



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

**UNIVERSITÉ IBN-KHALDOUN - TIARET-**

**ANNEXE SOUGUEUR**

**MÉMOIRE DE FIN D'ÉTUDES**

*Pour l'obtention du diplôme de Master*

*Filière : CHIMIE*

*Spécialité : Chimie Organique*

**Par :**

**CHAOU FATIMA ZOHRA**  
**LAIDI SAKINA**

**THÈME**

**Etude de QSAR sur des activités biologiques utilisant des produits D'origine naturels**

**Soutenue publiquement le : 30 / 09 /2020 devant le Jury composé de:**

|   |         |                      |              |
|---|---------|----------------------|--------------|
| Mr. ATMANI<br>Abdellali                   | M. C .B | Université de Tiaret | Président    |
| M <sup>me</sup> .<br>TANDJAOUI<br>Nassima | M. C .B | Université de Tiaret | Examineur    |
| M <sup>me</sup> . Maizi<br>Yasmina        | M. A. A | Université de Tiaret | Examineur    |
| M <sup>lle</sup> . LAOUD<br>Aicha         | M. A. B | Université de Tiaret | Encadreur    |
|   | M. C. B | Université de Tiaret | Co-Encadreur |

## **DEDICACE**

*Je dédie ce travail à :*

*Ma chère mère, symbole d'affection, d'amour et de sacrifices, avec toute ma reconnaissance.*

*Mon Père, Pour son amour, son encouragement et ses conseils.*

*Mes sœurs (NASSIRA, IKRAM) et la petite(BOUCHRA), je prie dieu pour la protégée.*

*Mes proche frères (ABD RAHMEN, SOUFIANE), que je les aime tous.*

*Mes amis (AHMED WAHBI, GHOFAN, YASMIN, HADJER, KHADIDJA, KARIMA, RAFIKA)*

*Toute ma famille : CHAOU, OMRANI.*

*Tous ceux que j'aime et que je respecte.*

**CHAOU Fatima-Z**

## **DEDICACE**

*Je dédie ce travail à :*

*A ma chère mère, symbole d'affection, d'amour et de sacrifices, avec toute ma reconnaissance (rahémahaallah).*

*A mon Père, Pour son amour, son encouragement et ses conseils.*

*A mes sœurs (HABIBA, SOUMIA)*

*A mes proche frères (LAKHDAR, MOUHAMED), que je les aime tous.*

*A mes amies (HADJIRA, FATIHA, LINDA, HADJER, KHADIDJA)*

*A toute ma famille : LAIDI.*

*A tous ceux que j'aime et que je respecte.*



**LAIDI SAKINA**

## **Remerciements**

*« Tout d'abord merci à mon Dieu »*

*Nous tenons à remercier en tout premier lieu, mon encadreur M<sup>lle</sup> LAOUD AICHA à l'université de Tiaret pour son encadrement, ses efforts, son soutien.*

*Nous tenons aussi à remercier ceux qui m'ont fait l'honneur de juger ce travail : M<sup>me</sup> MAIZI YASMINA, M<sup>me</sup> TANDJAOUI NASSIMA et Mr ATMANI ABDELALLI pour avoir accepté d'examiner ce travail et participer au Jury.*

*Nous voulons aussi remercier tous les enseignants de faculté des sciences de la matière, université IBN KHALDOUN Tiaret pour leur aide et sympathie.*

*Aussi, à tous nos amis et à toute personne chère.*

*A tous ceux qui nous ont soutenus de près ou de loin*

*Et surtout nous exprimons notre profonde gratitude et affection à nos chers parents qui ont sacrifié pour notre bonheur et réussite et à nos chers frères et sœurs, sans exception à toute la famille CHAOU et LAIDI.*

## **Résumé**

Les modèles quantitatifs de QSAR présentent une solution statistique du problème de la difficulté du calcul direct des propriétés physiques et biologiques à partir de la structure.

L'intérêt d'un modèle de QSAR est de tirer des informations à partir de l'ensemble des descripteurs numériques caractérisant la structure moléculaire et prédire ainsi les activités biologiques de nouvelles structures.

Cette mémoire décrit, en deux chapitres, le premier chapitre rapporte sur les glycosidases et leurs inhibiteurs, le second chapitre décrit la méthodologie de QSAR

**Mots Clés:** descripteurs, QSAR, glycosidases, inhibiteurs.

**Abstract :**

Quantitative structure- activity relationship (QSAR) models are a statistical solution to the problem of directly calculating physical and biological properties of molecules from their structure.

The goal of a QSAR model is to extract information from a set of numerical descriptors characterizing molecular structure and use this information to inductively develop a relationship between structure and activity.

This memory describes, by two chapters, the first chapter reports on glycosidases and their inhibitors, and the second chapter the methodology involved in QSAR.

**Keywords:** descriptors, QSAR, glycosidase, inhibitors.

## ملخص

في هذا العمل ، أجريت دراسات QSAR على ثمانية وعشرين من مشتقات الفلافونويد. المركبات المستخدمة لها نشاط بيولوجي واسع النطاق ، والذي يمكن أن يفسر على الأقل نشاطها التثبيطي  $\alpha$ -glucosidase. تم استخدام إجراء الانحدار الخطي المتعدد (MLR) لتطوير العلاقات بين الواصفات الجزيئية والنشاط البيولوجي لمشتقات الفلافونويد. تم تقييم الاستقرار والقوة التفسيرية والقدرة التنبؤية لأفضل نموذج تم الحصول عليه من QSAR من خلال التحقق الداخلي وتقنية التوزيع العشوائي Y والتحقق الخارجي

تظهر نتائجنا أن نموذج QSAR المطور يمكن تفسيره ، ويتمتع باستقرار جيد ، وقوة ، ولديه قدرة تنبؤية كبيرة.

الكلمات المفتاحية: MLR ، النشاط المثبط ، QSAR ، ألفا جلوكوزيداز ، الفلافونويد.

## *Liste des abréviations*

---

### **Liste des abréviations:**

|                                    |   |
|------------------------------------|---|
| <b>QSAR:</b>                       | Quantitative Structure Activity Relationship / Relation quantitative de structure activité.                     |
| <b>QSPR</b>                        | Quantitative structure-property relationship/ Relation quantitative de structure propriété.                     |
| <b>CAZ<sub>y</sub>:</b>            | Carbohydrate-Active Enzymes.  |
| <b>IC50:</b>                       | La concentration micro molaire d'une drogue, nécessaire pour inhiber 50% (la moitié) de l'activité enzymatique. |
| <b>K<sub>i</sub>:</b>              | la constante de dissociation du complexe (enzyme / inhibiteurs)   |
| <b>MLR:</b>                        | Méthode de Régression Linéaire  |
| <b>D:</b>                          | Les descripteurs  |
| <b>1D:</b>                         | une dimension   |
| <b>2D:</b>                         | Deux dimensions   |
| <b>3D:</b>                         | Trois dimensions  |
| <b>E<sub>LUMO</sub>:</b>           | Energie de l'orbitale la plus basse vacante   |
| <b>μ:</b>                          | Potentiel chimique  |
| <b>η:</b>                          | Dureté chimique   |
| <b>σ:</b>                          | Mollesse chimique   |
| <b>ω:</b>                          | Indice d'électrophile globale   |
| <b>AEI:</b>                        | Indice de l'énergie d'activation  |
| <b>MSD:</b>                        | Densité de Spin de Mulliken   |
| <b>E<sub>SOMO</sub>:</b>           | Energie de l'orbitale occupée par un électron   |
| <b>K<sub>ow</sub>:</b>             | Coefficient de partage (octanol-eau)  |
| <b>logP:</b>                       | Coefficient de partition  |
| <b>R<sup>2</sup>:</b>              | Coefficient de corrélation pour les molécules de la série de test.  |
| <b>R<sup>2</sup><sub>0</sub>:</b>  | Coefficient de corrélation entre les valeurs prédites et expérimentales pour la série de test.                  |
| <b>R<sup>2</sup>'<sub>0</sub>:</b> | Coefficient de corrélation entre les valeurs expérimentales et prédites pour la série de test.                  |
| <b>k`:</b>                         | Constante de la droite (à l'origine) de corrélation (valeurs expérimentales en fonction des valeurs prédites).  |
| <b>DA:</b>                         | Domaine d'applicabilité   |



## ***Liste des figures***

---

---

### **Chapitre I: Les glycosidases et leurs inhibiteurs**

|                   |  |           |
|-------------------|--|-----------|
| <b>Figure 1:</b>  | <i>Représentation graphique de Lineweaver-Burk pour l'inhibition compétitive</i>     | <b>11</b> |
| <b>Figure 2:</b>  | <i>Représentation graphique de Lineweaver-Burk pour l'inhibition incompétitive</i>   | <b>12</b> |
| <b>Figure 3:</b>  | <i>Représentation graphique de Lineweaver-Burk pour l'inhibition non compétitive</i> | <b>13</b> |
| <b>Figure 4:</b>  | <i>Structure chimique de l'acarbose</i>  | <b>14</b> |
| <b>Figure 5:</b>  | <i>Structures chimiques de Valiénamine, Valiolamine et Voglibose</i>                 | <b>15</b> |
| <b>Figure 6:</b>  | <i>Structure chimique de l'<math>\alpha</math>-méthylacarviosine</i>                 | <b>15</b> |
| <b>Figure 7:</b>  | <i>Structures de cinq grandes classes d'aminosucres</i>                              | <b>16</b> |
| <b>Figure 8:</b>  | <i>Structures chimiques de quelques aminosucres inhibiteurs de glycosidases</i>      | <b>17</b> |
| <b>Figure 9:</b>  | <i>Structures chimiques de swainsonine et de castanospermine</i>                     | <b>17</b> |
| <b>Figure 10:</b> | <i>Structures chimiques de salacinol et de kotalanol</i>                             | <b>18</b> |

### **Chapitre II : Méthodologie de QSPR/QSAR**

|                   |   |           |
|-------------------|---|-----------|
| <b>Figure 1:</b>  | <i>Présentation des étapes aboutissant</i>                          | <b>22</b> |
| <b>Figure 2 :</b> | <i>Procédure d'obtention et de validation d'un modèle QSPR/QSAR</i> | <b>23</b> |

## ***Liste des schémas***

---

### **Chapitre II: Les glycosidases et leurs inhibiteurs**

|                   |  |           |
|-------------------|--|-----------|
| <b>Schéma 1 :</b> | <i>Hydrolyse simplifiée de la liaison glycosidique</i>                     | <b>04</b> |
| <b>Schéma 2:</b>  | <i>Site actif des glycoside-hydrolases</i>                                 | <b>06</b> |
| <b>Schéma 3</b>   | <i>Mécanisme générale des glycosidases avec inversion de configuration</i> | <b>07</b> |
| <b>Schéma 4:</b>  | <i>Mécanisme générale des glycosidases avec rétention de configuration</i> | <b>08</b> |
| <b>Schéma 5:</b>  | <i>Etat de transition lors de l'hydrolyse glycosidique</i>                 | <b>09</b> |
| <b>Schéma 6:</b>  | <i>Equation décrivant la réaction enzymatique selon Michaëlis-Menten</i>   | <b>10</b> |
| <b>Schéma 7:</b>  | <i>Mécanisme réactionnel de l'inhibition compétitive.</i>                  | <b>11</b> |
| <b>Schéma 8:</b>  | <i>Mécanisme réactionnel de l'inhibition incompétitive</i>                 | <b>12</b> |
| <b>Schéma 9:</b>  | <i>Mécanisme réactionnel de l'inhibition non compétitive</i>               | <b>13</b> |

### **Plan du travail**

|                   |  |           |
|-------------------|--|-----------|
| <b>Schéma 1 :</b> | <i>Représentation du protocole appliqué pour la génération d'un modèle 2D-QSAR</i> | <b>33</b> |
|-------------------|--|-----------|

## *Liste des tableaux*

---

### **Chapitre II : Méthodologie de OSPR/OSAR**

|                  |  |           |
|------------------|--|-----------|
| <b>Tableau1:</b> | <i>Exemples de descripteurs électroniques de la chimie quantique</i> | <b>25</b> |
|------------------|--|-----------|

# TABLE DE MATIERE

---

---

|  |    |
|--|----|
| <i>Introduction générale</i> .....                                   | 01 |
| <b><u>Chapitre I: Les glycosidases et leurs inhibiteurs</u></b>      |    |
| <i>1-Les glycosidases</i> .....                                      | 04 |
| <i>Généralités</i> .....   | 04 |
| <i>Rôles des glycosidases</i> .....                                  | 05 |
| <i>Mécanismes d'action des glycosidases</i> .....                    | 05 |
| <i>Mécanisme avec inversion de configuration</i> .....               | 06 |
| <i>Mécanisme avec rétention de configuration</i> .....               | 07 |
| <i>Mécanismes particuliers</i> .....                                 | 08 |
| <i>Modèle général de l'état de transition</i> .....                  | 09 |
| <i>2- l'inhibition enzymatique</i> .....                             | 09 |
| <i>Inhibition compétitive</i> .....                                  | 11 |
| <i>Inhibition incompétitive</i> .....                                | 12 |
| <i>Inhibition non compétitive</i> .....                              | 12 |
| <i>3- Inhibiteurs de glycosidases</i> .....                          | 13 |
| <i>3-1 Les aminocyclitols</i> .....                                  | 13 |
| <i>Les aminosucres</i> .....   | 15 |
| <i>Les thiosucres</i> .....  | 18 |
| <i>Références Bibliographiques</i> .....                             | 19 |
| <b><u>Chapitre II : Méthodologie de QSPR/QSAR</u></b>                |    |
| <i>1- Introduction à la modélisation QSPR/QSAR</i> .....             | 22 |
| <i>2- Principe des méthodes QSPR/QSAR</i> .....                      | 23 |
| <i>3- Méthodologie générale d'une étude QSPR/QSAR</i> .....          | 23 |
| <i>3-1- Bases de données</i> .....                                   | 24 |
| <i>3-2- Descripteurs moléculaires</i> .....                          | 24 |
| <i>3-2-1- Les descripteurs moléculaires théoriques</i> .....         | 24 |
| <i>3-2-2- Les descripteurs moléculaires empiriques</i> .....         | 25 |
| <i>3-3- Méthodes d'analyse des données</i> .....                     | 26 |
| <i>3-4- Interprétation et validation d'un modèle QSAR/QSPR</i> ..... | 26 |
| <i>3-4-1- Validation interne</i> .....                               | 27 |
| <i>3- 4- 2- Validation externe</i> .....                             | 28 |
| <i>3-4-3- Domaine d'applicabilité</i> .....                          | 29 |
| <i>Références Bibliographiques</i> .....                             | 30 |

# **TABLE DE MATIERE**

---

---

## **Plan du travail**

|  |    |
|--|----|
| <i>Introduction</i> .....                | 32 |
| <i>Protocol de plan du travail</i> ..... | 33 |

# *Introduction générale*

## ***Introduction générale***

---

### ***Introduction générale :***

Les glycosidases sont une famille d'enzymes présente chez presque tous les êtres vivants. Elles ont un rôle capital dans 8600 processus biologiques connus [1]. Parmi ces enzymes qui catalysent l'hydrolyse des liaisons glycosidique, l' $\alpha$ -glucosidase impliquée dans l'étape finale de la digestion des carbohydrates de l'alimentation de l'être humain en permettant la dégradation complète des oligosaccharides libérés par l' $\alpha$ -amylase salivaire et pancréatique et non absorbés par le corps. Ce processus permet la libération du glucose pour être ensuite véhiculé par le sang vers les organes[2].

L'inhibition de l' $\alpha$ -glucosidase bloque le processus de la dégradation des oligosaccharides, stoppant la digestion des carbohydrates[3], et permettant ainsi de contrôler le niveau du glucose dans le sang chez les diabétiques. Cet effet bénéfique, des inhibiteurs des glycosidases, sur la santé a attiré l'attention des chercheurs et des firmes pharmaceutiques pour développer et enrichir la classe thérapeutiques des médicaments anti diabète type II [4].

QSAR [5] a été largement utilisé pendant des années pour fournir une analyse quantitative de la structure et des relations d'activité des composés. Des méthodes statistiques sont appliquées dans la modélisation QSAR pour établir des corrélations entre les structures chimiques et leurs activités biologiques. Une fois validés, les résultats peuvent être utilisés pour prédire les activités de composés non testés. Récemment, la conception de médicaments assistée par ordinateur basée sur QSAR a été utilisée avec succès pour développer de nouveaux médicaments.[6]

Dans ce travail, nous sommes intéressés à l'étude de l'activité inhibitrice de l' $\alpha$ -glucosidase de molécules dérivées de flavonoïdes en utilisant l'analyse 2D-QSAR, dans le but de développer la relation entre la structure et l'activité de cette substance chimique active et ses dérivés.

Le manuscrit est divisé en deux chapitres :

## *Introduction générale*

---

Dans le premier chapitre présente un aperçu général sur les enzymes glycosidases, leur rôle biologique, leurs mécanismes d'action, un bref rappel sur la cinétique enzymatique montrant les différents types d'inhibition. Un rappel bibliographique sur les inhibiteurs les plus significatifs des glycosidases est présenté dans ce même chapitre.

Le deuxième chapitre intitulé « méthodologie de QSAR » présente la méthode QSAR, son principe, son application et les différentes étapes à respecter pour la réaliser.

Enfin, nous terminerons ce manuscrit par un plan de travail.



## *Introduction générale*

---

### **Références bibliographiques :**

[1] D. G. Naumoff, *Biochemistry (Moscow)*, 2011, 76, 622-635.

[2] R. Quezada-calvillo, L. Sim, Z. Ao, B.R. Hamaker, A. Quaroni, G.D. Brayer, E.E. Sterchi, C.C. Robayo-torres, D.R. Rose, B.L. Nichols, *J. Nutr.* 2008,138, 685–692.

[3] L. EL BLIDI, “Synthèse chimio-enzymatique et évaluation d’inhibiteurs potentiels de glycosidases, analogues de la valioline” Thèse de doctorat, Université BLAISE PASCAL, France, 2006.

[4] A. K. Ghose, V. N. Viswanadhan, Ed. Marcel Dekker, Inc. Etats Unies, 2001.

[5] Hansch C, Mahoney PP, Fujita T, Muir RM. *La nature*. 1962,194, 178-180.

[6] Ring CS, Sun E, McKerrow JH. *Proc.Natl. Acad. Sci. USA*. 1993, 90, 3583–3587.

***CHAPITRE I :***

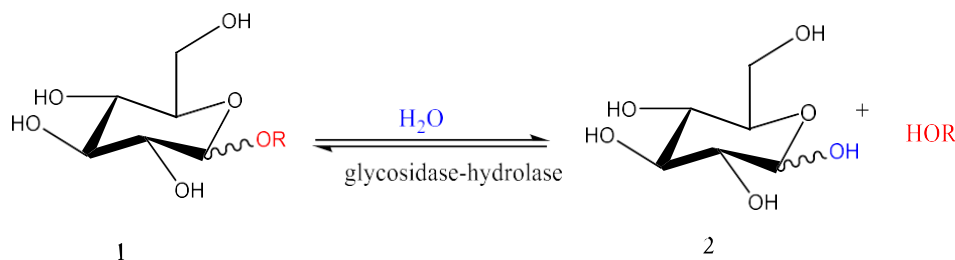
***LES GLYCOSIDASES ET LEURS***

***INHIBITEURS***

### 1- Les glycosidases :

#### Généralités :

Les glycosidases ou glycoside-hydrolases sont responsables de l'hydrolyse des oligo- et polysaccharides. Elles catalysent l'hydrolyse sélective des liaisons glycosidiques entre un sucre et un aglycone (R), qui est souvent un autre sucre (Schéma1).



*Schéma 1 : Hydrolyse simplifiée de la liaison glycosidique*[1].

Il existe différents types de glycosidases, qui sont spécifiques de l'unité glucidique hydrolysée, de sa série D ou L, de sa forme furanique ou pyranique et de la configuration de la liaison glycosidique  $\alpha$  ou  $\beta$ . Ainsi les  $\alpha$ -D-glucosidases hydrolysent les  $\alpha$ -D-glucosides, les  $\beta$ -D-galactosidases les  $\beta$ -D-galactosides etc. De plus, au sein de chaque famille, il existe de petites variations selon l'origine de l'enzyme[1].

On distingue également les exo-glycosidases et les endo-glycosidases. Les premières hydrolysent des liaisons glycosidiques terminales et les secondes des liaisons internes des chaînes polysaccharidiques.

Dans la nomenclature de l'Union Internationale de Biochimie et de Biologie Moléculaire (IUBMB) basée sur le type de réaction catalysée et la spécificité de substrat, chaque enzyme reçoit un numéro EC. Le premier chiffre de ce numéro classe l'enzyme suivant son activité. Le second chiffre désigne le type de fonction affectée par l'activité enzymatique. Le troisième et le quatrième chiffre permettent de préciser le substrat. Les numéros correspondant aux glycoside-hydrolases sont donc du type EC 3.2.1.x, "x" dépendant du substrat.

## ***Chapitre I: Les glycosidases et leurs inhibiteurs***

---

Depuis 1991, une autre classification de ces enzymes a été proposée par B. Henrissat[2], basée sur les séquences d'acides aminés, contient près de 200 familles de glycosidases, glycotransférases, polysaccharides lyases et carbohydate estérases[3].

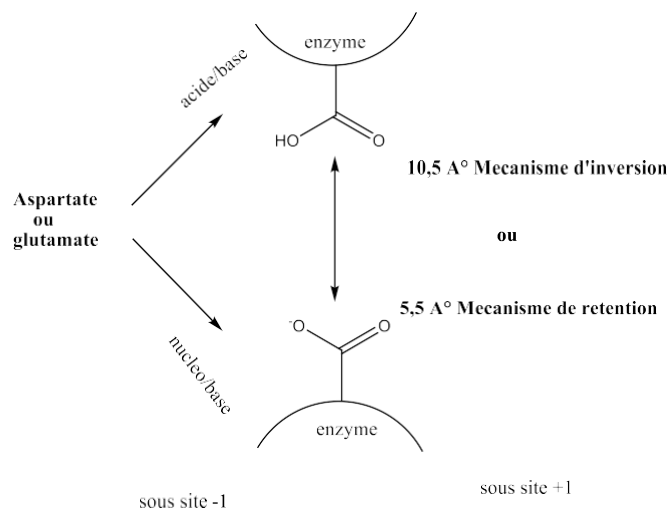
Cette classification diffère de la classification de l'IUBMB par le fait que des enzymes de spécificités différentes peuvent appartenir à la même famille, puisqu'il existe une relation directe entre similarités de séquences et repliements tridimensionnels. La classification CAZy (Carbohydrate-Active enZYmes) met ainsi en évidence les relations structure/fonction. Cette classification permet donc de prendre en considération la structure tridimensionnelle ainsi que le mécanisme moléculaire des glycoside-hydrolases.

### **Rôles des glycosidases :**

Les glycosidases sont des enzymes très répandues. Elles sont impliquées dans un grand nombre de processus biologiques fondamentaux et aussi variés que la digestion des polysaccharides alimentaires, la maturation des glycoprotéines ou bien le catabolisme des polysaccharides et des glycoconjugués. Les glycosidases sont donc essentielles à la croissance, ainsi qu'à la régulation et au développement cellulaire.

### **Mécanismes d'action des glycosidases :**

L'hydrolyse enzymatique des liaisons glycosidiques est une réaction acido-basique qui peut s'effectuer selon deux mécanismes catalytiques, proposés pour la première fois par Koshland en 1953[4]. Dans les deux cas, la réaction passe par un état de transition du type ion oxocarbenium. Cette réaction est généralement gouvernée par deux acides aminés (le plus souvent de type aspartate ou glutamate), l'un jouant le rôle de donneur de proton acide/base, l'autre agit en tant que nucléophile/base (**Schéma2**).

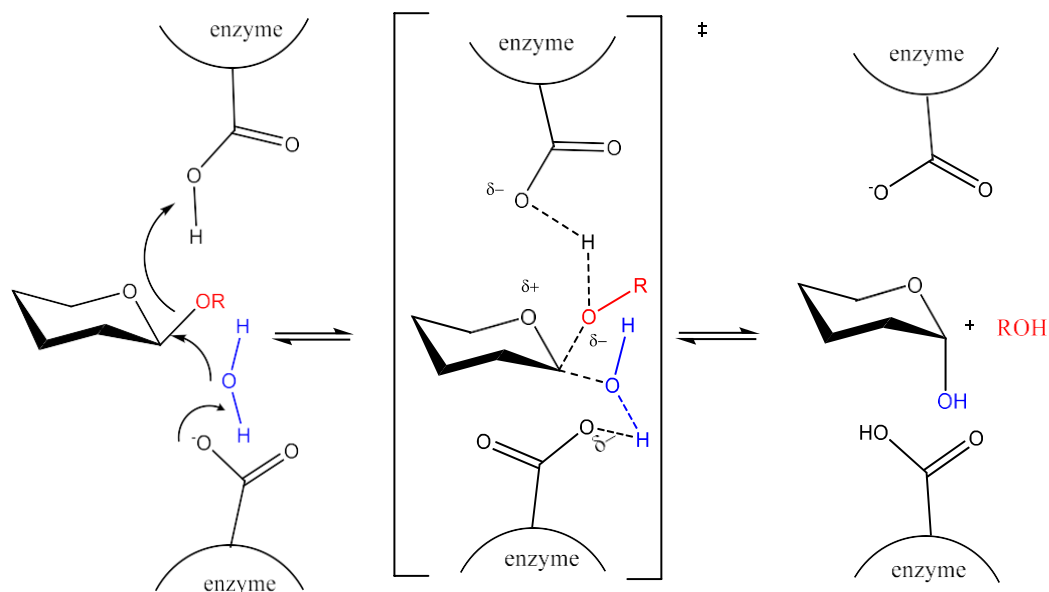


**Schéma 2 :** Site actif des glycoside-hydrolases[5].

La différence de mécanisme entre deux enzymes se traduit par des modifications structurales au sein du site actif, notamment au niveau de la distance entre les acides aminés catalytiques. En effet, les acides aminés catalytiques d'une enzyme agissant avec rétention de configuration sont distants en moyenne de 5,5 Å, alors que ceux d'une enzyme agissant avec inversion de configuration sont distants de 10,5 Å environ[5,6]. Cette distance plus importante est nécessaire au mécanisme d'inversion, puisqu'une molécule d'eau et le substrat doivent être présents simultanément entre ces deux acides aminés.

### Mécanismes avec inversion de configuration :

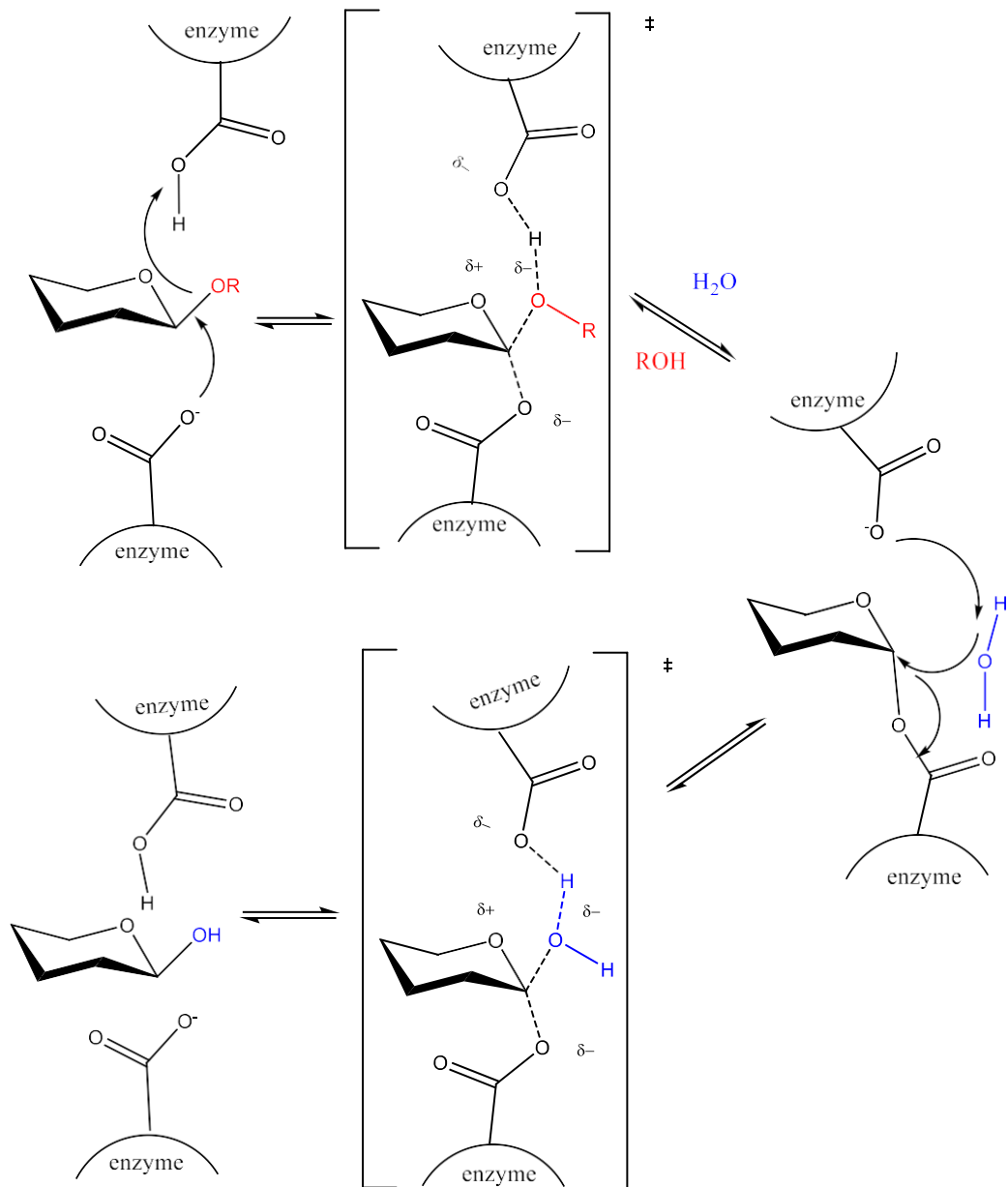
Hydrolyse avec inversion de configuration se déroule en une seule étape (**Schéma 3**). L'atome d'oxygène inter glycosidique est protoné par le résidu acide/base et le départ de l'aglycone s'accompagne de l'attaque d'une molécule d'eau activée par la base. Cette simple substitution conduit donc à un produit de configuration opposée à celle du substrat initial.



**Schéma 3:** Mécanisme générale des glycosidases avec inversion de configuration[7].

### Mécanismes avec rétention de configuration :

Ce mécanisme implique deux réactions d'inversion successives (**Schéma 4**). La première étape de l'hydrolyse est la protonation de l'atome d'oxygène inter glycosidique par le résidu acide/base et le départ de l'aglycone assisté par le nucléophile pour conduire à la formation d'un glycosyl-enzyme. Dans une seconde étape, l'intermédiaire covalent est hydrolysé par une molécule d'eau et libère ainsi un produit de même configuration que le substrat initial.



*Schéma 4: Mécanisme générale des glycosidases avec rétention de configuration[7].*

### Mécanismes particuliers :

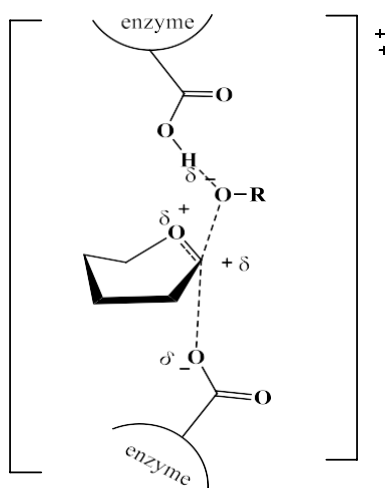
Il s'agit des cas particuliers où l'acide catalytique peut être remplacé par un phosphate inorganique. Différents groupes chimiques peuvent intervenir dans ce mécanisme et être à l'origine de la similarité de certaines glycosidases avec des enzymes présentant des activités apparemment très distinctes, comme par exemple : la maltose phosphorylase (EC 2.4.1.8, famille GH65), la sucrose phosphorylase (EC 2.4.1.7, famille GH13), ou la cellobiose phosphorylase (EC 2.4.1.20, famille GH94)[8,9]. Certaines glycosidases comme les

## Chapitre I: Les glycosidases et leurs inhibiteurs

myrosinases, les chitinases et les endo N-acétyl glucosaminidases présentent un mécanisme moléculaire d'hydrolyse n'impliquant l'intervention que d'un seul acide aminé catalytique[10]. D'autres impliquent des mécanismes d'oxydoréduction et/ou d'élimination[11].

### 1-4 Modèle général de l'état de transition :

La connaissance des mécanismes et donc des états de transition impliqués dans les réactions catalysées par des enzymes est fondamentale. En effet, de nombreuses stratégies de synthèse d'inhibiteurs sont orientées vers la création de molécules mimant les états de transition, ceux-ci étant considérés comme les structures les plus fortement liées par l'enzyme au cours du chemin réactionnel.



*Schéma 5 : Etat de transition lors de l'hydrolyse glycosidique.*

D'après cette représentation, deux caractéristiques de l'état de transition semblent essentielles : Une distorsion de cycle du substrat et une structure « pseudo-cationique » de type oxocarbémium.

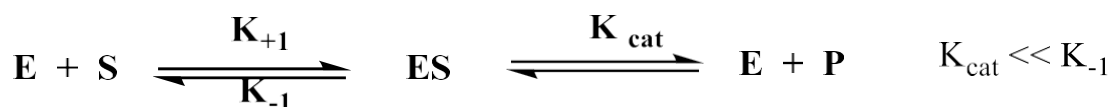
### 2- L'inhibition enzymatique :

L'étude cinétique de ces réactions permet de mieux comprendre les mécanismes de la catalyse et de déterminer l'activité des enzymes. En 1913, Leonor Michaelis et Maud Menten ont défini des équations permettant de décrire les réactions enzymatiques avec différents



## Chapitre I: Les glycosidases et leurs inhibiteurs

paramètres, dans lequel l'enzyme (E) en présence du substrat (S) forme un complexe dissociable (ES), qui libère le produit de la réaction (P) (**Schéma 6**).



*Schéma 6 : Mécanisme de la réaction enzymatique selon Michaëlis-Menten.*

L'enzyme en tant que catalyseur est en concentration plus faible que le substrat, ce qui conduit à un état stationnaire où la concentration complexe (ES) est constante. Ainsi la vitesse d'apparition du produit  $v$  est également constante :

$$v = v_{max} \times \frac{[S]}{K_M + [S]}$$

Avec  $K_M = \frac{K_{-1} + K_{cat}}{K_{+1}}$  : la constante de dissociation du complexe [ES] (appelée constante de

*Michaëlis-Menten*).

De façon générale, les inhibiteurs d enzyme ralentissent les réactions chimiques. On utilise deux paramètres pour décrire l'inhibition d'une enzyme par un inhibiteur:  $IC_{50}$  et le  $K_i$ [12]

Le  $K_i$  est la constante de dissociation du complexe (enzyme / inhibiteurs). Elle correspond à la concentration en inhibiteur nécessaire pour occuper 50% des sites de l'enzyme à l'équilibre. Cette constante est déterminée en faisant varier indépendamment la concentration en substrat et la concentration en inhibiteur.

L' $IC_{50}$  est la concentration en inhibiteur nécessaire pour diviser l'activité ou la vitesse de la réaction par 2. L' $IC_{50}$  et le  $K_i$  ne sont donc pas liés par une simple relation linéaire mais évoluent de façon similaire, leur relation dépend du mode d'inhibition.

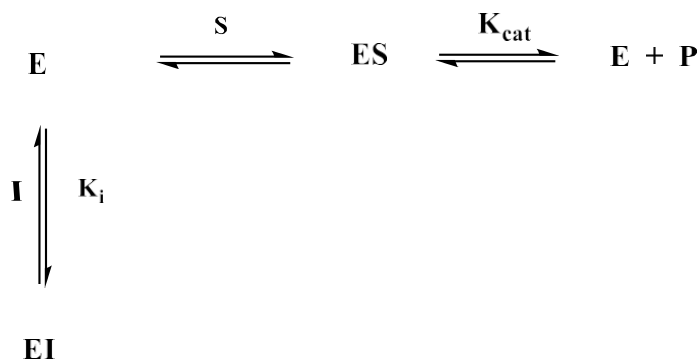
On distingue dans l'inhibition réversible trois types d'inhibiteurs, les inhibiteurs compétitifs, non compétitifs et incompétitifs.

## Chapitre I: Les glycosidases et leurs inhibiteurs

Pour connaître le type d'inhibition on peut tracer le graphe avec les courbes de qui sont caractéristiques l'inhibition, il s'agit de tracer  $1/v$  en fonction de  $1/(S)$  pour différents concentrations en inhibiteur (I), ou  $v$  est la vitesse de la réaction et (S) la concentration en substrat c'est pour cette méthode que l'on peut également déterminer le  $K_i$ .

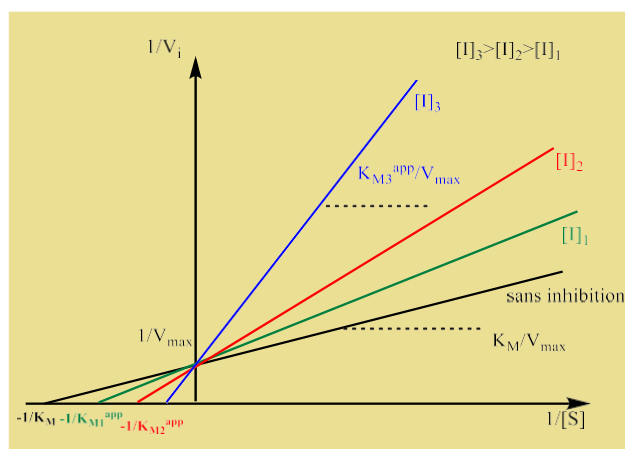
### **Inhibition compétitive :**

L'inhibition est compétitive si l'inhibiteur (I) se lie réversiblement à l'enzyme (E) et empêche le substrat (S) d'accéder au site actif (**Schéma 7**).



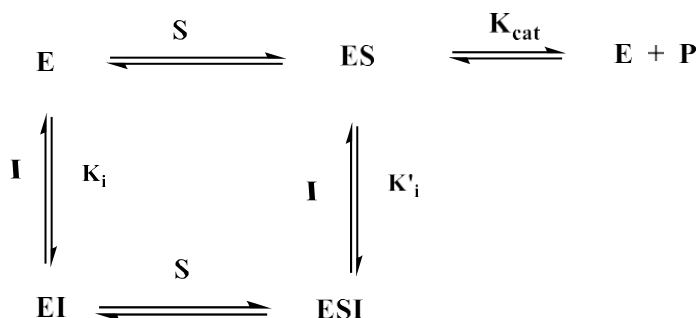
*Schéma 7: Mécanisme réactionnel de l'inhibition compétitive.*

Dans ce type d'inhibition les courbes de Lineweaver-Burk se coupent sur l'axe des ordonnées en  $1/V_{\text{max}}$ . La  $V_{\text{max}}$  ne varie pas en fonction de la concentration en inhibiteur (**Figure 1**).



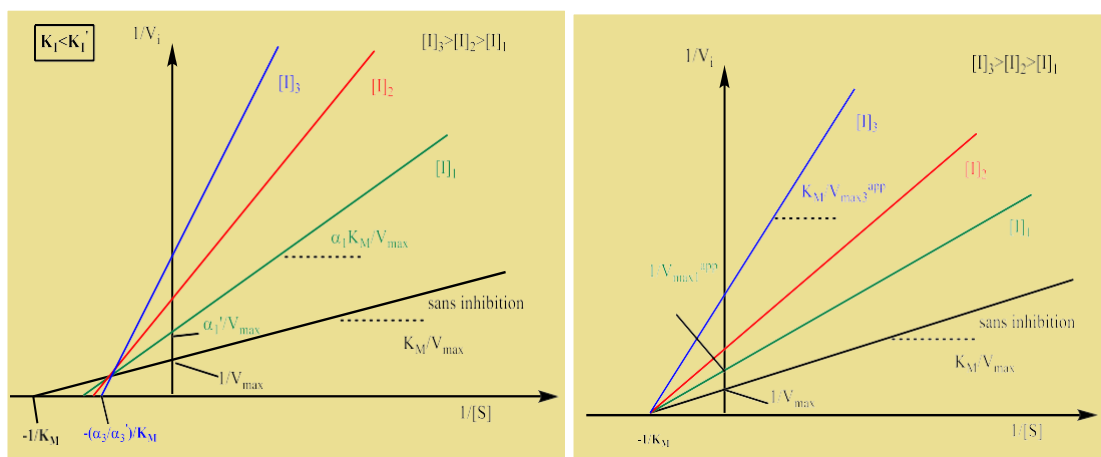
*Figure 1: Représentation graphique de Lineweaver-Burk pour l'inhibition compétitive.*





**Schéma 9:** mécanisme réactionnel de l'inhibition non compétitive.

Il existe deux types d'inhibition non compétitive : l'inhibition non compétitive pure et l'inhibition non compétitive mixte (**Figure 3**).



**Figure 3:** Représentation graphique de Lineweaver-Burk pour l'inhibition non compétitive.

### 3- Inhibiteurs des glycosidases :

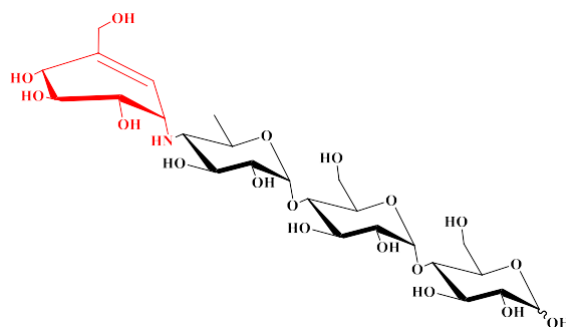
Les inhibiteurs de glycosidases qui ont été isolés ou synthétisés se sont souvent orientés vers la construction de molécules mimant l'état de transition, grâce à une distorsion de cycle, une charge positive ou en combinant les deux (**Figure 5**) [13]. Dans le reste de ce chapitre nous décrivons différents types chimiques d'inhibiteurs pour les  $\alpha$ -glucosidases comme les aminocyclitols, les iminosucres et les thiosucres.

#### Les aminocyclitols :

Les aminocyclitols sont analogues du sucre, possédant un atome d'azote exo cyclique protoné à PH physiologique. Parmi ces composés, l'acarbose (**Figure 4**) est un tétra

## Chapitre I: Les glycosidases et leurs inhibiteurs

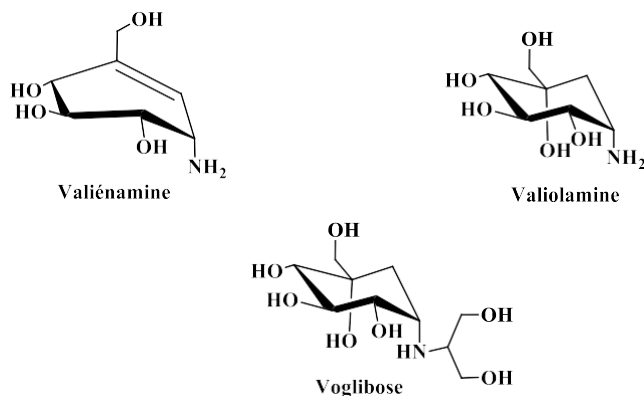
saccharide possédant une unité carbo sucre, inhibiteur puissant d' $\alpha$ -glucosidases intestinales[14].Isolé dans les années 70 d'une souche bactérienne *Actinoplanes sp.* EE50, Ainsi, ce composé a été mis à sur marche en 1990 à cet effet sous le nom de Glucor® ou Glucobay® par la société Bayer pour le traitement du diabète de type II.



**Figure 4 :** Structure chimique de l'acarbose.

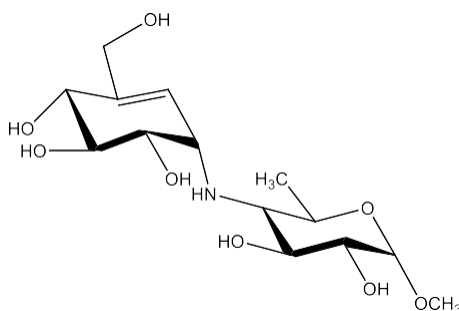
La similarité structurale de l'acarbose avec des oligosaccharides, en raison de ses résidus de glucose, est considérée comme étant responsable de la liaison de haute affinité pour les sites actifs des  $\alpha$ -glucosidases. L'incapacité des enzymes digestives à hydrolyser l'acarbose est due à la présence d'un pont amine, qui est considéré comme facteur déterminant de son effet inhibiteur.

Le cycle insaturé sur l'acarbose appelée valiénamine (**Figure 5**), est un bon inhibiteur d' $\alpha$ -glucosidases intestinales[15].Ce composé ainsi que la valiolamine (**Figure 5**) ont été isolés à partir d'une souche de *Streptomyces lygrosopicus*[16].Le voglibose, un des plus puissants inhibiteurs d' $\alpha$ -glucosidases ( $K_i = 3,8$  nM ; maltase de porc et  $K_i = 2,0$  nM ; saccharase de porc) (**Figure 5**). Il est commercialisé en Asie sous le nom Basen® par la société Takeda pour le traitement du diabète de type II.



**Figure 5 :** Structures chimiques de Valiénamine, Valiolamine et de Voglibose.

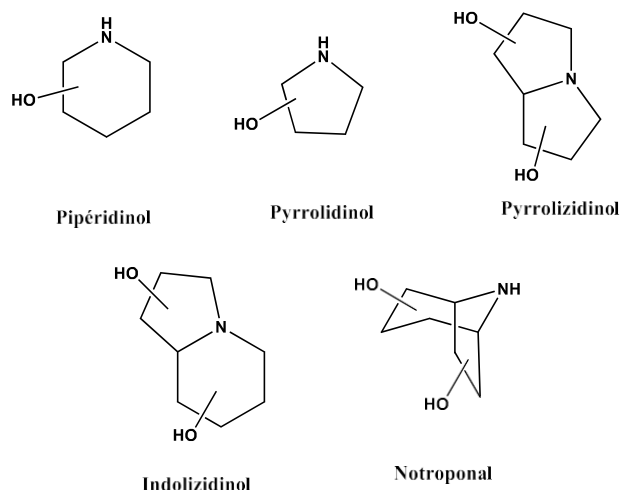
Parmi les analogues de l'acarbose il ya l' $\alpha$ -méthylacarviosine (**Figure 6**), un excellent inhibiteur d' $\alpha$ - glucosidases intestinales ( $IC_{50} = 0,16 \mu M$  ; saccharase de porc)[17].



**Figure 6 :** Structure chimique de l ' $\alpha$ -méthylacarviosine.

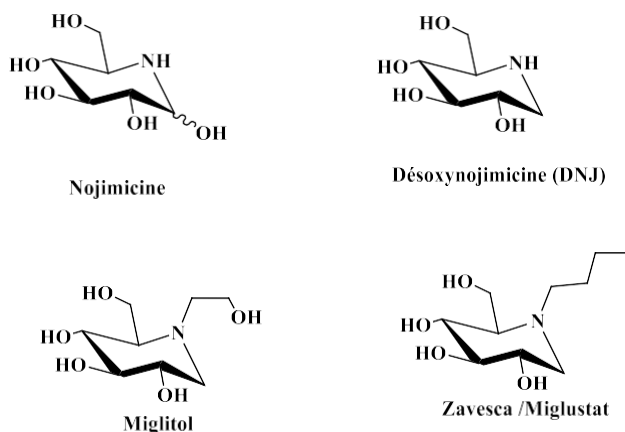
### Les aminosucres:

Les aminosucres sont des analogues de sucres dont l'oxygène endo cyclique est remplacé par un atome d'azote. Cet atome est protoné à pH physiologique et mime ainsi la charge positive. On peut les regrouper en 5 classes (**Figure 7**) qui se différencient par leur squelette (monocyclique ou bicyclique) et par la taille des cycles (5ou 6 membres).



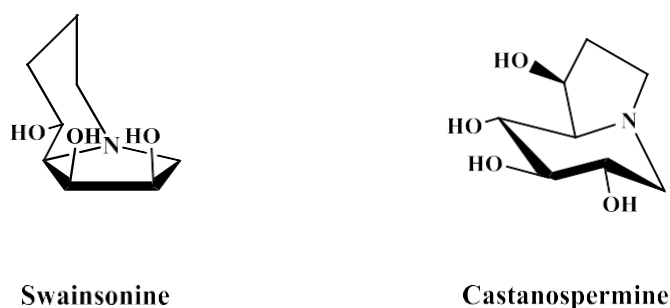
**Figure 7** : Structures de cinq grandes classes d'aminosucres.

Parmi ces aminosucres, la nojimyicine (**Figure 8**) a été découverte en tant qu'un premier azasucré naturel. Ce dernier, a été isolé d'une souche bactérienne *Streptomyces* par l'équipe d'*Inoye*[18]. En plus de ses propriétés antibiotiques, il inhibe les  $\alpha$ - et les  $\beta$ -glycosidases. Cependant ce composé est instable à cause de la présence du groupement hydroxyle porté par le carbone C1[19]. Ainsi, la désoxynojirimicine (DNJ) est un puissant inhibiteur d' $\alpha$  et de  $\beta$ -glucosidases [20,21]. En 1996 le miglitol ou N-hydroxyéthyl-1-désoxynojirimicine (**Figure 8**), puissant inhibiteur d' $\alpha$ -glucosidases intestinales[22], a été commercialisé sous le nom Diastabol® pour le traitement du diabète de type II. Le zavesca (miglustat, N-butyl-DNJ) est sur le marché depuis 2006 pour le traitement de la maladie de *Gaucher* de type I[23](**Figure 8**).



**Figure 8 :** Structures chimiques de quelques aminosucres inhibiteurs de glycosidases.

Les indolizidines poly hydroxylées, naturelles ou synthétiques, sont également de puissants inhibiteurs de glycosidases. Le premier alcaloïde de ce type découvert a été la swainsonine (**Figure 9**), isolé de la légumineuse *Swainsona canescens* en 1973, est un bon inhibiteur de l' $\alpha$ -mannosidase II située dans l'appareil de Golgi, enzyme impliquée dans la maturation des glycoprotéines[24]. Le castanospermine (**Figure 9**), isolé de *Castanospermum australe*[25] en 1981, est un inhibiteur d' $\alpha$ -et  $\beta$ -glucosidases. Le castanospermine pourrait se révéler très intéressant contre la maladie de *dengue*. En effet, en inhibant une  $\alpha$ -glucosidase de réticulum endoplasmique, il empêcherait la propagation du virus[26].



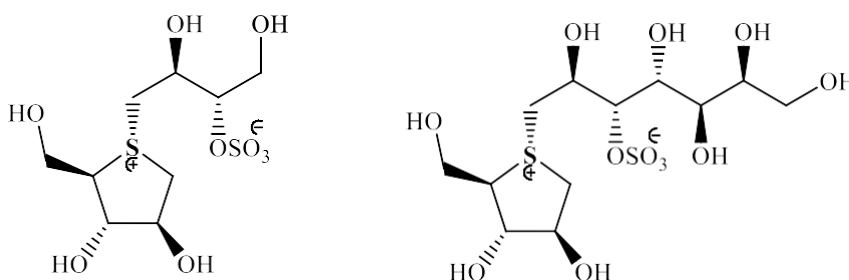
**Figure 9:** Structures chimiques de swainsonine et de castanospermine.



### Les thiosucres :

La famille des thiosucres, analogues de sucres dont l'oxygène endocyclique est remplacé par un soufre, a montré une activité inhibitrice puissante envers les  $\alpha$ -glucosidases avec la découverte du salacinol et le katalanol (**Figure 10**).

Le salacinol et le katalanol ont été isolés à partir de la plante *Salacia Reticulata*, que l'on trouve au sud de l'Inde et au Sri Lanka [27,28]. Les racines et les tiges de cette plante sont utilisées en médecine traditionnelle pour le traitement de diverses maladies dont le diabète. Ce n'est qu'en 1997 que Yoshikawa *et al*, ont montré que la fraction extraite au méthanol de cette plante inhibait *in vitro* la saccharase ( $IC_{50} = 26 \mu\text{g/mL}$ ) et la maltase ( $IC_{50} = 35 \mu\text{g/mL}$ ) intestinales de rat. Cela a permis de montrer que les propriétés traditionnelles antidiabétiques de cette plante étaient vraisemblablement dues à une activité inhibitrice vis-à-vis d' $\alpha$ -glucosidase intestinales [27]. Ces différentes séparations leur ont permis d'isoler le salacinol [29] et le katalanol [28], deux nouveaux inhibiteurs d' $\alpha$ -glucosidases ( $K_i = 0,93$  et  $0,54 \mu\text{M}$ ; maltase d'intestin grêle de rat) [28], possédant une structure zwitterionique inédite contenant un cation sulfonium stabilisé par un anion sulfate (**Figure 10**). Ces deux composés sont également isolés à partir d'autres variétés de *Salacia*: *Salacia oblonga* en Inde [30] et *Salacia chinensis* en Thaïlande.



**Figure 10** : Structures chimiques de salacinol et de kotalanol.

## Chapitre I: Les glycosidases et leurs inhibiteurs

---

### Références bibliographiques :

- [1] S. David, *InterEditions / CNRS Ed.* (1995) 58–61.
- [2] B. Henrissat, *M. Vegetales, F. Grenoble, Biochem. J.* 280 (1991) 309–316.
- [3] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic Acids Res.* 37 (2009) D233–D238.
- [4] J. D. E. Koshland, *Biol.Revs. Cambridge Phil. Soc.* 236 (1953) 416–436.
- [5] J. D. McCarter, G.S. Withers, *Curr Opin Struct Biol.* 4 (1994) 885–92.
- [6] Q. Wang, R.W. Graham, D. Trimbur, R.A.J. Warren, S.G. Withers, *J. Am. Chem. Soc.* 116 (1994) 11594–11595.
- [7] L.J. Pearson MSM, Mathe-Allainmat M, Fargeas V, *Eur J Org Chem.* (2005) 2159–91.
- [8] A. Vasella, G.J. Davies, M. Böhm, *Curr Opin Chem Biol.* 6 (2002) 619–629.
- [9] J.H. Davies, B.A.J. Evans, M.E.M. Jenney, J.W. Gregory, *J. Clin. Endocrinol. Metab.* 88 (2003) 6088–6097.
- [10] C.S. Rye, S.G. Withers, *Curr Opin Chem Biol.* 4 (2000) 573–580.
- [11] V.L.Y. Yip, A. Varrot, G.J. Davies, S.S. Rajan, X. Yang, J. Thompson, W.F. Anderson, S.G. Withers, *J. Am. Chem. Soc.* 126 (2004) 8354–8355.
- [12] B.T. Burlingham, T.S. Widlanski, *J. Chem. Educ.* 80 (2003) 214–218.
- [13] N. Asano, R. Nash, R. Moylneux, G. Fleet, *Tetrahedron:Asymmetry.* 11 (2000) 1645–80.
- [14] E. Truscheit, W. Frommer, B. Junge, L. Muller, D.D. Schmidt, W. Wingender, *Angew. Chemie Int. Ed. English.* 20 (1981) 744–761.

## ***Chapitre I: Les glycosidases et leurs inhibiteurs***

---

- [15] X. Chen, Y. Fan, Y. Zheng, Y. Shen, *Chem. Rev.* 103 (2003) 1955–1977.
- [16] H. Boo, S. Hwang, C. SikBae, Su-H. Park, B. GuHeo, S. Gorinstein, *Indust, Crop.Prod.* 40 (2012) 129–135.
- [17] S. Horii, H. Fukase, T. Matsuo, Y. Kameda, N. Asano, K.J. Matsui, *J.Med. Chem.* 29 (1986) 1038–1046.
- [18] N.T. Inouye S, Tsuruoka T, *J Antibiot.* 19 (1966) 288–92.
- [19] N.T. Inouye S, Tsuruoka T, Ito T, *Tetrahedron.* 24 (1968) 2125–44.
- [20] J.E. Legler G, *Carbohydr Res.* 128 (1984) 61–72.
- [21] C.W. Ekhardt, M.H. Fechter, P. Hadwiger, E. Mlaker, A.E. Stütz, A. Tauss, T.M. Wrodnigg, Wiley-VCH. (1999) 253–390.
- [22] K. Yasuda, H. Kizu, T. Yamashita, Y. Kameda, A. Kato, R.J. Nash, G.W.J. Fleet, R.J. Molyneux, N. Asano, *J.Nat. Prod.* 65 (2002) 198–202.
- [23] N. Asano, H. Kizu, K. Oseki, E. Tomioka, K. Matsui, M. Okamoto, M.J. Baba, *J.Med. Chem.* 38 (1995) 2349–2356.
- [24] A. Kato, E. Kato, I. Adachi, R.J. Molyneux, A.A. Watson, R.J. Nash, G.W.J. Fleet, M.R. Wormald, H. Kizu, K. Ikeda, N. Asano, *Tetrahedron Asymmetry.* 14 (2003) 325–331.
- [25] L.D. Hohenschutz, A.E. Bell, P.J. Jewess, D.P. Leworthy, R.J. Pryce, A. Edward, C. Jon, *Phytochemistry.* 20 (1981) 811–814.
- [26] M.S.D. K. Whitby, T. C. Pierson, B. Geiss, K. Lane, E. Engle, Y. Zhou, R. W. Doms, *Castanospermine, J. Virol.* 79 (2005) 8698–8706.
- [27] M. Yoshikawa, T. Murakami, H. Shimada, H. Matsuda, J. Yamahara, G. Tanabe, O.

## ***Chapitre I: Les glycosidases et leurs inhibiteurs***

---

*Muraoka, Salacinol, Tetrahedron Lett. 38 (1997) 8367–8370.*

[28] *M. Yoshikawa, T. Murakami, K. Yashiro, H. Matsuda, Kotalanol, Chem Pharm Bull. 46 (1998) 1339–40.*

[29] *M. Yoshikawa, T. Morikawa, H. Matsuda, G. Tanabe, O. Muraoka, Bioorganic Med. Chem. 10 (2002) 1547–1554.*

[30] *H. Matsuda, T. Murakami, K. Yashiro, J. Yamahara, M. Yoshikawa, Chem Pharm Bull. 47 (1999) 1725–9.*

## ***CHAPITRE II :***

# ***METHODOLOGIE DE QSPR/QSAR***

### 1- Introduction à la modélisation QSPR/QSAR :

Un modèle QSPR/QSAR est une formule mathématique qui permet de relier, d'une manière quantitative, la structure d'une molécule à une propriété ou à une activité donnée. Les méthodes QSPR/QSAR sont anciennes, elles remontent au 19<sup>ème</sup> siècle [1,2].

En effet, les premiers travaux utilisant la méthodologie QSPR/QSAR telle qu'employée actuellement sont dus à Hansch [3] et Free et Wilson [4]. D'un côté, des modèles reliant directement l'activité biologique des composés avec les propriétés hydrophobes, électroniques et stériques à l'échelle moléculaire ont été proposés par Hansch. D'un autre côté, des modèles empiriques, dits de contributions de groupes, pour l'étude de l'activité biologique ont été développés par Free et Wilson.

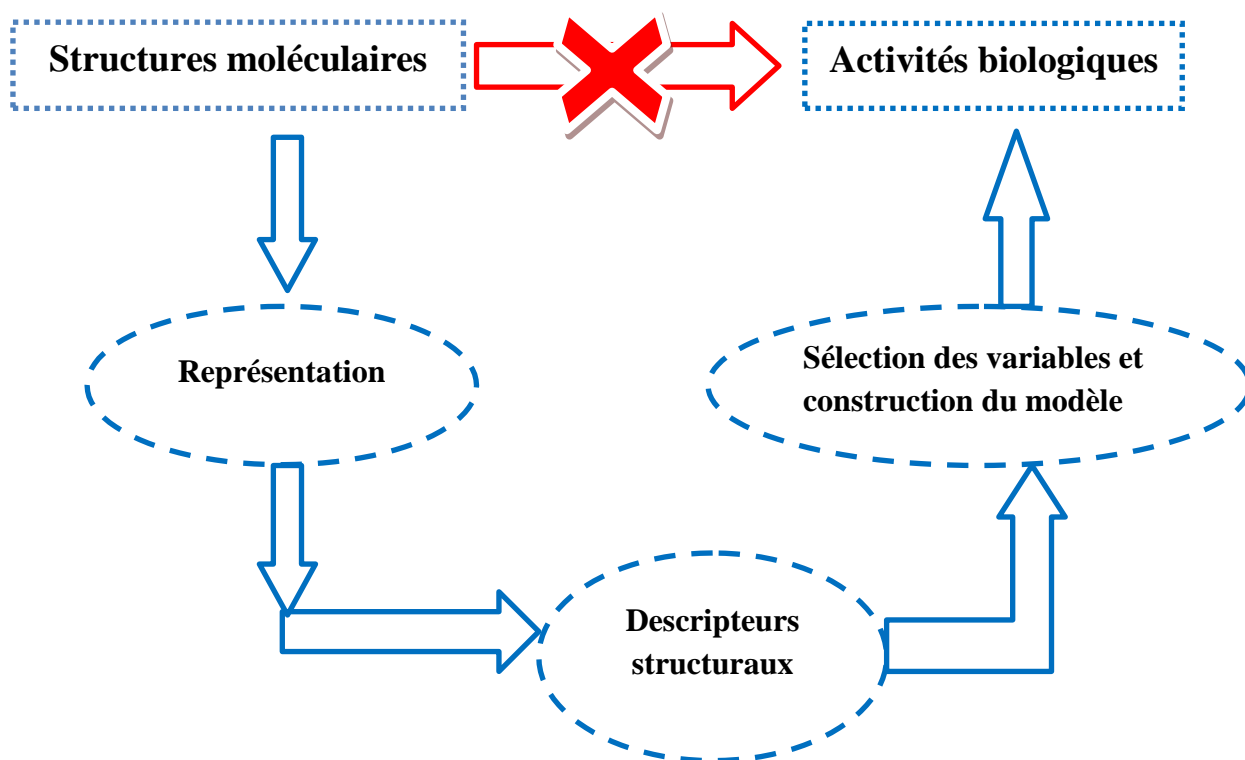


Figure 1 : Présentation des étapes aboutissant.

### **2- Principe des méthodes QSPR/QSAR:**

Le principe des méthodes QSPR/QSAR est de trouver une relation mathématique reliant de manière quantitative des propriétés moléculaires, appelées descripteurs, avec une propriété ou une activité biologique pour une série de composés chimiques similaires à l'aide des méthodes d'analyses de données. La forme générale d'un tel modèle est la suivante :

$$\text{Propriété/ Activité} = f(D_1, D_2, \dots, D_n, \dots) \quad (1)$$

$D_1, D_2, \dots, D_n$  sont des descripteurs des structures moléculaires.

### **3- Méthodologie générale d'une étude QSPR/QSAR :**

En générale pour prédire l'activité biologique par la méthode QSAR/QSPR, la construction du modèle mathématique est effectuée en 4 étapes de base (Figure 2).

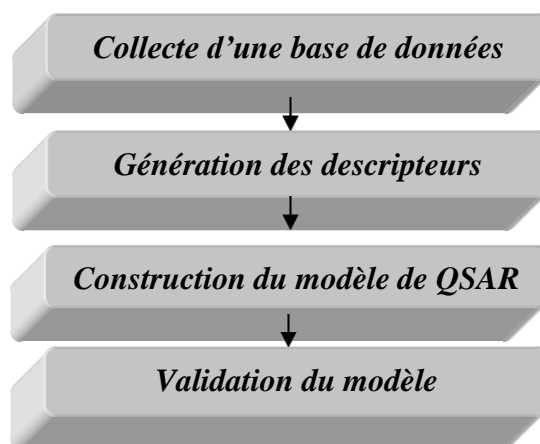
1- Collecte d'une base de données

2- Recherche de descripteurs adéquats pour l'activité étudiée

3- Le choix d'une méthode d'analyse des données

4- Validation du modèle

- ✓ Validation interne : on utilise la série d'apprentissage constituée de 2/3 de la base de données. Ce type de validation a pour but de vérifier la stabilité et le pouvoir explicatif du modèle QSAR retenu.
- ✓ Validation externe : on utilise la série de test constituée généralement de 1/3 de la base de données. Le but de cette validation est de vérifier le pouvoir prédictif du modèle élaboré.



**Figure 2 :** Procédure d'obtention et de validation d'un modèle QSPR/QSAR.

### 3-1- Bases de données :

Un modèle QSPR/QSAR est très dépendant des données expérimentales. Le choix de la base de données est décisif dans le développement de tel modèle. Une base de données doit être composée de données expérimentales aussi fiables que possible obtenues en suivant un protocole unique puisque les erreurs sur celles-ci se propageront dans le modèle final.

Plusieurs bases de données sont disponibles sur le net, on peut citer :

ZINC : <http://zinc.docking.org/catalogs/molport>

ECHA : <http://echa.europa.eu/fr>

Oracdatabase : <http://oracdatabase.com/list-of-foods/search/>

### 3-2- Descripteurs moléculaires :

Définition d'un descripteur: c'est l'information contenue dans la structure d'une molécule.

#### 3-2-1- Les descripteurs moléculaires théoriques :

*Les descripteurs 1D* : sont remportés directement à partir de la formule brute de la molécule. Exemple: le nombre d'atomes et la masse moléculaire,... etc.

*Les descripteurs 2D* : la formule développée d'une molécule peut nous offrir ce type de descripteurs. On distingue :

- *Les indices 2D constitutionnels* :

- *Les indices 2D topologiques* : Ils donnent des informations sur la taille, la forme globale et les ramifications d'une structure. Exemple: indice de Wiener [5], indice de Randić [6], indice de connectivité de valence de Kier-Hall [7]...etc

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

- *Les descripteurs 3D géométriques* : les plus importants sont le volume moléculaire, la surface accessible au solvant, le moment principal d'inertie.



## Chapitre II : Méthodologie de QSPR / QSAR

• Les *descripteurs 3D électroniques* : permettent de quantifier les différents types d'interactions inter- et intramoléculaires, de grande influence sur l'activité biologique des molécules. Le calcul de la plupart de ces descripteurs nécessite la recherche de la géométrie pour laquelle l'énergie est minimale, et fait souvent appel à la chimie quantique.

Exemple:

| Descripteur       | Définition                                    |
|-------------------|---|
| E <sub>LUMO</sub> | Energie de l'orbitale la plus basse vacante.  |
| E <sub>HOMO</sub> | Energie de l'orbitale la plus haute occupée   |
| $\mu$             | Potentiel chimique                            |
| $\eta$            | Dureté chimique                               |
| $\sigma$          | Mollesse chimique                             |
| $\omega$          | Indice d'électrophile globale                 |
| AEI               | Indice de l'énergie d'activation              |
| MSD               | Densité de Spin de Mulliken                   |
| E <sub>SOMO</sub> | Energie de l'orbitale occupée par un électron |

*Tableau.1. Exemples de descripteurs électroniques de la chimie quantique*

**Descripteurs thermodynamiques:** Les descripteurs thermodynamiques sont calculés sur la base de la fonction de partition totale Q de la molécule. Exemple: l'enthalpie de formation, l'enthalpie libre...

**3-2-2- Les descripteurs moléculaires empiriques :** Il existe plusieurs descripteurs qui peuvent être mesurés expérimentalement tel que le coefficient de partage (octanol-eau) (logP).

- **Coefficient de partage (octanol-eau)  $K_{ow}$ , log P :** Le coefficient de partition est défini comme le rapport de la concentration du soluté dans la phase huileuse  $C'$  à la concentration du soluté non ionisé dans la phase aqueuse  $C$ , à l'équilibre.

Il est largement utilisé dans des études de relations structure-activité quantitatives (QSAR<sub>s</sub>) dans les sciences pharmaceutiques, biochimiques, toxicologiques et dans les sciences de l'environnement.

**Détermination expérimentale de log P :** La méthode expérimentale utilisée pour mesurer le coefficient de partage d'une molécule est la méthode des flacons agités ou « shake-flask ». Cette méthode est préconisée comme procédure standard de caractérisation par l'OCDE [8] (Organisation de Coopération et de Développement Economique).

**Méthodes de calcul et d'estimation de log P :** Plusieurs méthodes d'estimation de log P ont été décrites dans la littérature depuis plusieurs décennies [9]. Les plus anciennes sont des méthodes fragmentales dans lesquelles une molécule est divisée en fragments prédéfinis et les contributions correspondantes sont sommées pour conduire à une valeur estimée du log P. Parmi ces méthodes on peut citer : Méthode de Hansch [10], Méthode de Rekker [11], La méthode de Ghose et Viswanadhan [12], La méthode de Klopman et Iroff [13], La méthode de Bodor [14] Plusieurs logiciels sont utilisés pour l'estimation de log P, on peut citer : ACD/LABS, ClogP, Hyperchem, KOWWIN, ChemSpider (online).

### **3-3- Méthodes d'analyse des données**

Une méthode d'analyse des données est nécessaire pour élaborer un modèle QSPR/QSAR. Il existe plusieurs méthodes pour construire un modèle et analyser les données statistiques de ce dernier, ces méthodes sont disponibles dans des logiciels tels que, Excel, Origin, Microcal, Minitab, Statistica, SPSS, R,...

### **3-4- Interprétation et validation d'un modèle QSAR/QSPR**

Une fois développé, le modèle doit être valide afin de vérifier sa stabilité, son pouvoir explicatif et son pouvoir prédictif.

### **3-4-1- Validation interne**

La méthode de la validation croisée, une méthode de validation interne, est la plus utilisée dans la plupart du temps dans QSAR. Elle est réalisée par différentes procédures: Leave-One-Out (*LOO*) et Leave-n-Out (*LMO*) où « n » représente le nombre de composés éliminés.

La procédure Leave-One-Out retire successivement une molécule de l'ensemble d'apprentissage (TSET). Un modèle QSAR est construit alors sur un ensemble de m-1 de composés et la molécule retirée est prédite par le modèle formé, « m » représente le nombre de molécules utilisées dans la construction du modèle (TSET). Cette procédure est répétée m fois afin de prédire les propriétés de toutes les molécules.

La procédure Leave-n-Out correspond à un découpage en plusieurs parties de l'ensemble de données. A tour de rôle, une partie de l'ensemble de données est attribuée pour un ensemble de test interne. Les autres constituent l'ensemble d'entraînement. Les molécules de chaque groupe éliminé sont prédites par le modèle formé. Cette procédure est répétée p fois pour prédire les propriétés de toutes les molécules, p est le nombre des groupes de molécules éliminés [15].

D'autres méthodes sont utilisées pour faire la validation interne, tel que la hasardisation de la réponse (Y-Randomization).

#### ***Y-Randomization***

Le test de randomisation permet d'affirmer que la corrélation de chance ne joue aucun rôle durant le développement du modèle. Ces tests sont exécutés quantitativement avec les modèles de QSAR et qualitativement avec les modèles issus de la classification. Les observations sont aléatoirement désorganisées dix fois. C'est-à-dire que la colonne des observations (propriétés) sera changée aléatoirement, en revanche la colonne des descripteurs

reste inchangée. A la fin, on obtient dix modèles avec des caractéristiques statistiques spécifiques [16].

La supposition sous-jacente de l'essai de randomisation est la suivante :

Si les capacités prédictives du modèle ne sont pas dues aux corrélations de chance, alors la désorganisation aléatoire des observations conduira à des modèles (quantitatifs ou qualitatifs) de prévisions faibles, et vice versa.

La validation externe du modelé est indispensable pour vérifier le pouvoir prédictif d'un modelé QSPR/QSAR. D'après certains chercheurs, la validation externe est une norme et une partie obligatoire dans la modélisation QSPR/QSAR [17, 18].

### **3- 4- 2- Validation externe**

Cette méthode consiste à prédire la propriété/activité d'une série de molécules appelée généralement série de test qui ne sont pas dans la série de développement du modelé, cette validation est caractérisée par les paramètres  $R^2$  (test) et  $R^2_{cv}$  (test).

Récemment plusieurs études [19,20] ont montré l'insuffisance des paramètres  $R^2$  (test),  $R^2_{cv}$  (test) pour vérifier la capacité prédictive 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 Tropsha» (Tropsha criteria) [20].

#### **Critères de validation Externe (série de test)**

$$R^2 > 0.7 \text{ (critère 1)}$$

$$R^2_{cv} > 0.6 \text{ (critère 2)}$$

$$\frac{|R^2 - R^2_{cv}|}{R^2} < 0.1 \quad \text{et} \quad 0.85 \leq k \leq 1.15 \quad \text{(critère 3)}$$

$$\frac{|R^2 - R^2_{cv}|}{R^2} < 0.1 \quad \text{et} \quad 0.85 \leq k' \leq 1.15 \quad \text{(critère 4)}$$

$$R^2 - R^2_{cv} \leq 0.3 \quad \text{(critère 5)}$$

Avec

$R^2$ : coefficient de corrélation pour les molécules de la série de test.

$R^2_0$ : coefficient de corrélation entre les valeurs prédites et expérimentales pour la série de test.

$R'^2_0$ : coefficient de corrélation entre les valeurs expérimentales et prédites pour la série de test.

$k$ : constante de la droite (à l'origine) de corrélation (valeurs prédites en fonction des valeurs expérimentales)

$k'$ : constante de la droite (à l'origine) de corrélation (valeurs expérimentales en fonction des valeurs prédites).

### **3-4-3- Domaine d'applicabilité**

Un modèle QSAR/QSPR ne peut pas être considéré comme un modèle universel, parce qu'il est développé sur un nombre limité de composés qui ne couvrent pas tout l'espace chimique. Par conséquent l'activité/propriété prédite d'un composé, chimiquement dissimilaire au jeu d'apprentissage, ne pourra pas être considérée fiable [21]. Le domaine d'applicabilité (DA) permet de définir la zone dans laquelle un composé pourra être prédit avec confiance. Le DA correspond donc à la région de l'espace chimique incluant les composés de la série d'apprentissage et les composés similaires, proches dans ce même espace [22].

### **Références Bibliographiques**

- [1] A. Crum Brown, T.R. Fraser. Trans. Roy.Soc. Edinburgh. 1868, 25, 151.
- [2] CrumBrown,A.and Fraser,T.R.J.Anat.Physiol.1868, 2, 224.
- [3] C. Hansch, A. Leo, D. Hoekmann. American. Chemical. Society. 1995.
- [4] S.M. Free, J.W. Wilson. J. Med. Chem., 1964, 7, 395.
- [5] H. Wiener. J. Am. Chem. Soc. 1947, 69, 17.
- [6] M. Randic. J. Am. Chem. Soc. 1975, 97, 6609.
- [7] L.B. Kier, L.H. Hall, W.J. Murray. J. Pharm. Sci. 1975, 64,1971.
- [8] OECD Guidelines for Testing of Chemicals No. 107, OECD,Paris, 1992.
- [9] B. Lee, F.M. Richards. J. Mol. Biol. 1971, 55, 379.
- [10] R.S. Pearlman, Eds., Pergamon, New York, 1986, 3.
- [11] R. F. Rekker, H. M. de Kort. (1979). Eur. J. Med. Chem. – Chim. Ther. 1979, 14, 479.
- [12] Ghose AK, Crippen GM. J. Chem. Inf. Comput. Sci. 1987, 27,21.
- [13] G.Klopman, L.D. Iroff, J.Comp.chem. 1981, 2, 157.
- [14] Bodor N, Buchwald P. J. Phys. Chem B. 1997, 101, 3404
- [15] P. Bultinck; H. D. Winter; W. Langenaeker; J. P. Tollenaere: "Computational Medicinal Chemistry for Drug Discovery" Ed. Marcel Dekker. Etats Unies, 2004.
- [16] L. Zhang; H. Zhu; T. I. Oprea; A. Golbraikh; A. Tropsha Pharm. Res. 25 (8)(2008) 1902 -1914.
- [17] A. Tropsha, P. Gramatica, V.K. Gombar, QSAR Comb.Sci. 2003, 22, 69
- [18] P.P. Roy, S. Paul, I. Mitra, K. Roy, Molecules, 2009, 14, 1660.
- [19] A. Golbraikh, A. Tropsha, Beware of q<sup>2</sup>!, J. Mol. Graph. Model. 2002, 20, 269.
- [20] T. M. Martin, P. Harten, D. M. Young, E.N. Muratov, A. Golbraikh, H. Zhu and A.Tropsha,J. Chem. Inf. Model. 2012, 52, 2570.

## *Chapitre II : Méthodologie de OSPR / OSAR*

---

- [21] I.V. Tetko, I. Sushko, A.K. Pandey, H. Zhu, A. Tropsha, E. Papa, T. Oberg, R. Todeschini, D. Fourches, Varnek A. J. Chem. Inf. Model. 2008, 48, 1733.
- [22] J. Jaworska, N.N. Jeliaskova, T. Aldenberg Altern. Lab. Anim 2005, 33, 445.

# *Plan de travail.*



### ***1-Introduction***

Le modèle **QSAR** (quantitative structure-activité Relationship) étudie la corrélation basée sur un calcul statistique, de paramètres physico-chimiques avec l'activité biologique.

La comparaison de l'activité biologique de certaines molécules et de leur structure a permis d'établir dans de nombreux cas des corrélations entre les paramètres structuraux et les propriétés d'une molécule. L'association de la variation de l'activité aux paramètres a permis d'obtenir un système d'équations qui donne, pour une série chimique donnée et pour une activité définie, une équation de corrélation.

L'intérêt essentiel de cette équation est qu'elle doit permettre de déterminer la valeur des paramètres qui correspondent à une activité maximale et ainsi de prédire l'activité des molécules qui n'ont pas encore été synthétisées.

### **2-Protocol de génération d'un modèle QSAR:**

Dans cette partie nous présentons un plan de travail sur l'étude linéaire de la relation quantitative structure-activité inhibitrice de l' $\alpha$ -glucosidase par les dérivés de flavonoïdes.

## Plan du travail

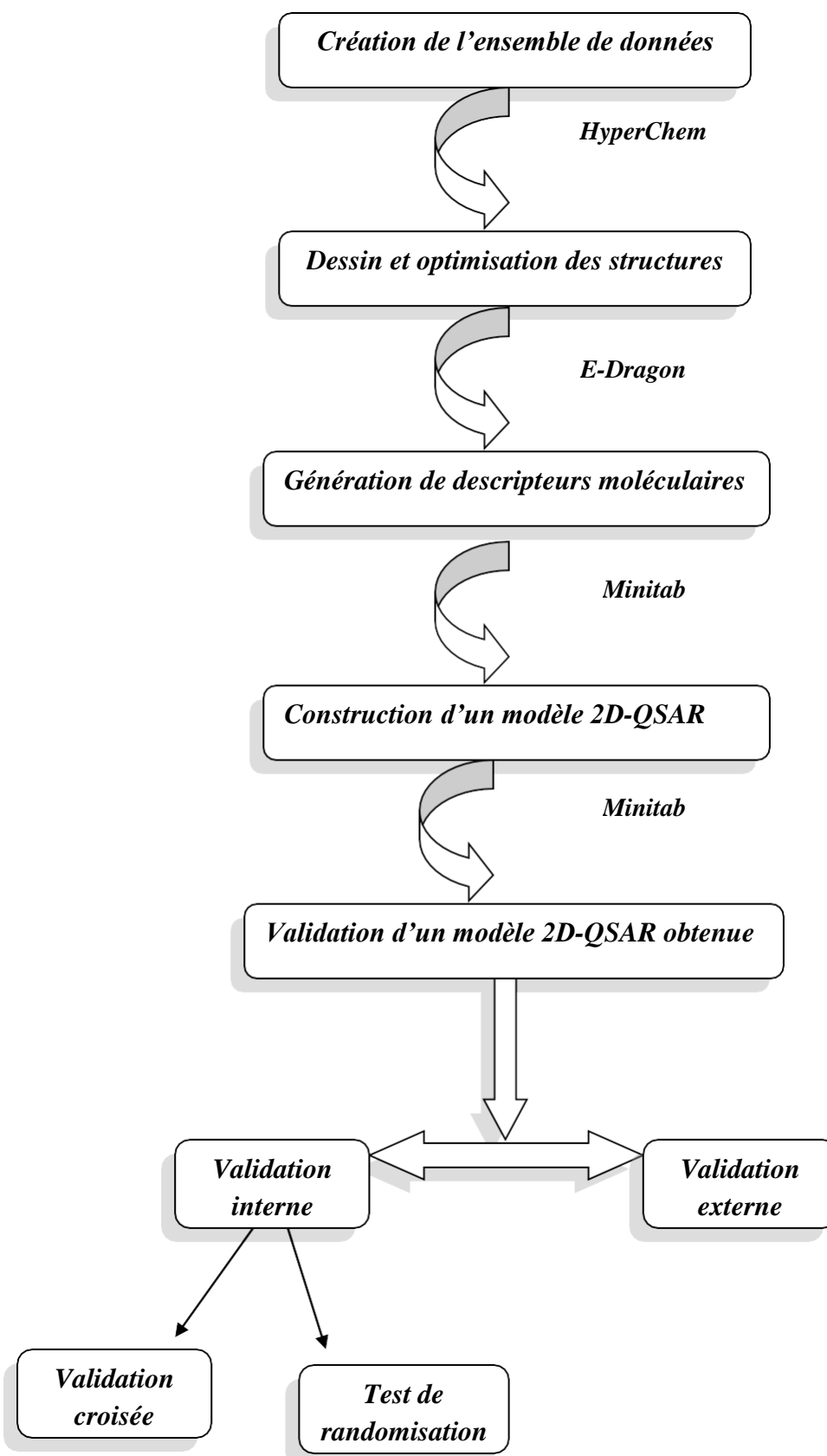


Schéma 1 : Représentation du protocole appliqué pour la génération d'un modèle 2D-QSAR.