouge proche à plus haute énergie, de14000 à 4000 cm^{-1}, peut exciter des vibrations harmoniques. L'infrarouge moyen de 4000 à 400 cm^{-1}, peut être utilisé pour étudier les vibrations fondamentales et les structures associées rotations-vibrations.

L'infrarouge lointain, (400–10 cm^{-1}), proche de la région des micro-ondes, est à basse énergie et peut être utilisé pour de la spectroscopie rotationnelle.

Le spectre infrarouge d'un échantillon est enregistré en faisant passer le rayon infrarouge à travers l'échantillon. L'examen de la lumière transmise révèle la quantité d'énergie absorbée dans chaque longueur d'onde. Cela peut être fait avec un rayon monochromatique, dont la longueur d'onde change avec le temps, ou en utilisant un instrument à transformée de Fourier pour mesurer toutes les longueurs d'onde en même temps. De là, on obtient un spectre de transmission ou d'absorption, montrant les longueurs d'ondes absorbées par l'échantillon. L'analyse de ces caractéristiques d'absorption révèle des détails sur la structure moléculaire de quand la fréquence de l'IR est la même que la fréquence de vibration d'une liaison, il y a absorption des échantillons présentant peu de liaisons et de hauts niveaux de pureté, donnent des spectres simples.**[21]**

Des structures moléculaires plus complexes donnent des spectres plus complexes.

Les tests réalisés par l'appareil définis dans la Figure suivante



Figure I.13. Appareil infrarouge

I.6.2Etude thermique ATD/ATG

I.6.2.1 La thermogravimétrie

Est une technique mesurant la variation de masse d'un échantillon lorsqu'il est soumis à une programmation de température, sous atmosphère contrôlée Cette variation peut être une perte de masse (émission de vapeur) ou un gain de masse (fixation de gaz, oxydation…).**[22]**

- L'échantillon est pesé et la ligne de base est établie
- L'échantillon est soumis à la montée en température
- à différentes températures, des réactions chimiques peuvent libérer des espèces gazeuses ou former des oxydes entraînant une variation de masse de l'échantillon.
- Ce changement de masse est enregistré en fonction de la température.

I.6.2.2 L'analyse thermique différentielle ATD

Un grand nombre de transitions (fusion, cristallisation, transition vitreuse…) Ne s'accompagnent pas d'une variation de masse.

Ces transitions qui ne sont pas détectée par l'analyse thermogravimétrique le sont par lesanalyses thermique et enthalpique différentielles.



Figure I.14. Appareil Couplage ATD/.ATG

L'analyse thermique différentielle : est une technique mesurant la différence de température entre un échantillon et une référence (matériau inerte thermiquement) en fonction du temps ou de la température, lorsqu'ils sont soumis à une programmation de température, sous atmosphère contrôlée. Avec la méthode DTA, toute transformation est détectable pour toutes les catégories de matériaux.

- L'échantillon et la référence sont soumis à la montée en température
- Quand l'échantillon subit un changement de phase, sa réponse en température va varier par rapport à la référence
- Le thermocouple va mesurer cette variation. Celle-ci va être enregistrée

$$DT = TSTR = f(T)$$

- Le type de réaction (exo ou endothermique) va produire un pic.

Les tests réalisés par l'appareil illustré dans la **Figure I.14.**

Références

[1] Nkongo. Les polymères dans l'industrie , mémoire de master, Université de Douala, Cameroun, 2005.

[2] Belaidi, Othman. Etude d'un réseau de polymères à base d'ethyl hexyl acrylate (EHA) et de tri propylene glycol diacrylate (TPGDA). Diss, Université Abou Bekr Belkaïd-Tlemcen. Faculté des Science, 2007.

[3] I. Zoubire. Elaboration et caractérisation d'un matériau composite polymère ZnO , mémoire de master, université Abou Bekr Belkaid - Tlemcen, 2017.

[4] O, H. Annad. Synthèse et caractérisation d'un polymère fonctionnel (semi-conducteur) , mémoire de master, Université Mohamed khider - Biskra, 2019.

[5] Fairouz M. Etude des polymères et de leurs mélanges (état fondu) additionnés de composés d'origine minérale, mémoire de master, Université Mohamed Khider-Biskra , 2014.

[6] D. A. Tomalia, H. Baker, J. Dewald, J. Hall, J. Kallos, G. Martin, S. Roeck, J. Ryder J. Smith, Polym . J. 17.117, 1985.

[7] Flory, P. J. Principles of polymerchemistry . Cornell university press.p 672 (ISBN 0801401348, 9780801401343), 1953.

[8] Mekelburger, H.-B.; Vögtle, F.; Jaworek, W. Dendrimers, Arborols, and Cascade Molecules: Break through into Generations of New Materials . Angew. Chem. Int. Ed., vol 31, p 1571–1576, 1992.

[9] Newkome, G.R.; Keith, J.M.; Baker, G.R.; Escamilla, G.H.; Moorefield, C.N . Chemistry within a Unimolecular Micelle Precursor: Boron Superclusters by Site- and Depth-Specific Transformations of Dendrimers . Angew. Chem. Int. Ed.vol 33, p 666–668, 1994.

[10] Newkome, G.R.; Yao, Z.; Baker, G.R.; Gupta, V.K. Micelles. Part 1. Cascade molecules :a new approach to micelles. A [27]-arborol . J. Org. Chem. Am. Chem. Soc. 2002, 50, 2003–2004.

[11] S. Adkins, M. Burmeister, Analytical Biochemistry , Article, vol 240, P 17-234, 1996.

[12] N. Nemiche . Synthèse et caractérisation de copolymères séquencés à séquences hydrophiles de poly (1,3-dioxolane), de poly (éthylène glycol) et de poly (N-vinylpyrrolidone), mémoire de magister, Université Ahmed Ben Bella - Oran, 2013.

[13] A. Methia, M. A. Rezzoug .étude de la diffusion de l'inganox 1076 dans un polymèrere cyclé à base de Pehd. mémoire de master, université de Bejaia, 2017.

[14] Hamri S. Etude thermophysique de la diffusion de polymères acrylique. doctorat, Université Abou Bekr Belkaid, Tlemcen, 2013.

[15] Marvin J. Weber et al. (préf. Marvin J. Weber). Handbook of Optical Materials. Boca Raton, CRC Press ,536 p (ISBN 0-8493-3512-4 et 978-0849335129), 2002.

[16] L. Djoudi. Notion fondamentale sur les polymères, Article, Université Mohamed Khider - Biskra, 2007.

[17] S. Hamel. Elaboration et caractérisation de pastilles de polymères acryliques réticulés obtenus par photo-polymérisation UV-Application à la rétention de molécules de faible poids moléculaire, mémoire de master, Université Abou Bekr Belkaid - Tlemcen, 2016.

[18] Fond C. Les polymères. http://espace2christophe.chez.com/polymères.pdf. Consulter le 18/04/2022.

[19] Lecomte J.-Beckers, Physiques des matériaux, polymères , université de liège, 2009.

[20] Van Krevelen, D. W., & TeNijenhuis, K. Properties of polymers: their correlation with chemical structure; their numerical estimation and prediction from additive group contributions. Elsevier, 2009.

[21] REACH for Polymers, meilleures méthodes d'essai et techniques disponibles .smithers Rapra, 2011.

Chapitre II :

Méthodes de Calcul

II.1 La modélisation moléculaire

La modélisation moléculaire comprend toutes les méthodes, tant théoriques que computationnelles, utilisées pour modéliser ou imiter le comportement des molécules. Les méthodes sont utilisées dans les domaines de la chimie computationnelle, de la conception de médicaments, de la biologie computationnelle et de la science des matériaux pour étudier des systèmes moléculaires allant des petits systèmes chimiques aux grandes molécules biologiques et aux assemblages de matériaux. Les calculs les plus simples peuvent être effectués manuellement, mais des ordinateurs sont nécessaires pour effectuer la modélisation moléculaire de tout système de taille raisonnable. Une caractéristique commune des méthodes de modélisation moléculaire est la description du niveau atomique des systèmes moléculaires. Cela peut inclure le traitement des atomes comme la plus petite unité individuelle (approche de la mécanique moléculaire), ainsi que la recherche de la relation entre la structure des molécules et les propriétés physiques de divers matériaux (QSAR/QSPR).

La chimie quantique consiste en l'utilisation des méthodes basées sur la résolution de l'équation de Schrödinger indépendante de temps (état stationnaire). En résolvant l'équation aux valeurs propres et vecteurs propres $\hat{H}\Psi = E\Psi$, ou H est l'Hamiltonien non relativiste, E l'énergie totale et Ψ la fonction d'onde du système , il sera alors possible de déterminer toutes les informations du système étudié .Il n'est cependant pas possible de résoudre rigoureusement une telle équation dans la théorie quantique proposée dès les années 20 afin de pouvoir résoudre cette équation de façon approchée **[1].**

Bien que l'étude des relations structure-activité ait commencé à la fin du 19^{éme} siècle, ce n'est qu'au début des années 60 que les travaux de Corwin Hansch**[2],** ont proposé un modèle mathématique pour corréler l'activité biologique et la structure chimique. Pendant ces quarante dernières années, ce domaine a largement été étudié et les données bibliographiques disponibles sur cette approche sont maintenant importantes **[3].**

La méthode QSAR inclut toutes les méthodes statistiques par lesquelles des activités biologiques (le plus souvent exprimées par des logarithmes des activités

molaires équipotentielles) sont reliées avec les éléments structuraux (analyse de Free Wilson), les propriétés physico-chimiques (analyse de Hansch) ou différents paramètres liés à la notion de champ aidant à la description de la structure (3D QSAR) **[4].**

II.2Les méthodes de calcul

II.2.1 Les méthodes empiriques

Parmi les méthodes empiriques " non quantique", la mécanique moléculaire et la dynamique moléculaire classique sont les plus utilisées. La première est une méthode d'analyse conformationnelle basée sur l'utilisation de champs de forces empiriques et la minimisation d'énergie. Elle permet de décrire l'énergie en fonction des coordonnées atomiques et de chercher des minima de l'énergie correspondant à des conformères stables **[5].**

Toutefois, elle ne permet pas de décrire l'énergie en fonction du temps et de la température (entropie, énergie libre) et de franchir des barrières énergétiques. La dynamique moléculaire classique a pour but de calculer les mouvements des molécules, le plus souvent à partir des énergies de la mécanique moléculaire, en appliquant les lois de la mécanique classique. Elle permet de simuler l'évolution des systèmes dans le temps **[6].**

II.2.2 Les méthodes semi-empiriques

Les méthodes semi-empiriques sont exclusivement basées sur l'expérience. Elles sont dérivées des méthodes *ab-initio* dont elles reprennent le principe en ne considérant que les électrons de la couche de valence de chaque atome et en négligeant certaines intégrales rendant compte du recouvrement des orbitales entre elles. Ces approximations sont alors compensées par la mise en jeu des grandeurs physiques expérimentales, les méthodes semi empiriques sont généralement utilisées pour traiter de grosses molécules, afin d'obtenir une première description de ses propriétés moléculaires (structure, réactivité). On distingue plusieurs variantes : CNDO, INDO, MINDO/3, MNDO, RM1, PM3, AM1 **[6,7].**

II.2.3 Les méthodes quantiques

Les méthodes quantiques pures basées sur la mécanique quantique, 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 **[7].**

Ces méthodes sont des techniques de résolution de l'équation de Schrödinger de systèmes à plusieurs électrons elles utilisent des données ajustées sur des résultats expérimentaux afin de simplifier les calculs, les méthodes semi-empiriques contiennent une base minimale par défaut (STO-3G) par contre les méthodes quantiques pures utilisent des bases et des corrélations différentes ça dépend le type de calcul **[6,7].**

II.2.3.1Théorie de la fonctionnelle de la densité (DFT)

Les bases de la théorie de la fonctionnelle de la densité (DFT) ont été élaborées en 1927 par Thomas et Fermi qui calculèrent l'énergie d'un atome en représentant son énergie cinétique en fonction de la densité électronique **[7].** En 1928, Dirac introduit le terme d'échange prédit par Hartree mais il n'y a toujours aucune prise en compte de la corrélation électronique qui fût finalement ajoutée par Wigner. Dans ce modèle, les n électrons dépendants de $3n$ coordonnées d'espace sont remplacés par leur densité $\rho\ (r)$ qui ne dépend plus que de 3 variables.

L'état fondamental de ce système est décrit par la fonction d'onde $\Psi_0(r_1, r_2, \ldots . r_n)$ qui correspond à une unique densité électronique$\rho\ (r)$. Cette fonction d'onde, et l'énergie E_0 qui lui est associée, sont déterminées par la minimisation de l'énergie totale du système. Le potentiel externe $V_{ext}(r)$ créé par les N noyaux du système est alors complètement déterminé et fixe donc l'hamiltonien. Ainsi, le nombre d'électron n et le potentiel$V_{ext}(r)$, définissent toutes les propriétés de l'état fondamental **[7].**

II.2.3.2 Ab-initio

Les méthodes *ab-initio* sont caractérisées par l'introduction d'une base arbitraire pour étendre les orbitales moléculaires et alors le calcul explicite toutes les intégrales exigées qui impliquent cette base **[5].**

Les méthodes *ab-initio* se divisent en deux sous familles : les méthodes Hartree - Fock (HF, RHF, UHF, ROHF) (Hartree, 1928 ; Fock, 1930), et les méthodes post Hartree-Fock, (MPn, CAS, …) (Moller, 1934). La principale différence entre ces deux méthodes est que les interactions électroniques sont négligées dans les méthodes HF et réintroduites dans les méthodes post HF. Ces méthodes ne peuvent être appliquées qu'à des systèmes de quelques dizaines d'atomes pour les méthodes HF et d'une dizaine d'atomes seulement pour les méthodes post HF **[6,7].**

II.2.4Les méthodes QSAR (relation quantitative structure-activité)

Le modèle QSAR (Quantitative Structure-Activity Relation ships) est une corrélation basée sur un calcul statistique, de paramètres physico-chimiques avec l'activité biologique

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 physico-chimique ,etc.) , pour une série des molécules biologiquement actives à l'aide des méthodes d'analyse de données "une équation de corrélation"**[8].**

II.3Modèle QSAR / QSPR

II.3.1 Introduction

Les premiers essais de modélisation des relations structure-activité ont commencé à la fin du 19ème siècle, lorsque Crum-Brown et Frazer ont postulé que l'activité biologique d'une molécule est une fonction de sa constitution chimique **[9].** Mais ce n'est qu'au début des années 60 que les travaux de Corwin Hanch ont proposé un modèle mathématique reliant l'activité biologique à la structure chimique

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

Sans l'utilisation de grands instruments analytiques, les résultats des études QSAR/QSPR peuvent fournir des informations utiles pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent être alors utilisées dans la prédiction des activités biologiques de nouveaux composés ainsi que dans la conception de nouvelles structures **[10].**

II.3.2 Principe de QSAR/QSPR

Le principe des méthodes QSAR/QSPR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative la structure moléculaire, codée par des propriétés moléculaires appelées descripteurs, avec une activité en utilisant des méthodes d'analyse de données **[8,10].** Par ces relations nous pouvons développer des modèles prédictifs de la forme générale suivante :

$$Activité = f\ (descripteurs\ moléculaires)$$

$$Propriété = f\ (descripteurs\ moléculaires)$$

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

En pratique, le développement d'un modèle commence par la collecte de données expérimentales fiables et en nombre important, ensuite il est nécessaire de générer un nombre de descripteurs, caractérisant les structures moléculaires, utilisés dans le développement du modèle QSAR. Une fois le modèle construit, il doit être évalué par des méthodes de validation (validations interne et externe) afin d'estimer sa robustesse et son pouvoir prédictif**[10].**

II.3.3Outils et Méthodologie de QSAR/QSPR

II.3.3.1 Activité biologiques /propriété physique

Les données biologiques (propriétés physiques) 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 (propriété) (log 1/C) sont également utilisés pour

obtenir des valeurs mathématiques plus élevées lorsque les structures sont biologiquement très efficaces.

II.3.3.2 Les descripteurs moléculaires

Les descripteurs moléculaires sont généralement classés en trois catégories ; les descripteurs physicochimiques, topologiques et électroniques. Ces descripteurs sont caractéristiques de la structure bidimensionnelle ou tridimensionnelle de la molécule, l'étude QSAR-2D permet de rationaliser une propriété avec des descripteurs calculés à partir de la formule développée avec une équation de régression linéaire, les études QSAR- 3D, décrites pour la première fois par Cramer et ses collaborateurs **[11]**permettent d'établir une relation entre une propriété et des descripteurs calculés à partir d'une conformation spatiale de la molécule définie.

- **Les descripteurs physico-chimiques**

Les descripteurs physico-chimiques caractérisent généralement la structure - bidimensionnelle de la molécule.

Parmi ces descripteurs, certains reflètent la composition moléculaire du composé, soit le nombre et le type d'atomes et de liaisons présents dans la molécule, son nombre de cycle, son poids moléculaire… : ce sont les descripteurs constitutionnels **[12,13].**

Ces descripteurs ne fournissent pas assez d'information sur la structure des molécules pour l'élaboration des modèles prédictifs plus complexes ; il est nécessaire d'ajouter d'autres types de descripteurs.

- **Les descripteurs électroniques**

Ces descripteurs caractérisent la distribution de charge des molécules (polarité des molécules) mais aussi les paramètres de la chimie quantique : le moment dipolaire, les énergies HOMO (Orbitale moléculaire occupée de plus haute énergie) et LUMO (Orbitale moléculaire non-occupée de plus basse énergie)**[14].**

- **Les descripteurs topologiques**

Ces descripteurs décrivent les connectivités atomiques dans la molécule. Ils sont pour la plupart représentés par des indices comme ceux de Balaban, Wiener… et

sont issus de la théorie des graphes. Cette théorie a été développée par Euler en 1736 avec les sept ponts de Königsberg. Un graphe est un ensemble de point, certains reliés par des lignes ; il permet de représenter la topologie des molécules sans se soucier de la géométrie spatiale exacte de cette dernière **[13].**

II.3.4Les méthodes statistiques

Les techniques statistiques ou chimio-métriques constituent la base mathématique de la construction d'un modèle QSAR. La méthode la plus facile à interpréter est l'analyse de régression linéaire parmi diverses méthodes statistiques pour QSAR. Ces régressions représentent la corrélation directe des variables indépendantes (x) avec une variable dépendante (y). Ce modèle peut être considéré pour la prédiction de y à partir des données de variables x. Cela peut appartenir à un ensemble de systèmes qualitatifs ou quantitatifs. **[12] :**

II.3.4.1Régression linéaire simple (RLS)

Que variable dépendante. Cette technique s'avère très prometteuse pour générer des relations de structure et d'activité en explorant certains des descripteurs les plus importants utilisés pour gouverner l'activité, alors que certaines des interactions de plusieurs descripteurs ont été négligées. La régression linéaire simple peut être exprimée par l'équation 1 :

$$y = a + bx \qquad \textbf{(1)}$$

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

II.3.4.2 Régression linéaire multiple (RLM)

Cette méthode est l'extension de RLS à plus d'une dimension. Dans cette méthode, des calculs de régression multi variables standard sont effectués.

L'identification d'une propriété de médicament est effectuée sur tous les descripteurs sous enquête. La possibilité de corrélation est vérifiée par la valeur du coefficient de corrélation multiple (r), valeur t à l'aide e la méthode « leave-one-out ». La corrélation est vérifiée par des valeurs r^2ou Q^2 qui sont généralement connues sous le nom de coefficient de corrélation croisé.

La relation est exprimée en équation linéaire unique à plusieurs termes comme suit **[12]** :

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_mx_m + e. \qquad \textbf{(2)}$$

II.3.4.3 L'analyse des moindre carrée partiels(PLS)

L'analyse des moindre carrée partiels (Partial Least Squares) (PLS) est dans son objectif identique à la régression linéaire mais toutes les variables sont prises en compte dans l'équation grâce à une procédure proche de l'analyse factorielle. Cette méthode d'analyse est aujourd'hui la plus utilisée, puisqu'elle permet de prendre en compte un grand nombre de descripteurs. Elle est particulièrement adaptée pour les descripteurs 3D issus d'analyses CoMFA ou d'analyses de similarité moléculaire. L'analyse PLS est une méthode qui utilise des «vecteurs » représentant l'ensemble des variables. Plusieurs critères permettent de valider un modèle. Le premier de ces critères est la «signification statistique». Le second critère est la valeur du PRESS (Predicted Residual Error Sum of Squares) qui doit être la plus faible possible pour le vecteur considéré.

$$PRESS = \Sigma (Y_{Théorique} - Y_{prédite})^2 \quad \textbf{(3)}$$

Le calcul de ce critère est directement en relation avec la méthode qui détermine un autre critère qui est le «cross validation ». Le dernier critère est ler^2 qui donne l'aspect de la linéarité du modèle.

II.3.4.5Réseaux de Neurones Artificielle (RNA)

Les réseaux de neurones artificiels sont devenus en quelques années des outils précieux dans des domaines très divers de l'industrie et divers services, Ils sont particulièrement utilisés pour résoudre des problèmes de classification, de prédiction, de reconnaissance des formes, de catégorisation, de mémoire associative et d'optimisation **[15].**

Réseau de neurones artificiels (RNA) est une technique de reconnaissance qui ressemble de près le fonctionnement interne du cerveau qui est essentiellement composé de neurones interconnectés. Ce phénomène est affecté par la conception architecturale de RNA où les unités neuronales sont interconnectés les uns aux autres.

La **figure II.1**montre la structure d'un neurone artificiel. Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones amont. A chacune de ces entrées est associé un poids w abréviation de weight (poids) représentatif de la force de la connexion. Chaque processeur

élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals. A chaque connexion est associée un poids.

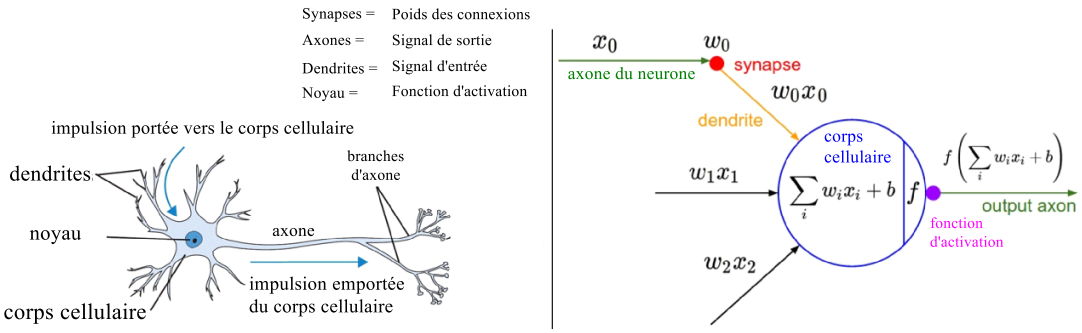


Figure II.1. Mise en correspondance neurone biologique / neurone artificiel

Une architecture couramment utilisé comme représenté sur la **figure II.2** contient un réseau d'anticipation à trois couches qui est composé d'une couche d'entrée, une couche cachée, et une couche de sortie. La couche d'entrée passe essentiellement des informations des variables indépendantes dans le système RNA ; par conséquent, le nombre d'unités de neurones présents dans la couche d'entrée est égal au nombre de variables indépendantes dans l'ensemble de données. Les connexions entre les neurones sont affectées des valeurs numériques connues comme poids.

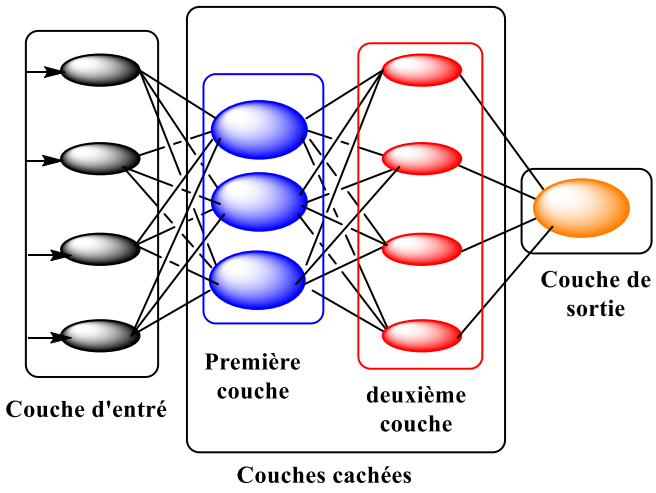


Figure II.2. Représentation d'un réseau de neurone

Les informations provenant de la couche d'entrée sont transmis à la couche cachée pour la configuration de traitement seront ensuite transmises de la couche cachée à la couche de sortie. Dans un algorithme de rétro-propagation, l'erreur calculée est dérivée de la différence entre la valeur prédite et la valeur réelle, et si elle

est acceptable, le processus d'apprentissage ne cesse autrement signaux seront envoyés vers l'arrière de la couche cachée à un traitement ultérieur et réajustements de poids. Ceci est réalisé de manière itérative jusqu'à ce qu'une solution soit trouvée et l'apprentissage est terminé.

II.3.5 Test de la signification totale de la régression

II.3.5.1 Coefficient de détermination (R^2)

Le coefficient de déterminationR^2 mesure l'adéquation entre un modèle issu d'une régression linéaire simple ou multiple et les données observées (ou les réalisations des variables aléatoires) qui ont permis de l'établir **[16].**Le R^2 se définit alors comme suit :

$$R^2 = \frac{\text{ESS}}{TSS} = \frac{TSS-RSS}{TSS} = 1 - \frac{RSS}{TSS} \qquad \textbf{(4)}$$

TSSest la somme des carrés totaux : $TSS = \sum(Y_{obs} - \bar{Y})^2$

ESSest la somme des carrés expliqués : $ESS = \sum(Y_{cal} - \bar{Y})^2$

RSSest la somme des carrée résiduels : $RSS = \sum(Y_{obs} - Y_{cal})^2$

II.4.5.2Coefficient de corrélation (r)

Le coefficient de corrélation $\boldsymbol{r}$mesure la direction et la force d'une relation linéaire. Le calcul de $\boldsymbol{r}$ est assez complexe. Nous nous concentrons sur la compréhension de ce que $\boldsymbol{r}$ signifie pour un nuage de points.

Voici quelques propriétés de $\boldsymbol{r}$:

- Il a toujours une valeur comprise entre -1 et +1.
- Les relations linéaires positives fortes ont des valeurs de $\boldsymbol{r}$ plus proches de +1.
- Les relations linéaires négatives fortes ont des valeurs de $\boldsymbol{r}$plus proches de -1
- Les relations plus faibles ont des valeurs de $\boldsymbol{r}$plus proches de 0.
- $R = (ESS/TSS)^{0.5} = (1 - ESS/TSS)^{0.5}$ **(5)**

II.3.5.3Ecart type (s)

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

$$s = \sqrt{\frac{RSS}{n-p-1}}$$

Avec n : Nombre d'observations dans l'échantillon **(6)**

Dont Pest le nombre de variables indépendantes (les descripteurs).

II.3.5.4 Le coefficient de détermination ajusté (R^2_{adj})

Le coefficient de détermination ajusté tient compte du nombre de variables. En effet, le principal défaut du R^2est de croître avec le nombre de variables explicatives. Or, on sait qu'un excès de variables produit des modèles peu robustes. C'est pourquoi on s'intéresse davantage à cet indicateur qu'au R^2. Mais ce n'est pas un véritable carré et il peut même être négatif.

$$R^2_{adj} = R^2 - \frac{k(1-R^2)}{n-k-1} = 1 - \frac{(1-R^2)(n-1)}{n-k-1} \quad \textbf{(7)}$$

Avec n : Nombre d'observations dans l'échantillon

k :Nombre de variables indépendantes

II.3.5.5. Coefficient de Prédiction (Q^2)

Si R^2 est une mesure de la justesse de la corrélation, Q^2 est une mesure de la justesse de la prédiction. Il mesure la capacité prédictive d'un modèle et donné par **[17]:**

$$Q^2 = 1 - \frac{PRESS}{SSY} \quad \textbf{(8)}$$

II.3.6Validation du modèle

II.3.6.1 Validation Interne

Quel que soit le but ultime du modèle QSAR, il doit être obligatoirement validé avant d'être interprété ou utilisé à fins prédictives. Il existe différentes solutions pour s'assurer de validité d'un modèle. La validation croisée LOO (leave-

one-out) est un processus qui permet de tester la précision prédictive d'un modèle, cette méthode basée sur le calcul de quelques paramètres statistiques tels que : $PRESS$,S_{PRESS}, R^2_{CV} et R^2_{adj}; ces paramètres sont calculés à partir des équations suivantes **[18]** :

$$PRSS = \sum (Y_{obs} - Y_{cal})^2 \quad \textbf{(8)}$$

$$TSS = \sum (Y_{obs} - Y_{cal})^2 \quad \textbf{(9)}$$

$$R^2_{adj} = (1 - R^2)(\frac{n-1}{n-p-1}) \quad \textbf{(10)}$$

$$R^2_{CV} = 1 - \frac{PRSS}{TSS} \quad \textbf{(11)}$$

$$S_{PRESS} = \sqrt{\frac{PRESS}{n}} \textbf{(12)}$$

Avec:

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

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

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

$\boldsymbol{R^2_{adj}}$**:** Estimation de la qualité d'ajustement du modèle à la population.

$\boldsymbol{R^2_{CV}}$**:** Est la variation de la statistique du R^2 obtenue en ajoutant ou en enlevant une variable indépendante.

II.3.6.2Validation externe

Cette méthode consiste à prédire l'activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modèle, cette validation est caractérisée par le paramètreR^2_{test}. Récemment plusieurs études [52,53] ont montré l'insuffisance des paramètres R^2, R^2_{CV} pour vérifier le pouvoir prédictif des modèles QSAR. Par conséquent, d'autres paramètres doivent être

vérifiés pour cet objectif. Ces paramètres sont connus sous le nom «critères de validation externe» ou souvent appelés «critères de Trophsa » **[18].**

II.3.7Les applications de l'étude QSAR

Il existe un grand nombre d'applications de ces modèles tels que **[19]** :

- 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 prédiction de la toxicité pour les espèces environnementales.
- La sélection des composés ayant des propriétés pharmacocinétiques optimales, que ce soit la stabilité ou la disponibilité dans les systèmes biologiques.
- La prédiction d'une variété de propriétés physico-chimiques 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.

Références

[1] J.P.Doucet, J. Weber, Computer –aided molecular design :Teory and applications, Academic Press, London, 266, (ISBN 0080529747, 9780080529745), 1996.

[2] Hansch, T. Fujita, J. Am. Chem. Soc. 86, p 1616-1626, 1964.

[2] H.N.G.Phuong, Thèse de doctorat, Synthèse et étude des relations structure – activité quantitatives (QSAR/2D) d'analogues benzo [c] phénathridiniques, Université D'angers, 2007.

[3] A. K. Debnath, Quantitative Structure-Activity Relationship (QSAR)

[4] G. Ohanessian, Applications de la Chimie Quantique en Chimie-Cours, École Polytechnique, Orsay, 2005.

[5] I.N.Levine, Quantum Chemistry, 5th edition, 455, 1999.

[6] R. Pollet, Méthodes de la Chimie Quantique, Cours de Master de Chimie– 2èmeannée, Université Paris-Sud 11 et ENS Cachan, 2006.

[7] H. Patrick, W. Langenacker, J. Tollenaere. Computational Medicinal Chemistry for Drug Discovery, marceldekker, New York, 1995.

[8] C. D. Selassie. History of quantitative Structure-Activity relationships, in Book History of quantitative Structure-Activity relationships, D. J. Abraham, Editor, John Wiley& Sons, Inc. California, 2003.

[9] A. Crum Brown, T.R. Fraser, Trans. Roy.Soc. 25, 151-203, 1868.

[10] H. Pajouhesh, G.R. Lenz. Medicinal chemical properties of successful central nervous system drugs.J. Am. Soc. Exp. NeuroThera., 2, 541-553, 2005.

[11] R,D, Cramer. D,E,Patterson, and J,D, Bunce .Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of stereoids to carrier proteins. J.Am.Chem. Soc. 110, 5959-5967, 1988.

[12] Roy, K., Kar, S., & Das, R. N.. A Primer on QSAR/QSPR Modeling: Fundamental Concepts: Springer International Publishing, (ISBN 3319172816, 9783319172811), 2015.

[13]Todeschini, R. Consonni, V. Mannhold, R. Kubinyi, H. & Timmerman, H. Handbook of Molecular Descriptors: Wiley, (ISBN 3527613110, 9783527613113), 2008.

[14] L,B, Kier. Molecular Orbital Theory in Drug Research, Academic Press, New York, 1981.

[15] P.J. DREW, et J.R.T. MONSON, Artificial neural networks, Surgery 127: 3-11, 2000.

[16] Z. Almi. Etude qualitative et quantitative des relations structures activités dans des hétérocycles à intérêt pharmaceutique. Thèse de doctorat. Université Mohamed Khider Biskra, 2015.

[17] F. Soualmia. Analyse multivariée des relations structures –activités dans des nouveaux composés organiques à utilisation médicinales. Thèse de doctorat. Université d'Oran, 2018.

[18] A. Tropsha, P. Gramatica, and V.K. Gombar, ―the importance of Being Earnest: Validation is the Absolute Essential for Successful Application and interpretation of QSPR Models, QSAR and Combinatorial Sciences, 22(1), p 69–77, 2003.

[19] J, leszczynski. Challenges and advances in computationalchemistry and physics, department of chemistry, jackson state university, U.S.A, 2010.

Chapitre III :

Application et interprétation des résultats

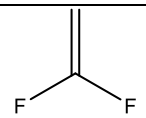
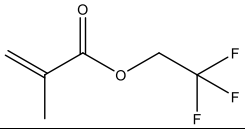
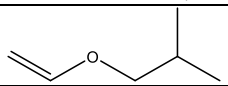
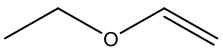
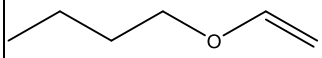
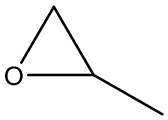
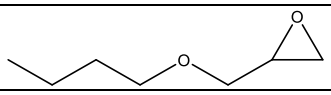
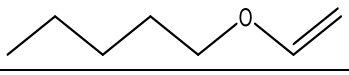
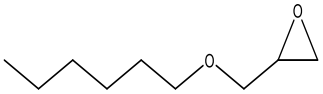
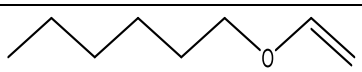
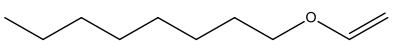
Ce chapitre est divisé en deux parties : dans la première partie nous expliquerons comment créer un modèle linéaire et interpréter les résultats obtenus dans la deuxième partie nous expliquerons comment créer un modèle non linéaire à l'aide de différents programmes, nous essaierons également de expliquer et discuter les résultats obtenus

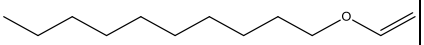
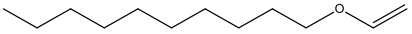
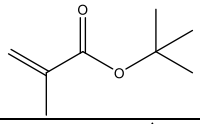
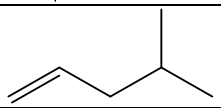
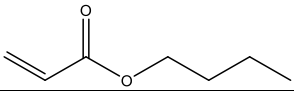
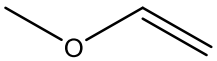
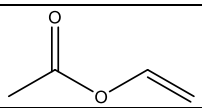
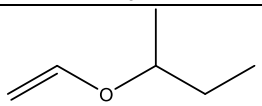
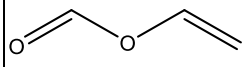
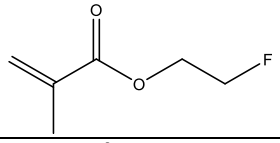
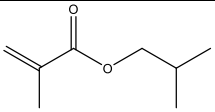
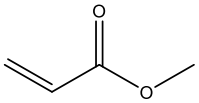
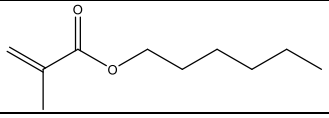
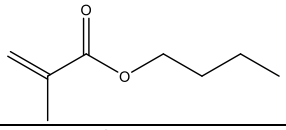
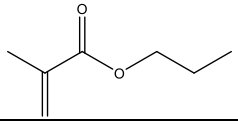
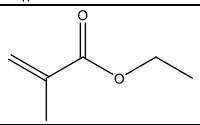
III.1 Préparation du modèle

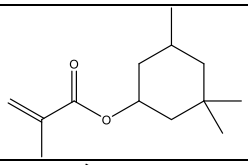
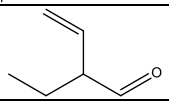
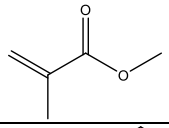
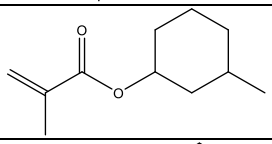
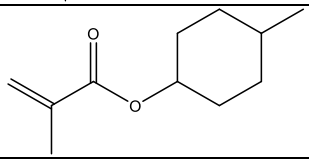
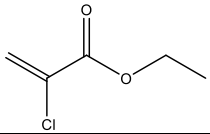
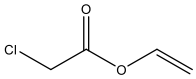
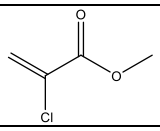
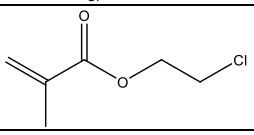
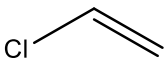
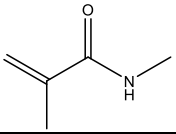
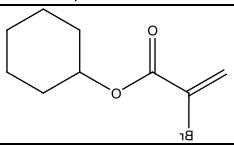
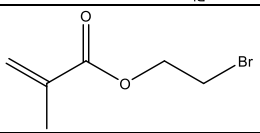
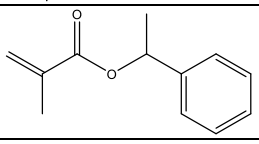
III.1.1Construction de la base de données

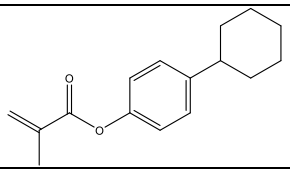
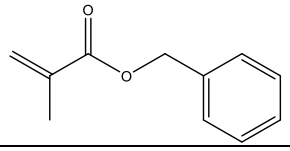
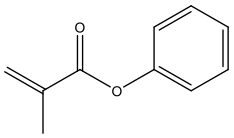
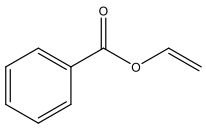
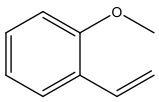
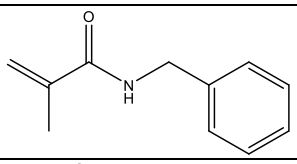
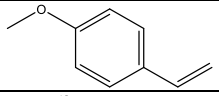
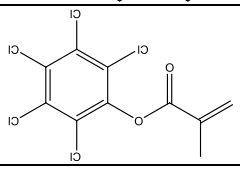
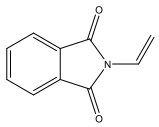
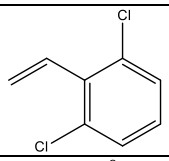
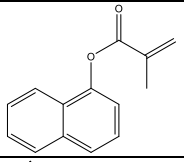
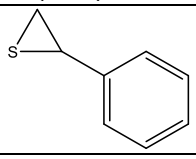
Dans la présente étude, un ensemble de données de 53 de polymères a été choisi en fonction de la disponibilité des données de la littérature et de la gamme de valeurs IR. Cet entrainement comprenait des méthacrylates, des acrylates, des esters, des amides et polymère des halogénures, **[1].** Le **tableau III.1** présente les structures chimiques de ces polymères et leurs indices de réfraction.

Tableau III.1. Structures chimique de polymères étudiées.

N°	Polymères	IR expérimental	Structures d'unités
1	Poly (vinylidenefluoride)	1.420	
2	Poly (trifluoroethylmethacrylate)	1.437	
3	Poly (vinylisobutylether)	1.451	
4	Poly (vinylethylether)	1.454	
5	Poly(vinyl n-butylether)	1.456	
6	Poly(propyleneoxide)	1.457	
7	Poly(3-butoxypropylene oxide)	1.458	
8	Poly(vinyl n-pentylether)	1.459	
9	Poly(3-hexoxypropylene oxide)	1.459	
10	Poly(vinyl n-hexylether)	1.459	
11	Poly(vinyl n-octylether)	1.461	

12	Poly(vinyl 2-ethylhexyl ether)	1.463	
13	Poly(vinyl n-decylether)	1.463	
14	Poly(t-butylmethacrylate)	1.464	
15	Poly(4-methyl-1-pentene)	1.465	
16	Poly(n-butyl acrylate)	1.466	
17	Poly(vinylmethylether)	1.467	
18	Poly(vinylacetate)	1.467	
19	Poly(vinyl sec-butylether)	1.474	
20	Poly(vinyl formate)	1.476	
21	Poly(2-fluoroethyl methacrylate)	1.477	
22	Poly(isobutylmethacrylate)	1.477	
23	Poly(methyl acrylate)	1.479	
24	Poly(n-hexylmethacrylate)	1.481	
25	Poly(n-butylmethacrylate)	1.483	
26	Poly(n-propylmethylacrylate)	1.484	
27	Poly(ethylmethacrylate)	1.485	

28	Poly(3,3,5-trimethylcyclohexyl methacrylate)	1.485	
29	Poly(vinylbutyral)	1.485	
30	Poly(methylmethacrylate)	1.489	
31	Poly(3-methylcyclohexyl methacrylate)	1.495	
32	Poly(4-methylcyclohexyl methacrylate)	1.498	
33	Poly(ethyl a-chloroacrylate)	1.502	
34	Poly(vinylchloroacetate)	1.513	
35	Poly(methyl a-chloroacrylate)	1.517	
36	Poly(2-chloroethyl methacrylate)	1.517	
37	Poly(vinylchloride)	1.539	
38	Poly(N-methylmethacrylamide)	1.540	
39	Poly(cyclohexyl a-bromoacrylate)	1.542	
40	Poly(2-bromoethyl methacrylate)	1.543	
41	Poly(1-phenylethyl methacrylate)	1.549	

42	Poly(p-cyclohexylphenylmethacrylate)	1.558	
43	Poly(benzylmethacrylate)	1.568	
44	Poly(phenylmethacrylate)	1.571	
45	Poly(vinyl benzoate)	1.578	
46	Poly(o-methoxystyrene)	1.593	
47	Poly(N-benzylmethacrylamide)	1.597	
48	Poly(p-methoxystyrene)	1.597	
49	Poly(pentachlorophenylmethacrylate)	1.608	
50	Poly(N-vinylphthalimide)	1.620	
51	Poly(2,6-dichlorostyrene)	1.625	
52	Poly(a-naphthylmethacrylate)	1.641	
53	Poly(styrenesulfide)	1.657	

III.1.2Calcul des descripteurs

Le programme ChemBioDraw Ultra 12.0 **[2]** a été utilisé pour optimiser la géométrie des molécules choisis, en appliquant la méthode de champ de force MM2qui a été développé par Norman Allinger **[3]**dans un premier temps pour l'analyse conformationnelle des hydrocarbures et autres petites molécules organiques. Il a été conçu afin de reproduire les géométries « covalentes » d'équilibre aussi précisément que possible. Les structures optimisée ont été sauvegardé avec une extension "*.mol".

La représentation numérique de la composition chimique (descripteurs moléculaires) est une étape importante. La performance du modèle développé et la précision des résultats sont étroitement liées à la manière dont ces descripteurs sont définis. Nous avons utilisé PADEL **[4]** pour calculer des descripteurs, le programme calcul plus que 1700 descripteurs pour chaque composé qui expriment principalement un ensemble de propriétés topologiques, géométriques et électroniques.

Le modèle conçu a été développé selon les étapes suivantes :

1- Suppression de toutes les colonnes qui ne contiennent pas de valeurs numériques (contiennent des zéros)
2- Suppression de toutes les colonnes avec des valeurs fixes ou semi-fixes.

III.2 Le modèle RLM

L'analyse par la méthode de régression linéaire multiple a été réalisée au moyen de logiciel SPSS 19 pour Windows **[5]**.

Les modèles rapportés par SPSS sont présentés sur le **tableau III. 2**

Tableau III.2. Les meilleurs descripteurs sélectionnés

Résumé du modèle			
Model	R	R^2	R^2_{adj}
1	0.688[a]	0.473	0.462
2	0.791[b]	0.625	0.609
3	0.828[c]	0.685	0.664
4	0.862[d]	0.743	0.72
5	0.884[e]	0.781	0.756
a. Predictors: (Constant), ASP0			
b. Predictors: (Constant), ASP0, ATSC0v			
c. Predictors: (Constant), ASP0, ATSC0v, maxHother			
d. Predictors: (Constant), ASP0, ATSC0v, maxHother, MLFER_E			
e.Predictors:(Constant),**ASP0, ATSC0v, maxHother, MLFER_E, ETA_Shape_Y**			

Il est très clair que le cinquième modèle est préférable en raison des valeurs élevées de $R^2 = 0.781$et $R^2_{adj} = 0.756$. Ce modèle a été construit à partir de cinq descripteurs : ASP-0, ATSC0v, maxHother, MLFER_E, ETA_Shape_Y.Ces descripteurs ont été utilisés pour le développement des modèles prédictifs. La description de ces descripteurs est donnée sur le **tableau III.3.**

Tableau III.3. La signification des descripteurs utilisés

Type	Descripteur	La description
2D	ASP-0	Chemin simple moyen, ordre 0 **[6]**
2D	ATSC0v	Autocorrélation Broto-Moreau centrée - décalage 0 / pondérée par les volumes de van der Waals**[7]**
2D	maxHother	Maximum d'atome H. État E : H sur aaCH, dCH2 **[8]**
2D	MLFER_E	Réfraction molaire excessive **[9]**
2D	ETA_Shape_Y	Indice de forme Y **[10,11]**

Afin d'évaluer la multi-colinéarité entre les cinq descripteurs pertinents. Leurs facteurs de variation d'inflation (VIF) ont été calculés. Si VIF est égal à 1. Aucun inter corrélation n'existe pour chaque variable ; si VIF est situé dans l'intervalle de 1 à 5. La sélection des descripteurs est acceptable si VIF est supérieur à 10. Le modèle associé aux descripteurs sélectionnés est instable et une nouvelle sélection est nécessaire.

D'après le **tableau III.4**, tous les descripteurs ont des valeurs VIF inférieures à 10 ce qui indique que la sélection effectuée est bonne et que les descripteurs obtenus ne sont pas corrélés entre eux **[12, 13].**

Tableau III.4. Les valeurs de VIF

	ASP0	ATSC0v	maxHother	MLFER_E	ETA_Shape_Y
VIF	4.482	1.905	1.722	3.793	1.977

La matrice de corrélation a été établie. La valeur du coefficient de corrélation de chaque paire de descripteurs sélectionnés a été examinée (**Tableau III.5**). La plus grande valeur du coefficient de corrélation appartenant aux descripteurs MLFER_E / ETA_Shape_Y étant égale à -0,684. Cela signifie que tous les descripteurs sélectionnés sont indépendants les uns des autres.

Tableau III.5. Matrice des coefficients de corrélation des descripteurs sélectionnés

	ASP0	ATSC0v	maxHother	MLFER_E	ETA_Shape_Y
ASP0	1.000	0.625	-0.604	0.761	-0.492
ATSC0v	0.625	1.000	-0.242	0.333	-0.293
maxHother	-0.604	-0.242	1.000	-0.562	0.433
MLFER_E	0.761	0.333	-0.562	1.000	-0.684
ETA_Shape_Y	-0.492	-0.293	0.433	-0.684	1.000

Le modèle RLM obtenu pour la prédiction des valeurs de IR expérimentales, en utilisant 53 polymères, est donné par la relation linéaire ci-dessous.

$$y = 2{,}115 - 0{,}585\ (ASP0) - 5{,}412E - 005\ (ATSC0v) - 0{,}185\ (maxHother)$$

$$+ 0{,}088\ (MLFER_E) - 0{,}148\ (ETA_Shape_Y) \quad \textbf{(01)}$$

Avec R^2= 0.781 et $R^2_{ajusté}$= 0.756

Les signes des coefficients de régression suggèrent le sens de l'influence des descripteurs dans un modèle donné. A partir des coefficients de régression de l'équation, il convient de noter que les coefficients de régression du descripteur MLFER_E ont des signes positifs. Ces descripteurs ont une influence positive et les valeurs les plus élevées contribueront à la prévision des valeurs de IR.

Ce modèle montre l'importance de la forme, le nombre d'hydrogènes et la réfraction molaire dans la prédiction de l'indice de réfraction.

Les valeurs de IR prédites par le modèle RLM pour l'ensemble des données (53 composés) Sont présentées dans le **Tableau III.6** et la **figure III.1**

Tableau III.6. Les valeurs de RI obtenues par le modèle RLM

Numéro de l'observation	IR (Exp)	IR (Préd)	Résidu
1	1.420	1.400	0.020
2	1.437	1.508	-0.071
3	1.457	1.500	-0.043
4	1.454	1.478	-0.024
5	1.456	1.489	-0.033
6	1.457	1.456	0.001
7	1.458	1.497	-0.039
8	1.459	1.486	-0.027
9	1.459	1.494	-0.035
10	1.459	1.462	-0.003
11	1.461	1.481	-0.020
12	1.463	1.444	0.019
13	1.463	1.477	-0.014
14	1.464	1.465	-0.001
15	1.465	1.472	-0.007
16	1.466	1.452	0.014
17	1.467	1.473	-0.006
18	1.467	1.478	-0.011
19	1.474	1.479	-0.005
20	1.476	1.476	0.000
21	1.477	1.471	0.006
22	1.477	1.463	0.014
23	1.479	1.476	0.003
24	1.481	1.486	-0.005
25	1.483	1.463	0.020

26	1.484	1.484	0.000
27	1.485	1.484	0.001
28	1.485	1.491	-0.006
29	1.485	1.490	-0.005
30	1.489	1.484	0.005
31	1.495	1.494	0.001
32	1.498	1.516	-0.018
33	1.502	1.488	0.014
34	1.513	1.506	0.007
35	1.517	1.539	-0.022
36	1.517	1.508	0.009
37	1.539	1.528	0.011
38	1.540	1.539	0.001
39	1.542	1.551	-0.009
40	1.543	1.547	-0.004
41	1.549	1.548	0.001
42	1.558	1.570	-0.012
43	1.568	1.575	-0.007
44	1.571	1.570	0.001
45	1.578	1.512	0.066
46	1.593	1.582	0.011
47	1.597	1.599	-0.002
48	1.597	1.598	-0.001
49	1.608	1.601	0.007
50	1.620	1.634	-0.014
51	1.625	1.637	-0.012
52	1.641	1.511	0.130
53	1.657	1.627	0.031

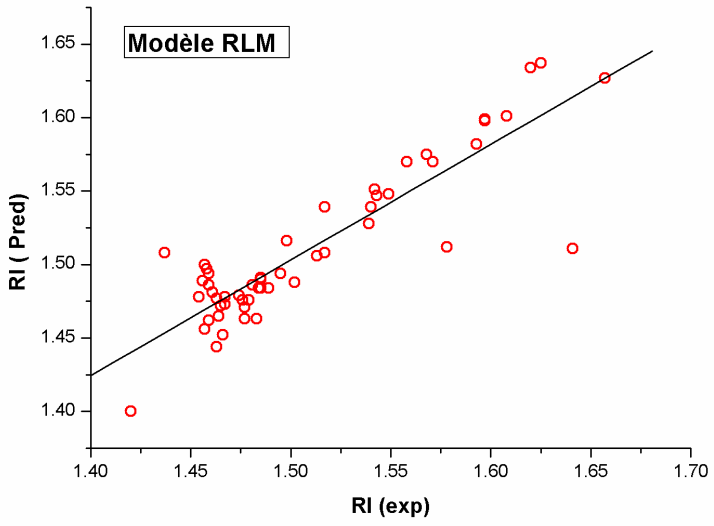


Figure III.1. Graphique des valeurs de RI observées par rapport aux valeurs prédites du modèleRLM pour l'ensemble d'apprentissage

III.3Le modèle RNA

Dans le cadre de ce travail, des réseaux de neurones artificiels (RAN) ont été utilisés pour construire un modèle non linéaire basé sur les mêmes descripteurs utilisés dans le cas du modèle RLM.

Pour l'obtention du model RNA en a utilisé le logiciel STATISTICA (**STATISTICA 10.0**).

L'objectif général d'un RNA est de trouver la configuration des poids de connexion entre neurones pour qu'il associe à chaque configuration d'entrée. Une réponse adéquate.

L'utilisation d'un RNA se fait en deux temps. Tout d'abord une phase d'apprentissage qui est chargée d'établir des valeurs pour chacune des connexions du réseau. Puis une phase d'utilisation proprement dite **[14].**

La base de données a été divisée en deux ensembles : 80 % du nombre total de composés pour l'ensemble d'entraînement et 20 % pour l'ensemble de test. L'ensemble d'entraînement était dédié au développement du modèle QSAR. L'ensemble de test, qui n'avait pas été utilisés pour le développement du modèle QSAR, a été laissé pour la validation externe du modèle.

La fonction sigmoïde et la fonction d'identité ont été utilisées comme fonctions de transfert de la couche d'entrée vers la couche cachée et de la couche cachée vers la couche de sortie, respectivement. Le réseau a été formé en utilisant l'algorithme des méthodes quasi-Newton BFGS.

(53 composés) de la base de données des polymères sélectionnés, été divisé en deux groupes : un ensemble d'apprentissage de 43 composé et un ensemble de validation composé de10 composés. Pour optimiser le nombre de nœuds dans la couche cachée, on a basé sur les valeurs des paramètres *Training perf* et *Test Perf*, qui indiquent la performance du réseau sur les sous-ensembles de données utilisés pendant l'apprentissage(**Tableau III.7)**.Par conséquent, le modèle final de ce travail a été choisi avec une architecture MLP 5-5-1 (**Figure III.2**)et un Le coefficient de corrélation linéaire $\boldsymbol{R^2_{Train} = 0.97}$et $\boldsymbol{R^2_{Test} = 0.99.}$

Tableau III.7. Les paramètres du meilleur réseau de neurones sélectionnés

architecture	Training perf.	Test perf.	Apprentissage erreur	Test erreur	Algorithme d'apprentissage
MLP 5-5-1	0.982564	0.995664	0.000057	0.000020	BFGS 125

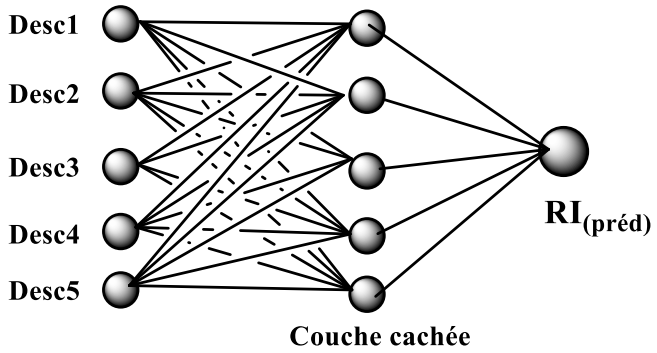


Figure III.2. L'architecture du réseau de neurones sélectionnés (MLP 5-5-1)

Les résultats prédictifs du modèle MLP pour l'ensemble des données (53 composés) ont été obtenus et présentés dans le **tableau III.8**, **figure III.3** et**figure III.4.**

Tableau III.8. Les valeurs IR (expérimentales et prédites) pour l'ensemble d'apprentissage et du test avec RNA

Numéro de l'observation	IR (Exp)	IR (Préd)	Etat
1	1.420	1.464	Train
2	1.437	1.439	Train
3	1.457	1.443	Train
4	1.454	1.454	Train
5	1.456	1.470	Train
6	1.457	1.470	Train
7	1.458	1.466	Train
8	1.459	1.467	Train
9	1.459	1.474	Train
10	1.459	1.470	Train
11	1.461	1.469	Train
12	1.463	1.461	Train
14	1.464	1.472	Train
15	1.465	1.465	Train
17	1.467	1.470	Train
18	1.467	1.466	Train
19	1.474	1.468	Train
20	1.476	1.468	Train
21	1.477	1.469	Train
22	1.477	1.476	Train
24	1.481	1.474	Train
25	1.483	1.470	Train
26	1.484	1.474	Train
27	1.485	1.474	Train
29	1.485	1.482	Train

31	1.495	1.480	Train
34	1.513	1.471	Train
35	1.517	1.520	Train
36	1.517	1.509	Train
38	1.540	1.550	Train
39	1.542	1.553	Train
40	1.543	1.554	Train
41	1.549	1.535	Train
42	1.558	1.569	Train
43	1.568	1.584	Train
44	1.571	1.569	Train
45	1.578	1.579	Train
46	1.593	1.595	Train
48	1.597	1.583	Train
49	1.608	1.594	Train
50	1.620	1.622	Train
51	1.625	1.627	Train
52	1.641	1.639	Train
13	1.463	1.466	Test
16	1.466	1.467	Test
23	1.479	1.466	Test
28	1.485	1.481	Test
30	1.489	1.483	Test
32	1.498	1.498	Test
33	1.502	1.497	Test
37	1.539	1.552	Test
47	1.597	1.594	Test
53	1.657	1.649	Test

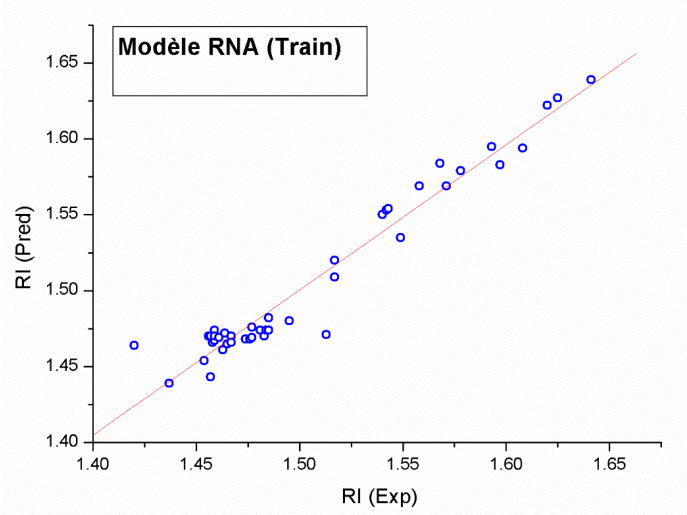


Figure III.3. Graphique des valeurs de RI observées par rapport aux valeurs prédites du modèleRNA pour l'ensemble d'apprentissage

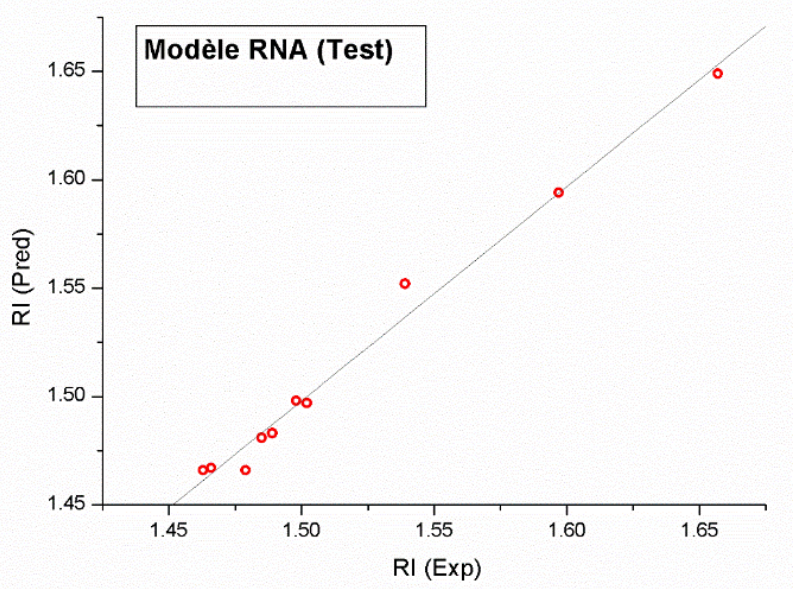


Figure III.4. Graphique des valeurs de IR observées par rapport aux valeurs prédites du modèle RNA pour l'ensemble du test

Une forte liaison entre les valeurs expérimentales et les valeurs prédites traduit par la valeur de R^2 ($\boldsymbol{R^2_{Train} = 0.97}$) ce qui signifie les 5 descripteurs pertinents pourraient expliquer 97 % de la variance de l'indice de réfraction (IR).

III. 4. Conclusion

Dans ce chapitre. Nous avons montré que le RNA est l'outil le plus puissant pour prédire l'activité inhibitrice en comparant les résultats des deux modèles. Les valeurs R et R^2les plus élevées ont été trouvées pour le modèle de RNA.

La validation de ce modèle a été faite par le calcul de R^2_{Test}**.** La valeur de $R^2_{Test} = 0.99$ signifie une très grande corrélation entre les valeurs expérimentales et les valeurs prédites.

Références

[1] Andrew J. Holdera, L in Yeb, J. David Eickb and Cecil C. Chappelowc "A Quantum-Mechanical QSAR Model to Predict the Refractive Index of Polymer Matrices"QSAR Comb. Sci. 25, No. 10, 905 – 911, 2006.

[2]https://www.cambridgesoft.com/ .Consulter le 22/04/2022.

[3] E. Eliel and N. L. Allinger. Topics in Stereochemistry, John Wiley & Sons. (ISBN 047014744X, 9780470147443), 2009.

[4]YAP, Chun Wei. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. Journal of computational chemistry, vol. 32, no 7, p 1466-1474, 2011.

[5] Spss 19 For Windows, SPSS software packages, SPSS Inc., 444 North Michigan Avenue, Suite 3000, Chicago, Illinois, 60611, USA. http://www.spss.com

[6] Kier, L. B., and Hall, L. H. Molecular connectivity in chemistry and drug research, (New York: Academic Press), 1976.

[7] Todeschini, R. and Consonni, V. Molecular descriptors for chemoinformatics, (Weinheim: Wiley VCH) p 27-37, 2009.

[8] Hall, L. H., and Kier, L. B. Electrotopological state indices for atom types: A novel combination of electronic, topological, and valence state information. J ChemInf, 1995.

[9]Platts JA, Butina D, Abraham MH, Hersey A. Estimation of molecular free energy relation descriptors using a group contribution approach. J ChemInfComput Sci.

[10] Roy K & Ghosh G, QSTR with Extended Topochemical Atom Indices. 2. Fish Toxicity of Substituted Benzenes. J. Chem. Inf. Comput. Sci., 44, p 559-567, 2004. http://dx.doi.org/10.1021/ci0342066

[11] Roy K, Das R N, on some novel extended topochemical atom (ETA) parameters for effective encoding of chemical information and modeling of fundamental physicochemical properties. SAR QSAR Environ Res 22, p 451-472, 2011. http://dx.doi.org/10.1080/1062936X.2011.569900

[12] Chen, M., Jabeen, F., Rasulev, B., Ossowski, M., &Boudjouk, P. A computational structure–property relationship study of glass transition temperatures for a diverse set of polymers. Journal of Polymer Science Part B:PolymerPhysics, 56(11),p 877-885, 2018.

[13] BITAM, S., HAMADACHE, M., et HANINI, S. QSAR model for prediction of the therapeutic potency of N-benzylpiperidine derivatives as AChE inhibitors. SAR and QSAR in EnvironmentalResearch, vol. 28, no 6, p 471-489, 2017.

[14] Bourret, P. Reggia, J. &Samuelides, M. Réseaux neuronaux (une approche connexionniste de l'intelligence artificielle), 1991.

Une étude de la relation quantitatives structure activité QSAR est appliquée à un ensemble de 53polymères, afin de prédire l'indice de réfraction (IR) et de trouver une corrélation entre les différents paramètres moléculaires (descripteurs), en utilisant la régression linéaire multiple RLM et les réseaux des neurones RNA.

Les résultats statistiques de RLM, du RNA indiquent que les coefficients de détermination R^2 étaient respectivement de 0.87, 0.97. Ce critère utilisé pour valider ces modèles a révélé la supériorité du modèle RNA qui possède une architecture 5-5-1

Les deux modèles montrent l'importance de la forme, le nombre d'hydrogènes et la réfraction molaire dans la prédiction de l'indice de réfraction

Les résultats montrent que les méthodes de modélisation permettent de bien prédire (IR) et peuvent être utiles pour prédire l'indice de réfractiond'autres polymères.

A travers les différents résultats obtenus au cours de ce travail, nous pouvons dire que les traitements statistiques et les techniques informatiques utilisés, lors du développement et l'analyse des modèles de QSAR, ont donné de bons résultats, ce qui nous permet d'entrevoir des perspectives assez prometteuses dans ce domaine par l'amélioration des traitements statiques et par l'utilisation d'autres méthodes d'apprentissage.