

Drug discovery: Historique et développement :

Par le passé, la transmission des savoirs ancestraux en médecine traditionnelle a été, pendant longtemps, un facteur déterminant de la découverte d'un grand nombre de médicaments. Grâce à l'identification de principes actifs extraits de substances naturelles d'origine végétale, minérale ou animale qui étaient historiquement utilisées et sélectionnées par l'observation empirique de leurs effets sur le cours des maladies, l'exploration de ces savoirs ancestraux est toujours d'actualité et constitue une discipline à part entière, l'ethnopharmacologie¹. Ainsi, la morphine a été isolée du pavot au début du XIX^{ème} siècle et ses effets sont décrits par Friedrich Wilhelm ce qui a marqué un tournant historique dans la recherche de nouveaux médicaments².

Cinquante ans plus tard, l'essor des médicaments issus de synthèses débute avec la mise au point de l'acide acétylsalicylique puis sa commercialisation plus tard sous le nom de marque « Aspirine » par les Laboratoires Bayer³. Depuis, plusieurs médicaments ont été identifiés ainsi. Cependant, la nécessité de développer d'autres sources de découverte et d'établir une démarche de recherche systématique est rapidement apparue suite aux nombreuses avancées dans le domaine de la synthèse chimique et de la pharmacologie en permettant l'élaboration de nouveaux médicaments de plus grande efficacité⁴. Les découvertes de médicaments les plus marquantes ont eu lieu au cours des XIX^{ème} et XX^{ème} siècles et sont résumées dans figure 1⁵.

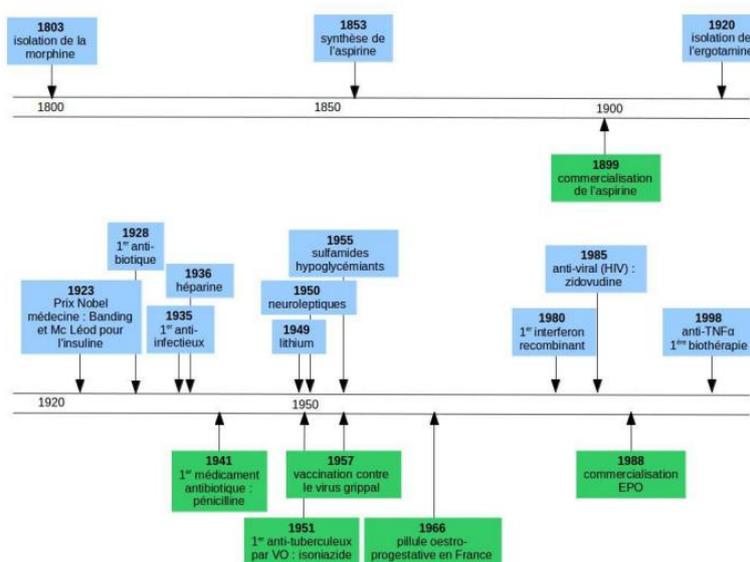


Figure II. 1. Quelques grandes étapes de la découverte de médicaments au cours du XIX^{ème} et XX^{ème} siècle

Chapitre II. Criblage virtuel : voie d'accès à des substances médicamenteuses

Avec la chimie et la pharmacologie, deux autres disciplines ont totalement révolutionné la recherche de nouveaux médicaments et abouti aux processus actuellement utilisés dans les phases précoces de recherche et développement (R&D). Il s'agit de la génétique qui a permis la rationalisation du choix et de l'utilisation de cibles biologiques définies au niveau moléculaire et de la bioinformatique avec la mise en pratique de nouveaux outils de découverte.

La découverte d'une molécule médicamenteuse « drug discovery » est un processus souvent extrêmement long, fastidieux et incertain (Figure 2). Pour chaque nouveau médicament, le temps s'écoulant entre la première étape, au cours de laquelle une cible biologique pertinente dans un processus pathologique donné est identifiée, jusqu'à la mise sur le marché d'un médicament est estimé à une moyenne de 12 à 14 ans avec un coût global minimum de 800 millions de dollars⁶.

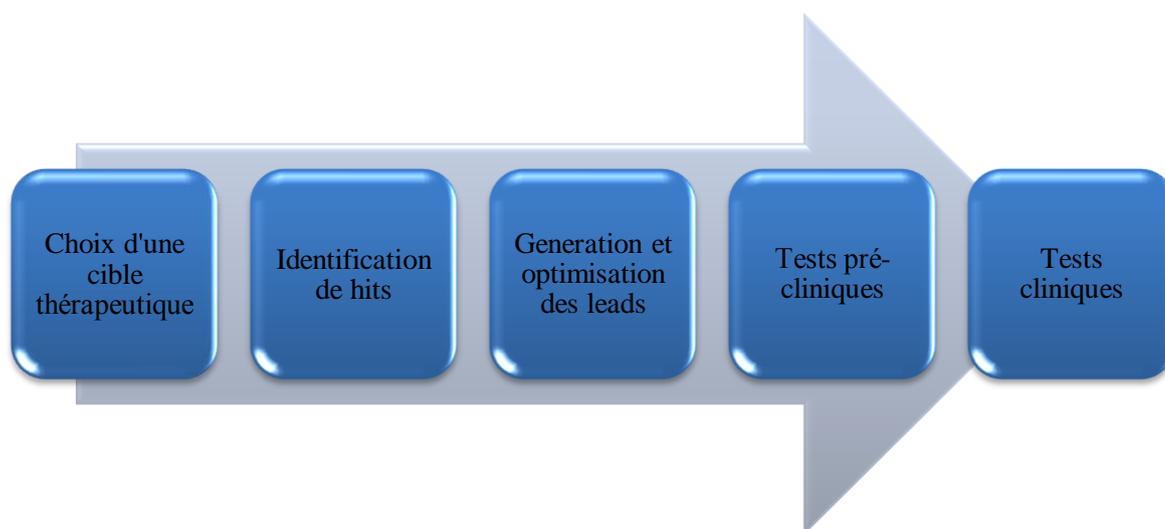


Figure II. 2. Étapes du processus de recherche et développement (R&D) des médicaments

En effet, le processus de (R&D), dans l'industrie pharmaceutique, n'est couronné de succès que pour environ une molécule sur 10000 testées. Les nouvelles méthodes permettant la découverte de nouveaux médicaments se doivent donc d'innover afin de mettre en évidence des molécules encore inconnues ayant un certain potentiel d'activité sur des cibles biologiques connues et les outils mis en place doivent être capables de guider les chimistes dans le choix des molécules à cribler et à synthétiser^{4, 7}. Ce processus suit les étapes suivantes:

Choix d'une cible thérapeutique :

Le processus de découverte d'un nouveau médicament doit logiquement débiter par la définition d'une maladie pour laquelle le défaut de traitement adapté est efficace et engendre un réel besoin médical. Il s'agit le plus souvent de pathologies touchant une grande partie de la population, largement étudiées et avec un fort potentiel commercial, comme par exemple les maladies cardiovasculaires, neurodégénératives, les cancers, le diabète, etc....

Une fois la maladie définie, une cible thérapeutique, à l'aide de données issues de la littérature, doit être identifiée. Cette cible est une biomolécule, en général une protéine ou un complexe protéique, impliquée dans un processus pathologique.⁸

Avant de rechercher des molécules capables d'agir sur la cible biologique identifiée, il est nécessaire de procéder à la validation de cette cible. Cette validation consiste d'une part à s'assurer de l'effet bénéfique de la modulation de la cible sur la pathologie étudiée tout en vérifiant d'autre part que les conséquences de ces altérations ne seront pas néfastes. Ces paramètres régulent ce que l'on appelle la « druggabilité » de la cible^{9, 10}.

De nombreux outils *in vitro* et *in vivo* sont utilisés lors de cette étape de validation, parmi lesquels les animaux transgéniques, les petits ARN interférents (ou small interfering RNA siRNA), les anticorps monoclonaux ou encore la chémogénomique (dont le but est de fournir une petite molécule pour chaque protéine codée par le génome pour explorer les fonctions cellulaires et guider la découverte de nouveaux médicaments).^{11, 12}

Identification de hits :

Dans cette seconde étape, les effets de cible en question sont modulés pour pouvoir agir favorablement sur le processus pathologique en question, et ceci par interaction avec des molécules. Ces composés ayant la capacité d'interagir avec la cible et susceptibles de moduler ses effets sont appelés des « touches » ou « hits ». Les hits sont sélectionnés par le criblage « screening » de bibliothèques de composés « chimiothèques » ciblées ou d'autres non ciblées. Le screening s'effectue, soit par le biais des tests expérimentaux, qui nécessitent le plus souvent une connaissance approfondie des systèmes étudiés. Ce type de screening, peut s'avérer très compliqué à mettre en œuvre et parfois irréalisable au vu des coûts financiers et temporels très importants, surtout lors d'un criblage à haut débit HTS «High Throughput Screening»¹³. Pour pallier ces difficultés, des techniques de criblage virtuel sur ordinateur «*in silico*» ont été développées. Ce criblage, contrairement aux tests expérimentaux, est

souvent plus aisé à mettre en œuvre, peu coûteux et relativement rapide. En revanche, ce criblage donne des résultats prédictifs qui doivent donc ensuite être impérativement validés par des tests expérimentaux, ceci représente l'inconvénient majeur de ce genre de screening.¹⁴

Par conséquent, la combinaison des deux approches de criblage est utilisée très fréquemment lors des processus actuels de R&D de nouveaux médicaments en considérant le criblage *in silico* en tant que premier filtre des chimiothèques afin d'identifier les molécules les plus prometteuses et restreindre le nombre de composés à cribler expérimentalement¹⁵.

1.3 Génération et optimisation des leads :

A ce stade, parmi les hits précédemment choisis, quelques molécules se qualifient têtes de séries ou « leads ». Les leads idéaux sont des composés plus actifs et plus sélectifs que les hits dont ils sont issus, tout en présentant des propriétés pharmacocinétiques optimales, qui vont ensuite permettre de créer des séries de composés ayant une activité démontrée sur la cible choisie¹².

L'intervention des chimistes et des chémoinformaticiens est déterminante à cette étape pour obtenir ces leads, en effectuant de vastes études de relations structure-activité (SAR) basées principalement sur la variation et la modification des groupements fonctionnels tout en gardant leurs squelettes de bases des hits. L'analyse des résultats issus de ces études SAR sont actuellement assistés par des approches bioinformatiques¹⁶, ce qui permet de guider la sélection de nouveaux leads en optimisant les propriétés favorables d'activité « drug-like » (activité, sélectivité, toxicité, propriétés physico-chimiques,...)¹⁷ tout en tentant d'améliorer l'affinité avec la cible et de prouver que l'effet biologique observé est bien induit par interaction du composé avec la cible¹⁸.

1.4 Le tests précliniques et cliniques :

Le rôle d'un chimiste au cours du processus R&D se limite dans les étapes précédemment décrites à identifier des médicaments candidats susceptibles de devenir un jour un médicament commercialisé¹⁰. Une fois les candidats identifiés, place aux tests précliniques et cliniques qui représentent les étapes les plus longues et les plus coûteuses, afin d'évaluer l'efficacité, la sélectivité et la sûreté de ces candidats.

Les tests précliniques s'effectuent sur les animaux, *in vivo*, pour assurer une analyse du métabolisme, de l'absorption et de la distribution du médicament et permettre également d'obtenir plus d'informations toxicologiques dans le but de prédire ces effets secondaires et

Chapitre II. Criblage virtuel : voie d'accès à des substances médicamenteuses

déterminer les dosages adéquats pour les essais cliniques. De nos jours, pour épargner des vies animales, ces tests sont réalisés *in vitro* en remplaçant les modèles animaux par des cellules humaines dérivant de différents tissus et organes¹². Cette étape est très importante puisque la poursuite ou non du développement des composés et leur entrée en phase clinique sont décidées à cette étape. Cette décision représente pour l'équipe de R&D une double prise de risque aux vues du grand coût financier de telles études mais aussi et surtout de la grande responsabilité incombée puisque les candidats médicaments vont être testés à l'échelle humaine.

Les tests à échelle humaine appelés «tests cliniques», constituent l'étape la plus critique et la plus longue de tout processus de R&D. Cette étape se passe en plusieurs phases en évaluant l'efficacité du candidat médicament chez l'homme, ainsi que sa pharmacocinétique et sa sécurité d'emploi pour une éventuelle mise du médicament sur le marché.

Dans une première phase, les tests sont effectués sur un faible nombre de volontaires sains pour s'assurer de la sécurité du médicament candidat en recherchant les effets secondaires qui apparaissent lors de l'administration croissante du candidat médicament et pour déterminer la dose minimale active et la dose maximale tolérable. Cette phase n'est pas autorisée quand il s'agit de molécules anticancéreuses et antibiotiques. Dans ces cas, les chercheurs passent directement à la phase II. Réalisée sur un groupe de patients malades pour évaluer l'efficacité (la biodisponibilité) du candidat, cette phase permet de déterminer la posologie du traitement et d'ajuster les dosages à prescrire en fonction des réponses obtenues des patients. Les premiers effets secondaires du médicament peuvent être également déterminés lors de cette phase, et si les effets secondaires sont trop importants le processus R&D est suspendu définitivement à cette phase. Par conséquent, seul un tiers des essais cliniques accèdent à la phase III. Celle-ci s'effectue sur une large population de malades et consiste à évaluer le potentiel thérapeutique réel du candidat médicament et établir le rapport bénéfices/risques. Les résultats sont ensuite comparés avec ceux d'un traitement de référence pour valider et confirmer l'efficacité du médicament⁵.

A la fin de ces trois phases, toutes les informations et les résultats obtenus sont rassemblées dans un dossier déposé pour obtenir une Autorisation de Mise sur le Marché (AMM), et seul un candidat médicament sur dix entrent en phase clinique obtient une AMM et est commercialisé. Une fois le médicament commercialisé, une dernière phase IV débute et des tests de surveillance ou de pharmacovigilance sont mis en œuvre. Le médicament

Chapitre II. Criblage virtuel : voie d'accès à des substances médicamenteuses

administré à une large population pour révéler les effets secondaires (indésirables ou non) apparaissant plus tard dans la prise du médicament et inexplorés lors des phases précédentes pour garantir aux patients une plus grande sécurité d'utilisation¹².

2. Criblage virtuel «in silico», méthodes et alternatives :

Le terme criblage virtuel regroupe un ensemble de techniques computationnelles ayant pour objectif l'exploration de bases de composés à la recherche de molécules d'intérêt¹⁹. Une analogie souvent utilisée compare ces techniques à des filtres qui permettraient de constituer des ensembles de molécules partageant certaines propriétés, de sélectionner les plus susceptibles d'interagir avec une cible donnée et d'éliminer les composés supposés inactifs ou les molécules indésirables¹⁵.

Actuellement, le criblage virtuel est utilisé en amont du processus R&D de nouveaux médicaments (Figure 3) notamment dans les premières étapes, ce qui permet un gain important en temps et en argent. D'ailleurs, un certain nombre de médicaments mis sur le marché proviennent d'une conception rationnelle basée sur des stratégies de criblage virtuel comme, par exemple, des inhibiteurs de l'aldose réductase, obtenus suite à une recherche dans des bases de données de composés, un inhibiteur d'un élément de réponse de la transactivation de la HIV-1 RNA: obtenu après une étude docking rigide d'une base de données de composés, un inhibiteur de la thrombine: obtenu après une étude docking sur des chimiothèques combinatoires^{20, 21}

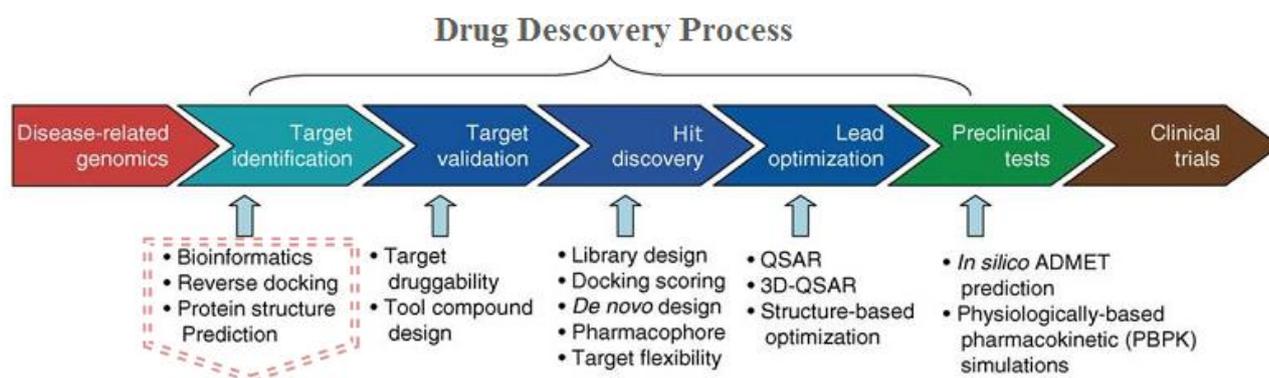


Figure II. 3. Intervention des méthodes in silico dans le processus R&D de nouveaux médicaments²².

On regroupe les méthodes in silico en deux grandes familles, le criblage virtuel « structure-based » et le criblage virtuel « ligand-based ». Comme son nom l'indique, le

Chapitre II. Criblage virtuel : voie d'accès à des substances médicamenteuses

criblage virtuel « structure-based » est basé sur la structure de la cible en évaluant la capacité des ligands à établir des interactions avec le site de liaison étudié et ainsi sélectionner les molécules capables de se lier à cette cible¹³.

En revanche, les techniques de criblage virtuel « ligand-based » utilisent les ligands connus comme guides pour définir les propriétés responsables de leur affinité en analysant les relations structure-activité de ces molécules pour découvrir de nouveaux composés susceptibles d'être actifs²³. Lorsque ces deux types de données sont disponibles simultanément, les méthodes « ligand-based » et « structure-based » peuvent toutes deux être utilisées l'une à la suite de l'autre²⁴. Ces différentes méthodes de criblage virtuel peuvent ainsi être utilisées dans les premières phases de R&D de nouveaux médicaments par les chimistes (figure 4), pour guider la sélection des composés les plus prometteurs, que ce soit lors des phases d'identification des hits ou d'optimisation des leads.

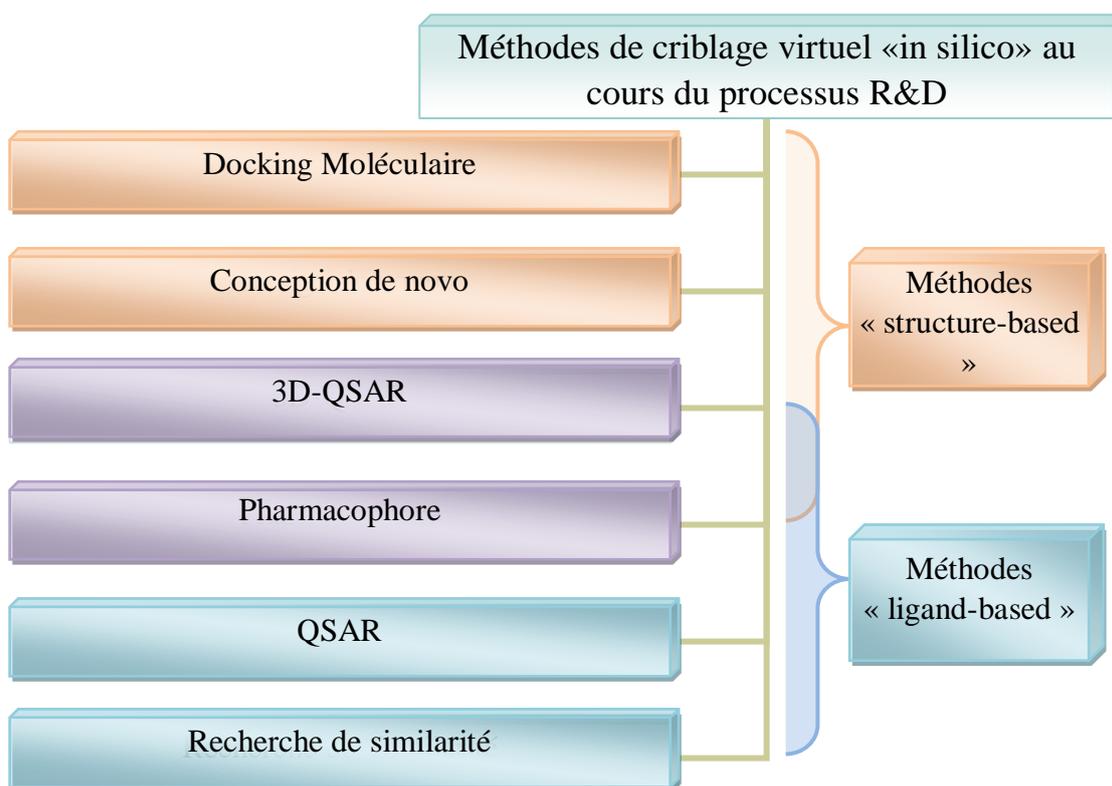


Figure II. 4. Classification des méthodes de criblage virtuel « ligand-based » et « structure-based », les plus utilisés au cours du processus R&D notamment par les chimistes.

2.1. Criblage virtuel « ligand-based » :

Un criblage virtuel « ligand-based » peut être réalisable, si au moins un ligand de la cible étudiée est déjà connu. L'hypothèse que les molécules similaires vont avoir tendance à présenter des profils d'activités très proches à similaires, est le principe de base de toutes les méthodes « ligand-based »²⁵.

Ce type de criblage très populaire est employé au cours de la phase d'identification de nouveaux hits et les phases d'optimisation des hits et des leads. Qui permet de rationaliser les processus de découverte de nouveaux composés, lorsque des données sont disponibles sur un ou plusieurs ligands actifs de référence.

2.1.1. Recherche de similarité :

La recherche de similarité est la méthode à employer lorsque très peu de ligands ont été rapportés pour la cible biologique choisie. Cette méthode repose sur l'utilisation de descripteurs et de métriques de similarité permettant de comparer des molécules à cribler à un ou plusieurs ligands de référence pour prédire leur profil d'activité²⁵.

Les descripteurs de similarité permettent de définir au sein d'une base de données quels ligands sont les plus ressemblants aux ligands actifs connus. Ce sont des nombres ou des vecteurs, souvent classés selon leur dimension (1D, 2D et 3D), qui représentent des caractéristiques structurales clés de composés. Ces descripteurs peuvent être obtenus en encodant la molécule par une représentation symbolique ce qui transforme l'information chimique en un nombre utile, qui est une procédure mathématique logique. En revanche, une procédure expérimentale peut également être utile en utilisant le résultat obtenu comme descripteur²⁶.

L'obtention des vecteurs de descripteurs pour chaque molécule de la chimiothèque ainsi que les molécules actives de référence permet ensuite de comparer ces molécules par métriques de similarité²⁷. Ces métriques sont classées en trois catégories: les mesures directes de similarité, les mesures de distances ou de dissimilitude et les mesures de corrélation. Ils sont traduits en coefficients dont la valeur est comprise dans l'intervalle [0;1] comme par exemple le cas du coefficient de Tanimoto «T_c» qui est le plus populaire et le plus couramment utilisé. Les valeurs de coefficients T_c les plus élevées sont associées aux molécules les plus similaires²⁸.

2.1.2. Modèles pharmacophoriques « ligand- based » :

Un pharmacophore est dit basé sur le ligand (ou « ligand-based ») lorsqu'il est déterminé à partir de la structure de composés actifs de référence, sans connaître ou sans prendre en compte la structure du récepteur. Lorsque la structure du récepteur est utilisée pour construire le pharmacophore, celui-ci est dit basé sur la structure ou « structure-based ».

En 1898, Ehrlich a développé le concept de pharmacophore, même si le terme pharmacophore n'était pas employé, en admettant que certains groupements chimiques dans une molécule sont responsables de l'action biologique ou pharmacologique. En 1960, la première définition moderne du pharmacophore a été établie en remplaçant « groupements chimiques » par le terme « caractéristiques abstraites »²⁹. Quelques années plus tard, Kier a publié le premier modèle de pharmacophore pour des agents muscariniques avec de mesures précises des distances entre groupements constituant le pharmacophore³⁰.

La définition officielle de l'IUPAC de 1998 indique que le pharmacophore se constitue de l'ensemble des propriétés stériques et électroniques d'une molécule, nécessaires pour assurer l'établissement d'interactions supramoléculaires optimales avec une cible biologique spécifique et engendrer ou bloquer une réponse biologique³¹. Ceci indique que les molécules partageant le même pharmacophore pour une cible donnée devraient donc se lier de manière identique à ce récepteur et présenter des profils d'activité similaires. Le pharmacophore généré est donc utilisé pour cribler la chimiothèque à la recherche de molécules se superposant à ce pharmacophore. L'une des caractéristiques majeures de ce type de méthodes est qu'un pharmacophore est défini par des points pharmacophoriques complémentaires les uns des autres, qui sont des groupes fonctionnels et non plus des groupes d'atomes. Les différents points pharmacophoriques recherchés sont les donneurs et les accepteurs de liaisons hydrogènes, les groupements chargés positivement ou négativement qui forment des interactions électrostatiques avec ceux de charge opposée, et les groupements aromatiques en tant que groupement hydrophobes³².

On distingue les approches pharmacophoriques 2D et 3D selon le format dans lequel sont présentés les ligands utilisés à la recherche du pharmacophore.

2.1.3. QSAR : Modèles de relations quantitatives structure-activité :

Les premiers essais de modélisation des relations structure-activité ont commencé à la fin du 19^{ème} siècle, lorsque Crum-Brown et Frazer ont postulé que l'activité biologique

d'une molécule est une fonction de sa constitution chimique³³. Mais ce n'est qu'au début des années 60 que les travaux de Corwin Hansch ont proposé un modèle mathématique reliant l'activité biologique à la structure chimique³⁴. Aujourd'hui, l'utilisation de QSAR n'a cessé de progresser³⁵. 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 chères³⁶. Sans l'utilisation de grands instruments analytiques, les résultats des études QSAR peuvent fournir des informations utiles pour obtenir une meilleure connaissance des structures moléculaires et probablement le mode d'action au niveau moléculaire. Ces informations peuvent être alors utilisées dans la prédiction des activités biologiques de nouveaux composés ainsi que dans la conception de nouvelles structures³⁷.

Le principe des méthodes QSAR est, comme leur nom l'indique, de mettre en place une relation mathématique reliant de manière quantitative la structure moléculaire, codée par des propriétés moléculaires appelée descripteurs, avec une activité en utilisant des méthodes d'analyse de données³⁸. Par ces relations nous pouvons développer des modèles prédictifs de la forme générale suivante :

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

L'objectif de ces méthodes est alors d'analyser les données structurales afin de déterminer les facteurs influençant l'activité mesurée³⁹.

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 QSAR⁴⁰. Une fois le modèle construit, il doit être évalué par des méthodes de validation (validation interne et externe) afin d'estimer sa robustesse et son pouvoir prédictif^{41, 42}. Enfin, pour tout modèle, il est important de connaître quel type de molécules ce modèle est utilisable (connaître le domaine d'applicabilité)⁴³.

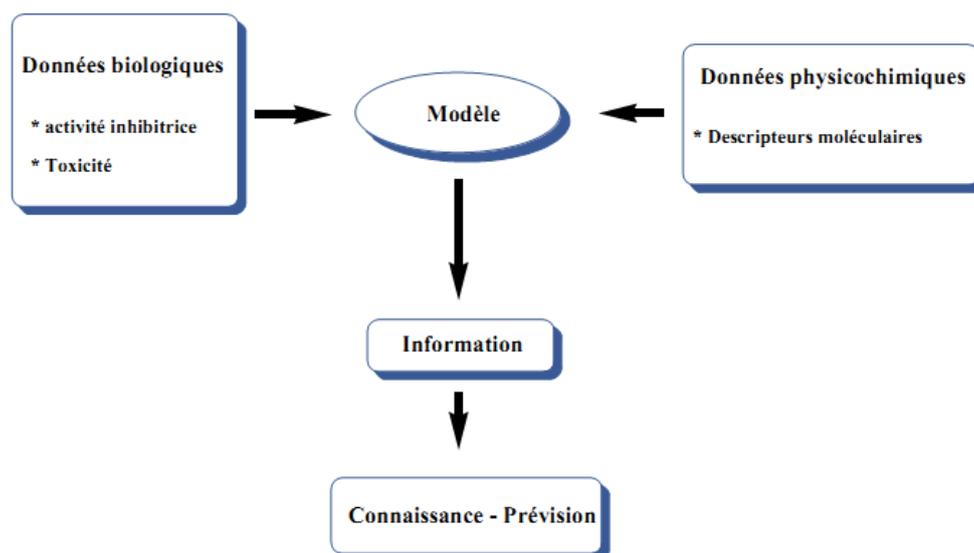


Figure II. 5. Présentation de la méthodologie de QSAR

2.1.4. 3D-QSAR :

Des modèles 3D-QSAR dit « receptor dependent » (RD-QSAR) peuvent être mis en œuvre lorsque la structure 3D de la cible biologique est résolue⁴⁴. Les approches 3D-QSAR ont été développées pour corréler l'activité biologique d'une série de composés actifs de référence avec l'arrangement spatial de nombreuses propriétés de la molécule telles que les propriétés stériques, lipophiliques et électronique et pour fournir des indications pour l'optimisation par pharmacomodulation et la conception de nouveaux composés avec des profils d'activité améliorés⁴⁵.

En 1979, La première approche 3D-QSAR a été proposée en décrivant les propriétés de champs moléculaires de composés, les calculer sur une grille régulière puis les corréler à leur activité biologique par analyse en «composante principale» (PCA)⁴⁶. Actuellement, différentes méthodes 3D-QSAR sont utilisées tel que CoMFA (Comparative Molecular Field Analysis) et CoMSIA (Comparative Molecular Similarity Indices Analysis). Il est à noter que toutes ces méthodes nécessitent un alignement minutieux des ligands de référence. Cependant, ces approches sont considérées à la fois « ligand-based » et « structure-based ».

2.2. Criblage virtuel « structure-based » :

Le criblage virtuel « structure-based » est considéré comme un équivalent *in silico* d'un test expérimental étudiant la liaison ligand-cible biomoléculaire⁴⁷. Cependant, ce criblage dépend essentiellement de la disponibilité de la structure 3D de la cible biologique

qui est obtenue, soit par méthodes expérimentales (RX et RMN) dans des bases de données (tel que PDB: Protein Database), soit par des méthodes de prédiction de la structure 3D par homologie de séquence⁴⁸.

Différentes approches peuvent être employées pour réaliser ce criblage: la construction de modèles pharmacophoriques 3D, l'établissement de modèles 3D-QSAR (section II.1.4), la conception de novo « de novo design » et les méthodes de docking moléculaire qui sont les plus populaires.

2.2.1. Les approches pharmacophoriques 3D :

Les pharmacophores 3D décrivent l'arrangement spatial des propriétés chimiques nécessaires pour l'activité biologique à partir d'un ensemble de ligands actifs de référence⁴⁹. Les pharmacophores 3D sont obtenus également à partir de la structure 3D du récepteur ou la structure 3D d'un complexe ligand-récepteur. Les pharmacophores générés sont donc basés sur la structure et peuvent être ensuite utilisés pour cribler une chimiothèque à la recherche de molécules potentiellement actives⁵⁰.

Le processus de détermination du pharmacophore se divise en plusieurs étapes successives (Figure 6) qui peuvent être assistées par plusieurs logiciel tel que CATALYST, MOE, PHASE et SCAMPI⁵¹:

- Sélectionner les ligands de référence.
- Effectuer des recherches conformationnelles.
- Déterminer et représenter les points pharmacophoriques de chaque ligand.
- Déterminer les modèles de pharmacophore.
- Attribuer des scores à chaque model obtenu afin de choisir le ou les meilleurs pharmacophores.