

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'enseignement supérieur et de la recherche scientifique

UNIVERSITÉ MOULAY TAHAR DE SAIDA

Faculté des Sciences et de la technologie

Département de Chimie



Mémoire de fin d'étude  
*En vue de l'obtention du*  
**DIPLÔME DE MASTER EN CHIMIE**

Option : *Chimie Informatique*

**Thème**

---

**L'outil informatique dans la conception des médicaments**

---

Présentée par :

*M<sup>elle</sup> TOUHAMI Moufida*

Devant le jury composé de :

<b>M<sup>f</sup> B.ARICHE</b>	<b>Président</b>	<b>Université de Saida.</b>
<b>M<sup>f</sup> AEK.ALLALI</b>	<b>Examineur</b>	<b>Université de Saida.</b>
<b>M<sup>f</sup> N.DOUMI</b>	<b>Examineur</b>	<b>Université de Saida.</b>
<b>M<sup>f</sup> A.RAHMOUNI</b>	<b>Encadreur</b>	<b>Université de Saida.</b>

*Année Universitaire 2012-2013*

## **Résumé :**

L'innovation thérapeutique progresse traditionnellement par la combinaison du criblage expérimental et de la modélisation moléculaire. En pratique, cette dernière approche est souvent limitée par la pénurie de données expérimentales, particulièrement les informations structurales et biologiques. Aujourd'hui, la situation a complètement changé avec le séquençage à haut débit du génome humain et les avancées réalisées dans la détermination des structures tridimensionnelles des protéines. Cette détermination permet d'avoir accès à une grande quantité de données pouvant servir à la recherche de nouveaux traitements pour un grand nombre de maladies. À cet égard, les approches informatiques permettant de développer des programmes de criblage virtuel à haut débit offrent une alternative ou un complément aux méthodes expérimentales qui font gagner du temps et de l'argent dans la découverte de nouveaux traitements.

Appliqué aux grandes bases de données moléculaires, le criblage virtuel à haut débit permet de limiter le criblage expérimental en fournissant, pour chaque cible biologique visée, des molécules potentiellement intéressantes au moyen de méthodes informatiques adaptées. Cependant, la plupart de ces approches souffrent des mêmes limitations. Le coût et la durée des temps de calcul pour évaluer la fixation d'une collection de molécules à une cible, qui est considérable dans le contexte du haut débit, ainsi que la précision des résultats obtenus sont les défis les plus évidents dans le domaine. Le besoin de gérer une grande quantité de données hétérogènes est aussi particulièrement crucial.

## **Title :**

Computer tool in drug design

## **Abstract :**

Therapeutic innovation traditionally increased by the combination of experimental screening and molecular modeling. In practice, this latter approach is often limited by the lack of experimental data, particularly the structural and biological information. Today, the situation has completely changed with the high-throughput sequencing of the human genome and advances in determining the three-dimensional structures of proteins. This determination provides access to a large amount of data that can be

used to research new treatments for many diseases. In this regard, computational approaches for developing programs virtual high throughput screening offer an alternative or complement to experimental methods that save time and money in the discovery of new treatments.

Applied to large molecular databases, virtual high throughput screening limits the experimental screening by providing, for each biological target, potentially interesting molecules through appropriate computational methods. However, most of these approaches suffer from the same limitations. The cost and duration of the computation time to evaluate the establishment of a collection of molecules to a target, which is significant in the context of broadband, as well as the accuracy of the results are the most obvious challenges in the field. The need to manage large amounts of heterogeneous data is also particularly crucial.

## ملخص:

الابتكار العلاجي زادت تقليديا من قبل مجموعة من الفحص التجريبي والنمذجة الجزيئية. في الممارسة العملية، وهذا النهج الأخير غالبا ما تكون محدودة بسبب عدم وجود بيانات تجريبية، ولا سيما المعلومات الهيكلية والبيولوجية. اليوم، فإن الوضع قد تغير تماما مع التسلسل عالية الإنتاجية للجينوم البشري والتقدم في تحديد هياكل ثلاثية الأبعاد للبروتينات. هذا التصميم يوفر الوصول إلى كمية كبيرة من البيانات التي يمكن استخدامها للبحث عن علاجات جديدة للعديد من الأمراض. في هذا الصدد، والنهج الحسابية لتطوير برامج افتراضية عالية الإنتاجية الفرز تقدم بديلا أو مكملًا للطرق التجريبية التي توفر الوقت والمال في اكتشاف علاجات جديدة.

تطبق على قواعد البيانات الجزيئية كبيرة، يحد الظاهري عالية الإنتاجية الفرز والفحص التجريبي من خلال توفير، لكل هدف بيولوجي، فإن الجزيئات يحتمل أن تكون مثيرة للاهتمام من خلال الأساليب الحسابية المناسبة. ومع ذلك، فإن معظم هذه المناهج تعاني من نفس القيود. التكلفة والمدة من الوقت اللازم للحساب لتقييم إنشاء مجموعة من الجزيئات إلى الهدف، والتي تعتبر مهمة في سياق واسع النطاق، فضلا عن دقة النتائج هي التحديات الأكثر وضوحا في هذا المجال. الحاجة لإدارة كميات كبيرة من البيانات غير المتجانسة هي أيضا حاسمة بشكل خاص.

*A Muma*

*A mes parents*

*A mes sœurs et mes frères*

*A mes amis et mes collègues*

*A tous les gens qui m'ont soutenu*

## **REMERCIEMENTS**

*Le présent travail a été réalisé au sein du Laboratoire de Modélisation et Méthodes de Calcul de l'université Dr Moulay Tahar de Saïda, sous la direction du Professeur RAHMOUNI Ali Professeur à l'université de Saïda.*

*Je tiens tout particulièrement à lui exprimer ma profonde reconnaissance d'avoir accepté de diriger ce mémoire. Je tiens aussi à lui témoigner toute ma gratitude pour l'intérêt qu'il a constamment manifesté pour ce travail.*

*J'adresse mes vifs remerciements à Mr B.ARICHE qui m'a fait l'honneur de présider le jury de ce mémoire*

*J'adresse mes plus sincères remerciements à Mr AEK.ALLALI et Mr N.DOUMI qui ont accepté d'examiner ce travail et de participer aux membres de jury.*

*Mes remerciements vont aussi aux enseignants de la spécialité « Chimie Informatique »*

*En fin, un grand merci à tous mes camarades, de l'option « chimie informatique », ils se connaissent tous, ils ont été à bien des égards des camarades d'exception. L'ambiance inoubliable, qu'ils ont su créer dans le groupe.*

# Sommaire

---

oSommaire :

Introduction générale.....1

## **Chapitre 1 : La mise au point d'un médicament**

I. La maladie.....5

I.1.Causes des maladies.....5

I.2.Identification et traitement.....6

II. Le médicament.....7

II.1.Composition d'un médicament.....9

II.2.Catégories thérapeutiques.....9

II.3.Origines des médicaments.....10

II.3.1.Origine naturelle.....10

II.3.2.Origine synthétique.....10

II.4.La dénomination des médicaments.....11

III. Identification et validation des cibles thérapeutiques.....11

III.1.Site actif.....12

IV. La genèse d'un médicament.....13

IV.1.Première étape.....13

IV.2.Deuxième étape.....14

IV.3.Troisième étape.....14

IV.4.Quatrième étape.....15

IV.5.Cinquième étape.....16

V .Processus de découverte.....16

V.1.Découverte fortuite.....16

V .2.Découverte ciblée.....19

VI. Développement prés-clinique.....20

VI.1Etude pharmacodynamique.....20

VI.2.Etude pharmacocinétique.....21

## Sommaire

---

VI.3.Etude toxicologique.....	21
VI.4.Etude technique.....	22
VI.5.Etude analytique.....	23
VII. Développement clinique.....	23
VII.1.Etude clinique phase 1.....	23
VII.2.Etude clinique phase 2.....	24
VII.3.Etude clinique phase 3.....	24
VII.4.Types d'études cliniques.....	25
VII.5.Parcours administratifs.....	25
VII.6.Pharmacovigilance.....	26
VIII. Fabrication.....	26
<b>Chapitre 2 : Identification d'une entité thérapeutique assistée par ordinateur</b>	
I. Introduction.....	29
II. Notion de chimiothèque.....	30
II.1.La chimie combinatoire.....	30
II.1.1.La synthèse de bibliothèque combinatoire.....	31
II.1.2.Conception de bibliothèques combinatoires et génération de diversité.....	32
II.1.3.Les caractéristiques les plus importantes de la chimie combinatoire.....	34
II.2. Les Chemins aléatoires (Random walks) .....	34
II.3. Constructions à partir des structures existantes.....	35
II.4. L'espace chimique des produits naturels.....	35
II.5. Assemblage des fragments .....	36
III. Les collections de ligands et de cibles pour le criblage .....	37
III.1. Les collections de ligands .....	37

## Sommaire

---

III.2. Les ressources pour les cibles .....	38
VI. Les descripteurs moléculaires et indices de similarité .....	38
VI.1. Descripteurs moléculaires des ligands .....	38
VI.2. Descripteurs moléculaires des cibles .....	42
V. La dynamique moléculaire .....	45
V.1. Principe .....	45
V.2. Intérêt de la dynamique moléculaire en amont du CHD .....	45
VI. Méthodes CHD basées sur la structure du ligand (ligand-based) .....	46
VI.1. Méthodes basées sur la topologie du ligand (2D) .....	46
VI.1.1. Empreinte (fingerprint) topologique .....	46
VI.1.2. Arbre de propriétés .....	46
VII. Méthodes de CHD basées sur la structure de la cible (structure-based) .....	46
VII.1. Amarrages moléculaires .....	46
VIII. Profilage de l'activité et criblage parallèle .....	47

### **Chapitre 3 : Conception d'une application informatique pour la génération d'une chimiothèque**

I. Introduction .....	49
II. Présentation du projet .....	49
II.1. Cahier des charges .....	49
II.1.1. Constatation de départ .....	49
II.1.2. Définition de l'application à réaliser .....	49
II.1.3. Fonctions essentielles que devra remplir l'application .....	50
II.2. Conceptualisation de l'application .....	50
II.2.1. L'organigramme .....	50

## Sommaire

---

II.2.2.L'algorithme .....	52
Conclusion.....	54

## **Liste des abréviations :**

TCSD: Tufts Center for the Study of Drug Development

DCI : Dénomination Commune Internationale

OMS : Organisation Mondiale de la Santé

SIDA : Syndrome d'Immunodéficience Acquise

IUPAC : International Union of Pure and Applied Chemistry

AMM : Autorisations de Mise sur le Marché

HTS: High Throughput Screening

DMT : Dose Maximale Tolérée

ADN : Acide désoxyribonucléique

FDA : Food and Drug Administration

EMA: Agence Européenne du Médicament

CHD: Criblage à Haut Débit

SOSA: Selective Optimization of Side Activities

SCONPs: Structural Classification of Natural Products

DM: Dynamique Moléculaire

PDB: Protein Data Bank

ADME/Tox: Absorption, Distribution, Métabolisme, Excrétion/Toxicité

## **Listes des figures :**

Figure 1 : La genèse d'un médicament

Figure 2 : Aperçue d'une plateforme de HTS robotisé et une plaque 96 puits

Figure 3 : Méthode « Split »

Figure 4 : Quelques exemples de descripteurs et leur classification en 1D, 2D, et 3D d'une molécule.

## **Chapitre 2 :**

**Identification d'une entité thérapeutique assistée par ordinateur**

## I. Introduction :

Pour des raisons économiques (liées au coût très élevé du criblage à haut débit ou de la chimie combinatoire souvent impossible dans le milieu de la recherche et dans les petites compagnies) et, pour des raisons scientifiques (liées à l'augmentation croissante du nombre de cibles, et notamment les cibles 3D, à la taille des bibliothèques chimique, à l'augmentation de la puissance de calcul des ordinateurs, à une meilleure prédiction des interactions protéine ligand et aux progrès réalisés en informatique), une nouvelle méthodologie présentant un réel gain en temps et en argent, le Criblage à Haut Débit(CHD), a été développée[15]. Le CHD est une technique informatique utilisée en recherche dans le domaine de la conception des médicaments.



Figure 2 : Aperçu d'une plateforme de HTS robotisée et une plaque 96 puits

Il consiste en un parcours test sur de larges librairies de molécules chimiques pour en sélectionner celles qui ont le plus de chances de se transformer en médicaments. Le CHD est le fruit des avancées scientifiques dans les domaines de la modélisation moléculaire [16], la chimie combinatoire et la biologie moléculaire. De nos jours, des millions de molécules doivent être testées en une courte période de temps, d'où la nécessité d'avoir des méthodes *in Silico* pour faire un criblage rapide et efficace. L'ordinateur se substitue alors à une partie de l'expérimentation pharmacologique.

Ces technologies de modélisation peuvent aussi contribuer à développer de nouvelles approches dans la recherche de molécules efficaces sur une cible donnée. Elles permettent de restreindre l'espace chimique des molécules intéressantes pour une maladie ou une cible thérapeutique et de se focaliser sur les molécules ayant le

plus de chances d'avoir de bons résultats aux tests expérimentaux ou d'aboutir à d'éventuels médicaments.

Le CHD a pris de plus en plus d'importance et constitue un réel apport permettant d'accélérer le processus de découverte de nouvelles molécules d'intérêt thérapeutique. On part alors de la structure de la cible et l'on tente de mettre au point le modèle de molécule susceptible d'interagir avec elle. Les produits les plus proches sont ensuite synthétisés et testés. Toutes les méthodes de CHD sont généralement utilisées dans les premières phases du développement de molécules têtes de série à fort potentiel thérapeutique, afin de le rendre plus efficace et moins coûteux en temps et en argent dans le processus de découverte de nouveaux médicaments. Les méthodes in Silico sont aujourd'hui bien insérées dans le processus de découverte de nouveaux traitements dans l'industrie pharmaceutique.

## **II. Notion de chimiothèque :**

### **II.1. La chimie combinatoire :**

La découverte de nouvelles substances bioactives, et plus particulièrement de molécules présentant des propriétés spécifiques en vue d'une application thérapeutique précise, repose sur la mise en œuvre de plusieurs stratégies : la synthèse en série de molécules individuelles soumises, l'une après l'autre, à des tests, notamment de leur activité potentielle pour une cible biologique donnée. Ces stratégies se sont développées sous l'approbation du précepte que notre aptitude à prédire des propriétés de molécules sur la base de leur structure est toujours à son début. [17]

Le besoin de découvrir rapidement et de manière productive des molécules dotées de nouvelles propriétés utiles a émergé ainsi dans le développement de la chimie combinatoire, révolutionnant les méthodes de travail en chimie médicinale. Plus de 10000 publications et 500 revues apparues durant les 15 dernières années témoignent du succès de ce domaine hautement pluridisciplinaire. Aujourd'hui, aucune entreprise pharmaceutique ne renoncerait à se servir pleinement de cet outil puissant et performant, indispensable à sa productivité et à sa compétitivité

La chimie combinatoire implique la synthèse rapide d'un grand nombre de composants. Cette collection de molécules, appelée bibliothèque combinatoire, peut être un mélange chimique de composants ou un ensemble de composants individuels purs. Cette collection est ensuite soumise au criblage (« screening ») pour une activité biologique, permettant d'identifier d'éventuels composants actifs. D'une manière simplifiée, l'approche combinatoire est donc caractérisée par deux étapes principales : la synthèse de la bibliothèque et l'identification du composant actif.

La chimie combinatoire a ses racines conceptuelles dans le système immunitaire. Dans le corps, lorsqu'un antigène entre en contact avec la large collection d'anticorps existante, l'anticorps capable de lier le mieux l'antigène est identifié et ensuite reproduit en grande quantité, déclenchant ainsi la réaction immunitaire.

### **II.1.1. La synthèse de bibliothèques combinatoires :**

La chimie combinatoire est la science de la synthèse divergente efficace. Plutôt que de n'utiliser qu'un seul réactif dans une réaction chimique, l'idée de base est de conduire des réactions avec plusieurs réactifs de départ à la fois, soit en parallèle (petites bibliothèques), soit avec des mélange (grandes bibliothèques ; > 10000 composants). Ainsi, toutes les combinaisons possibles sont formées systématiquement à chaque étape, générant une bibliothèque avec une grande population de composants à partir d'un petit nombre de réactifs de départ. [18]

La taille de la bibliothèque en nombre de composants croît de manière exponentielle avec le nombre de réactifs de départ et la quantité d'étapes mise en jeu. La diversité ainsi créée aux niveaux moléculaire et structural n'est, en principe, limitée que par la disponibilité et la diversité de ces réactifs de départ.

La synthèse combinatoire, également appelée synthèse à haute capacité (« high throughput synthesis »), a stimulé le développement de nouveaux outils puissants, notamment la synthèse hétérogène sur support solide. Elle requiert de nouvelles méthodes de synthèse telles que par exemple la méthode « divide-couple-and-recombine » (« split method ») (Figure 3), permettant de contourner les problèmes de purification et d'isolation et de générer des bibliothèques avec une distribution statistique idéale (équimolaire) des composants. La chimie combinatoire met en jeu de faibles quantités de réactifs et requiert des méthodologies de synthèse efficaces

et sélectives. Elle permet également une automatisation, en particulier l'utilisation des robots de synthèse, afin d'augmenter la productivité et d'optimiser la rentabilité.

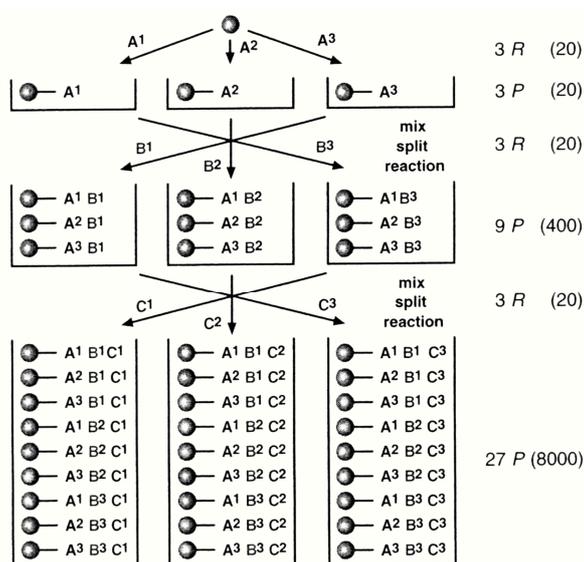


Figure 3 : Méthode « Split »

### II.1.2. Conception de bibliothèques combinatoires et génération de diversité :

Si l'on considère que le nombre total de composants organiques ayant une masse moléculaire inférieure à 750 g·mol<sup>-1</sup> est estimé à 10<sup>200</sup>, il est évident que, malgré le grand potentiel de la chimie combinatoire, il est impossible de synthétiser et de tester tous les composants potentiellement actifs. Un choix doit donc être fait par rapport aux réactifs de départ et aux produits désirés. La probabilité de trouver un composé actif augmente avec la diversité et la dissemblance des constituants de la bibliothèque. Il est important de souligner que la diversité d'une bibliothèque en nombre et en structure de composants est à distinguer de la variété qu'une telle bibliothèque présente pour une cible donnée. La variété dépend de manière critique du site d'interaction moléculaire. Une bibliothèque combinatoire peut donc avoir une grande variété pour une cible donnée, mais aucune variété pour une autre.

Si des informations structurales sur le site actif de la cible sont connues, l'évaluation des paramètres interactionnels à l'aide de méthodes assistées par ordinateur telle que la modélisation moléculaire, permet de concevoir intelligemment une bibliothèque ciblée (« rational design », « structure based combinatorial synthesis »). Ceci permet de focaliser la diversité des bibliothèques autour d'un point

d'intérêt spécifique. Malheureusement, de telles informations structurales ne sont pas disponibles pour la plupart des cibles biologiques. C'est alors que les méthodes d'essai et d'erreur (« random design ») trouvent leur intérêt, essayant de couvrir le plus grand espace interactionnel possible en créant des bibliothèques avec des réactifs de départ les plus variés possibles. Le succès d'une telle démarche repose en grande partie sur un bon équilibre entre le nombre et la diversité des composés testés, les critères de diversité considérés devant rendre compte, dans la mesure du possible, des différentes propriétés susceptibles de moduler l'activité biologique

La composition des bibliothèques varie selon le but visé : la recherche d'une piste, c'est-à-dire d'un produit ayant une activité intéressante sur la cible donnée, ou l'optimisation d'une piste (Tableau 1).

**Tableau 1. Caractéristiques de bibliothèques combinatoires.**

<b>Recherche de piste</b>	<b>Optimisation de piste</b>
Bibliothèques aléatoires	Bibliothèques ciblées
Plusieurs cibles	Une cible ou une famille de cibles
Synthèse sur support solide (synthèse de bibliothèques de petites molécules)	Synthèse en solution/support solide (la structure de la piste détermine la chimie)
Synthèse « split »	Synthèse parallèle automatisée
Mélanges de composants (« one bead, one compound »)	Composants individuelles
>> 10000 composants	<< 10000 composants
< 1 mg par produit	> 1 mg par produit
Criblage de mélanges (avec déconvolution)	Criblage de composants individ. (avec décodage)

Le principe mathématique sous-jacent au concept de la chimie combinatoire est celui de l'avantage de parallélisme (« parallelism advantage »). Imaginons une bibliothèque de tripeptides composés à partir de 20 acides aminés. Cette bibliothèque contient alors un ensemble de  $20 \times 20 \times 20 = 8000$  composants. Pour déterminer le composant actif dans cette bibliothèque, il suffit d'un maximum de  $20 + 20 + 20 = 60$  expériences ( $\equiv$  séquences de synthèse de sousbibliothèques et leur criblage), au lieu de 8000 dans le cas de produits isolés.

En plus de la réduction du nombre de tests, la rapidité et la sensibilité des tests biologiques ont été considérablement augmentées grâce au progrès en biologie moléculaire, en microbiologie et en technologie de gènes. Cela a permis la mise en place des systèmes automatisés de criblage à haut débit (« high throughput screening ») capables de tester les bibliothèques fournies par la synthèse combinatoire à haute capacité.

### **II.1.3. Les caractéristiques les plus importantes de la chimie combinatoire :**

Une bibliothèque combinatoire 'classique' consiste en un ensemble réel de différents composants organiques discrets, covalents, chimiquement stables, préparés de manière irréversible à partir d'une collection d'unités de base (réactifs de départ) en appliquant des séquences répétitives de synthèse propres et quantitatives (cas idéal), afin de créer la plus grande diversité structurale possible. Idéalement, cette bibliothèque statique est constituée de composants en fractions équivalents.

Si le composant actif ne se trouve pas ou en quantité très faible dans le mélange, le criblage de la bibliothèque restera négatif, bien que tous les éléments pour construire le substrat idéal puissent être contenues dans l'ensemble des constituants de la bibliothèque. De plus, il y a stricte séparation entre la synthèse des bibliothèques, qui se fait en absence de la cible, et leur criblage. Le criblage, à son tour, implique des procédures de déconvolution et/ou d'encodage pour la détermination du composant actif dans la bibliothèque

### **II.2. Les Chemins aléatoires (Random walks) :**

Cette méthode permet d'explorer la diversité moléculaire en utilisant la génération aléatoire des structures moléculaires à partir d'un sous-ensemble constitué de molécules ayant un intérêt thérapeutique précis. Étant donné les dimensions de

l'espace chimique des petites molécules, la probabilité de recherche fructueuse par les chemins purement aléatoires est faible, ce qui explique l'échec des premières mises en œuvre de la chimie combinatoire qui n'incluaient pas les considérations sur « la qualité » des médicaments.

De plus en plus, l'attention est portée sur les bibliothèques focalisées, basées soit sur des échafaudages moléculaires productifs (médicaments existant et structures privilégiées), soit sur l'imitation de la chimie combinatoire de la nature par des méthodes basées sur la diversité [19]

### **II.3. Constructions à partir des structures existantes :**

La base la plus fructueuse pour la découverte d'un nouveau médicament est de partir d'un médicament déjà connu. Clairement, il existe beaucoup d'exemples de médicaments qui sont des variations d'anciens médicaments : par exemple les  $\beta$ -bloquants, les antidépresseurs tricycliques, le 1,4-dihydropyridine bloquant des canaux de calcium, etc.. Parfois des médicaments sont produits en réponse à la concurrence ou pour des besoins économiques.

Cette approche produit fréquemment des molécules très similaires mais avec des propriétés pharmacocinétiques et pharmacodynamiques différentes. De plus, puisque ces molécules sont déjà connues dans l'occupation d'un espace pharmacologique validé, elles peuvent être employées comme « têtes de série » pour de nouvelles indications thérapeutiques. [20]

L'approche basée sur l'optimisation sélective des effets secondaires (SOSA) permet de cribler une bibliothèque de médicaments existants (avec des valeurs de toxicité et biodisponibilité connue chez l'homme) sur de nouvelles cibles. L'optimisation structurale subséquente permet de convertir l'effet secondaire pour un médicament existant en une activité principale d'une nouvelle entité thérapeutique.

### **II.4. L'espace chimique des produits naturels :**

Historiquement, les produits naturels et leurs dérivés ont été une source majeure d'agents thérapeutiques. Presque 50 % des nouvelles entités chimiques présentées dans les années 1980 et les années 1990 ont été tirées directement ou indirectement des structures de produits naturels.

Il y a maintenant un regain d'intérêt de la biosynthèse des squelettes de produit naturels et de l'amélioration des méthodes synthétiques qui peuvent plus aisément produire des structures complexes issues des produits naturels. Traditionnellement, les structures complexes des substances naturelles sont traitées sur la base "une molécule à la fois".

Des méthodes synthétiques plus récentes s'efforcent, grâce à la synthèse basée sur la diversité, de permettre, à partir d'un bloc chimique simple au départ, la génération d'une grande collection d'entités diverses et complexes. Pour faciliter l'exploration d'un espace chimique de composés basés sur les produits naturels, Waldmann et ses collaborateurs ont fourni une classification structurale des produits naturels (SCONPs) basée sur les squelettes sous-jacents présents dans ces produits naturels. [22]

Cette organisation hiérarchique basée sur le squelette structural des composés naturels peut fournir des conseils sur la sélection de motifs moléculaires spécifiques utilisés dans le développement de bibliothèques chimiques basées sur les composés naturels.

### **II.5. Assemblage des fragments :**

La découverte de nouveaux médicaments basée sur les fragments repose sur le concept de base que la complémentarité moléculaire est plus facilement et efficacement explorée avec les petits fragments moléculaires de taille allant jusqu'à 12 atomes, puisque le nombre de tels fragments, ~107, est beaucoup plus petit que le grand nombre de molécules semblables au médicament qui est supérieur à 1060. Des exemples récents d'application de cette technique ont été détaillés par Congreve et al. La synthèse guidée par un modèle peut être considérée comme une extension de l'assemblage par fragments : par exemple, la biosynthèse peut être utilisée comme modèle d'assemblage de fragments moléculaires afin de conduire à la formation des liaisons covalentes entre fragments et former une molécule active. Un excellent exemple de succès de cette approche est la synthèse de l'inhibiteur de l'acétylcholine estérase avec une affinité femto molaire à partir des motifs de la tacrine et du phenanthridinium. [23]

### III. Les collections de ligands et de cibles pour le criblage :

#### III.1. Les collections de ligands :

Aujourd'hui, il existe une pléthore de bases de données et de bibliothèques de molécules chimiques provenant, soit de la chimie combinatoire, ou des substances naturelles, soit des molécules, déjà connues comme intéressantes sur certaines cibles biologiques, et qui peuvent être potentiellement des médicaments « drug-like ». Tout cela contribue à élargir l'espace chimique virtuel. Une exploration systématique d'une petite partie de cet espace avec de molécules ayant jusqu'à onze atomes lourds a été récemment réalisée. Après l'exclusion des molécules peu convenables, plus de 13 millions de composés différents sont sélectionnés.

Une molécule de médicament typique peut être jusqu'à deux fois aussi grande que les composés examinés dans cette étude (masse moyenne de 340 Daltons, environ 24 atomes lourds). Le nombre de molécules « drug-like » accessibles aux procédures de synthèse actuelles est de l'ordre de 10<sup>6</sup> à 10<sup>10</sup>. Ces nombres indiquent que nous couvrons une fraction presque négligeable de l'espace chimique virtuel. Les composés pour le criblage peuvent être obtenus à partir des bases de données de structures connues, de bibliothèques combinatoires ou des programmes de conception de novo. [24]

En raison des problèmes de synthèse, on considère souvent seulement les structures connues. Des bases de données typiques avec des composés organiques de laboratoire par exemple MDL ou SPRESI ne sont pas des sources appropriées de composés de criblage en raison des propriétés non « drug-like » de la plupart des entrées. (En fait, ces bases de données sont utilisées comme des références pour des non médicaments).

De meilleures sources sont des collections disponibles en interne dans les laboratoires pharmaceutiques ou offertes par les vendeurs de composés chimiques, contenant des composés historiques et des bibliothèques combinatoires. Dans la base de données MDL de composés pour le criblage, plus de 3 millions de composés pour le criblage sont disponibles avec des informations de fournisseur. Malheureusement, toutes ces bases de données ont besoin d'un vaste nettoyage pour être appropriées pour le criblage de nouveaux médicaments. Très récemment,

la ZINC, une grande bibliothèque nettoyée commerciale de composés pour le criblage est devenue disponible. Les données de référence de composés pharmaceutiques aux différentes étapes de développement peuvent être prises du MDL Drug Data Report+, du World Drug Index++, ou de la base de données MDL Comprehensive Medicinal Chemistry+++. Ces bases de données, qui sont de grande taille, sont aujourd'hui d'un réel apport en diversité de molécules pour les méthodes in silico de découverte de nouvelles molécules thérapeutiques mais peuvent constituer une réelle limitation à cause de la quantité des données à traiter.

### **III.2. Les ressources pour les cibles :**

UniProt est le catalogue le plus complet de l'information sur les protéines. Il s'agit d'une collection de séquences de protéines et de fonctions créée en joignant les informations contenues dans Swiss-Prot, TrEMBL et PIR. UniProt a trois composantes, chacune optimisée pour différents usages. L'UniProt Knowledge base (UniProtKB) est le point central d'accès aux protéines ; il contient des informations très diverses, y compris la fonction, la classification et les références croisées. La base de données PDB est le dépôt mondial d'informations sur les structures tridimensionnelles de grandes molécules biologiques en complexe ou non, y compris des protéines et des acides nucléiques. Au mardi 23 mars 2010 la PDB contient 64 229 structures. [25]

## **VI. Les descripteurs moléculaires et indices de similarité :**

### **VI.1. Descripteurs moléculaires des ligands :**

La structure moléculaire d'un composé organique ou inorganique détermine ses propriétés. Les descripteurs moléculaires représentent les informations topologique, structurale, géométrique et physico chimique permettant de caractériser les molécules et macromolécules. Ils peuvent être calculés à partir de la structure (constitution, configuration et conformation moléculaires) ou des propriétés (physiques, chimiques, biologiques) appartenant aux molécules. Ils sont généralement utilisés pour la recherche de similarité, l'analyse de la diversité/similarité des bibliothèques de composés et de cibles. [26]

Les descripteurs moléculaires sont aussi fréquemment utilisés pour développer des modèles statistiques, pour la prédiction informatique de l'activité, la fixation du ligand au récepteur ou les propriétés toxicologiques de composés à partir de leurs structures.

#### **VI.1.1. Les descripteurs constitutionnels (0D) :**

Les descripteurs constitutionnels incluent l'information d'ordre des atomes et des liaisons. Ils sont importants pour décrire les propriétés et définir les règles de prédiction des molécules drug-like, telles que la règle des 5 de Lipinski, celle publiée par Veber et al et les modèles statistiques des propriétés pharmacodynamiques, pharmacocinétiques et toxicologiques des molécules.

#### **VI.1.2. Les descripteurs physico chimiques :**

Les descripteurs physico chimiques sont généralement utilisés en combinaison avec d'autres descripteurs pour représenter les éléments contribuant à des propriétés pharmacodynamiques, pharmacocinétiques, toxicologiques spécifiques et les composés qui se fixent à des sites catalytiques de certains récepteurs. Ces descripteurs sont aussi utilisés pour évaluer les modes de liaisons chimiques, pour estimer l'énergie libre et la constance de proportion des réactions chimiques ainsi que la présence ou l'absence de fragments.

#### **VI.1.3. Les descripteurs topologiques (2D) :**

Les indices topologiques sont utilisés pour représenter le squelette structural et les caractéristiques responsables d'une activité particulière, de la fixation au récepteur et du mode de liaison chimique de la molécule.

#### **VI.1.4. Les descripteurs géométriques (3D/4D)**

Les descripteurs géométriques concernent l'arrangement en 3D des atomes. Les descripteurs conformationnels (4D) représentent l'arrangement spatial thermodynamique stable des atomes dans une molécule. Les descripteurs géométriques sont utilisés pour prédire l'affinité de liaison et assigner la similarité moléculaire ainsi que pour prédire les propriétés pharmacodynamiques, pharmacocinétiques et toxicologiques particulières.

### VI.1.5. Les empreintes moléculaires (1D/2D/3D) :

L'empreinte chimique d'une molécule (fingerprint) est une chaîne de bits (séquence de 0 et de 1 chiffre), dans laquelle sont codées les informations sur la structure de la molécule. Elle permet de représenter de façon compacte la structure chimique d'une molécule.

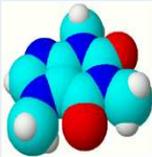
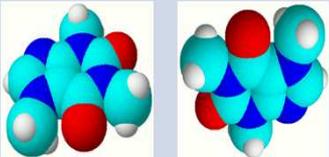
Représentation moléculaire	Descripteurs	Exemples
0D $C_8H_{10}N_4O_2$	Nombre d'atomes, Nombre liaison, poids moléculaire	Poids moléculaire, nombre atomes hydrogène, nombres atomes carbone, nombre hétéroatomes
1D $C_8H_{10}N_4O_2$	Nombre de fragments	Nombre de carbone primaire $sp^3$ , nombre de groupe ammonium, surface polaire basé sur les fragments
2D 	Descripteurs topologiques	Index zagreb, index wiener, vecteur autocorrelation 2D
3D 	Descripteurs géométriques	3D balaban index, fonction de distribution radiale, surface polaire, volume de la poche, analyse comparative du champs moléculaire (CoMFA), Coordonnées 3D
4D 	Descripteurs de conformations	Coordonnées 3D + échantillonnage des conformations

Figure 4 : Quelques exemples de descripteurs et leur classification en 1D, 2D, et 3D d'une molécule.

### VI.1.6. Sélection des descripteurs moléculaires :

Le calcul et la sélection des descripteurs sont des facteurs déterminants de la réussite du criblage virtuel de molécules. Beaucoup de questions doivent donc être posées. Si des propriétés physicochimiques sont utilisées, il faut fixer à l'avance lesquelles seront retenues et comment elles devront être calculées. Dans le cas de

descripteurs structuraux, il faut choisir le niveau de représentation (1D, 2D, 3D ou 4D) en sachant que l'approche 1D présente de nombreux avantages mais est d'un niveau descriptif incomplet. [27]

Les descripteurs 2D reflètent bien les propriétés physiques et la réactivité dans la plupart des cas, mais l'activité biologique est étroitement liée à la représentation 3D. Cependant, l'utilisation de structures 3D dans la caractérisation des molécules présente des problèmes de conformation, d'énergie et aussi de disponibilité des bases de données 3D. D'autre part, les tautomères et les ions présentent de nouvelles contraintes. Des approches dites « mixtes » sont très utilisées actuellement mais il faut choisir un groupe de descripteurs en veillant à leur indépendance et à leur utilité.

Dans ce choix, le problème à traiter est souvent non complet, c'est-à-dire un problème pour lequel le temps de résolution peut s'avérer exponentiel. Ainsi, l'usage de techniques d'apprentissage automatique (arbres de décisions, règles d'associations, etc..) semble nécessaire. En raison de l'existence de bases de molécules de plus en plus grandes, le facteur de vitesse de traitement ne pourra pas être négligé au moment de choisir la représentation optimale. Il est important de noter qu'il n'existe pas de « bon » ou de « mauvais » descripteur : l'utilité et l'efficacité sont étroitement liées aux types de molécules à traiter ainsi qu'aux calculs à effectuer. Par conséquent, la plupart des descripteurs connus aujourd'hui sont employés de préférence dans le contexte pour lesquels ils ont été créés.

#### **VI.1.7. Les indice de similarité des ligands :**

Pour mesurer la similarité moléculaire, on utilise des fonctions qui transforment les différences entre deux molécules en nombres réels, généralement dans l'intervalle unité [0-1]. Cette quantité fournit une mesure quantitative du niveau de ressemblance chimique pour un jeu de descripteurs donnés. Les mesures de similarité sont généralement constituées de deux éléments : une représentation mathématique de l'information chimique pertinente (en forme de groupes, graphes, vecteurs ou fonctions) et un index compatible avec la représentation.

Par exemple, si nous représentons une molécule  $M_i$  sous la forme d'un vecteur où chaque composante  $i$  correspond à un descripteur moléculaire individuel  $d_i$ . D'un

point de vue formel, ce vecteur positionne la molécule  $M$  dans un point de l'espace vectoriel  $V$ , dans lequel chacun des axes correspond à un descripteur. Cet espace vectoriel s'appelle « l'espace structural ». La similarité moléculaire entre deux molécules ( $M_1$ ,  $M_2$ ) sera intuitivement reliée à la distance entre les deux points dans cet espace particulier.

La règle de calcul de cette distance est appelée «métrique». Ainsi, toute mesure adéquate de la similarité doit être cohérente avec les propriétés d'une distance mathématique. L'évaluation de similarité peut être abordée par des corrélations, des mesures de distance ou des approches probabilistes ou associatives. La performance de différentes mesures de similarité est le sujet de nombreux travaux. Remarquons que l'évaluation de similarité se fait dans l'espace structural défini par les descripteurs choisie au moyen d'une métrique fixée et non par rapport aux distances interatomiques dans l'espace 3D.

## **VI.2. Descripteurs moléculaires des cibles :**

### **VI.2.1. La structure primaire :**

La structure primaire, ou séquence, d'une protéine correspond à la succession linéaire des acides aminés (ou résidus) la constituant. Les protéines sont donc des polymères d'acides aminés, reliés entre eux par des liaisons peptidiques.

### **VI.2.2. Structure secondaire :**

La structure secondaire décrit le repliement local de la chaîne principale d'une protéine. L'existence de structures secondaires vient du fait que les repliements énergétiquement favorables de la chaîne peptidique sont limités et que seules certaines conformations sont possibles. Ainsi, une protéine peut être décrite par une séquence d'acides aminés mais aussi par un enchaînement d'éléments de structure secondaire. De plus, certaines conformations se trouvent nettement favorisées car stabilisées par des liaisons hydrogène(s) entre les groupements amide (-NH) et carbonyle (-CO) du squelette peptidique. Il existe trois catégories principales de structures secondaires selon l'échafaudage de liaisons hydrogène(s) et donc selon le repliement des liaisons peptidiques : les hélices, les feuillets et les coudes

**VI.2.3. Structure tertiaire :**

La structure tertiaire d'une protéine correspond au repliement de la chaîne polypeptidique dans l'espace. On parle plus couramment de structure tridimensionnelle ou structure 3D. La structure 3D d'une protéine est intimement liée à sa fonction: lorsque cette structure est cassée par l'emploi d'agents dénaturants, la protéine perd sa fonction: elle est dénaturée. [28]

**VI.2.4. Structure quaternaire :**

La structure quaternaire des protéines regroupe l'association d'au moins deux chaînes polypeptidiques identiques ou différentes par des liaisons non-covalentes (liaison H, liaison ionique, interactions hydrophobes), mais rarement des ponts disulfures. L'effet hydrophobe est un facteur prépondérant dans l'assemblage des éléments structuraux, y compris dans l'association des sous-unités. Chacune de ces chaînes est appelée monomère (ou sous-unité) et l'ensemble est désigné sous le nom d'oligomère ou protéine multimérique.

**VI.2.5. Interactions responsables de la stabilité conformationnelle :**

Il est généralement admis que la structure d'une protéine "native" est thermodynamiquement la structure la plus stable. À l'exception des ponts disulfures qui n'existent que dans certaines protéines, principalement les protéines exocellulaires, les interactions qui stabilisent la conformation de ces molécules sont des interactions non covalentes. Toutes les interactions de ce type, qui interviennent dans les petites molécules, existent également dans les protéines.

D'autre part, les interactions non covalentes ont lieu entre les divers groupes d'une protéine, mais aussi entre ces groupes et les molécules de solvant. Ainsi, l'énergie conformationnelle d'une molécule protéique est la somme de plusieurs contributions. Certaines de ces contributions résultent de facteurs intrinsèques à la protéine : ce sont les interactions de Van der Waals (non-bonded interactions) qui comportent un terme d'attraction et un terme de répulsion, les potentiels de torsion, les énergies de contrainte dans les angles ou les longueurs de liaison.

D'autres proviennent d'interactions intramoléculaires influencées par le solvant, comme les liaisons hydrogène(s) et les interactions électrostatiques. D'autres enfin

sont principalement déterminées par le solvant, ce sont les interactions hydrophobes. Les liaisons hydrogène(s) et les interactions hydrophobes présentent une dépendance de signe opposé par rapport à la température.

Les liaisons hydrogène(s) sont plus stables à basse température, à l'inverse des interactions hydrophobes; par suite, la température correspondant au maximum de stabilité dépend de la proportion de ces interactions et, par conséquent, varie d'une protéine à l'autre. La structure native d'une protéine résulte d'un équilibre subtil entre différentes interactions stabilisantes et l'entropie conformationnelle qui tend à déstabiliser l'ensemble

#### **VI.2.6. Effet hydrophobe :**

La séquence d'une protéine comporte une certaine proportion d'acides aminés polaires (hydrophiles) et non polaires (hydrophobes). Leurs interactions avec les molécules d'eau conditionnent la manière dont la chaîne polypeptidique se replie. Les acides aminés non polaires auront tendance à éviter l'eau. Inversement, les résidus polaires vont chercher à rester à proximité du solvant aqueux. Ainsi, dans le cas des protéines solubles, il se forme un cœur hydrophobe au centre de la structure tertiaire, tandis que les groupes polaires restent plutôt en surface.

Dans le cas des protéines transmembranaires, le problème est inverse. L'environnement membranaire est globalement hydrophobe. Ainsi, les acides aminés hydrophiles vont se retrouver au cœur de la protéine tandis que les acides aminés hydrophobes vont se retrouver en surface. Des résidus hydrophiles peuvent se retrouver à la surface des protéines membranaires, en contact avec le milieu hydrophobe. Dans ce cas, il y a de fortes chances que ces résidus soient impliqués dans des interactions avec d'autres résidus hydrophiles de la même ou d'une autre protéine. [29]

#### **VI.2.7. Les descripteurs moléculaires de la cavité :**

Un certain nombre de propriétés sont utilisées pour décrire la cavité des protéines présentant un intérêt dans le processus de fixation de la protéine au ligand. Ainsi la cavité peut être représentée par sa forme, sa surface ou son volume. Les groupements chimiques et les propriétés de la cavité nécessaires à l'activité de la

protéine peuvent être codés sous forme de bit (0,1) en une empreinte (fingerprint) permettant de caractériser le site de fixation de la protéine considérée.

## **V. La dynamique moléculaire :**

### **V.1. Principe :**

Une simulation de dynamique moléculaire consiste à simuler par le calcul informatique l'évolution d'un système de particules au cours du temps. Ces simulations servent de modèles structuraux et dynamiques pour la compréhension ou la prédiction de résultats expérimentaux. Dans la pratique, cela revient concrètement à simuler le mouvement d'un groupe d'atomes dans le temps.

Les simulations par dynamique moléculaire sont majoritairement utilisées pour étudier l'espace conformationnel des macromolécules biologiques ou des petites molécules chimiques. Elles sont valables pour mieux appréhender le comportement des protéines et leurs ligands dans une période de temps donnée. Il est également possible d'étudier l'effet de molécules de solvant explicites sur la structure protéique ou du complexe protéine-ligand. [30]

Des progrès également dans les interfaces (membrane simulée, solvant explicite) permettent de minimiser l'évolution du système dans un environnement plus complexe que le vide. De meilleurs champs de forces ont par ailleurs vu le jour, impliquant un meilleur traitement des interactions électrostatiques à longue distance.

Toutefois, certains problèmes peuvent gêner l'utilisation de ces méthodes, en particulier le piégeage du complexe dans un minimum local qui ne serait pas représentatif de l'espace conformationnel et dans lequel évolue l'ensemble.

Les programmes couramment utilisés dans la simulation des biomolécules par dynamique moléculaire sont AMBER35, CHARMM36, GROMOS37 et NAMD38

### **V.2. Intérêt de la dynamique moléculaire en amont du CHD :**

La faiblesse majeure des algorithmes de CHD est l'absence ou le peu de prise en compte de la flexibilité de la protéine et du ligand lors de l'opération d'amarrage entre le ligand et la protéine. Ceci ne permet donc pas une co-adaptation optimale du ligand et du récepteur. La dynamique moléculaire est capable de traiter la flexibilité

de la protéine et du ligand. Par conséquent, le couplage des deux puissantes techniques que sont le CHD et la DM doit théoriquement augmenter le pouvoir prédictif de notre modèle.

## **VI. Méthodes CHD basées sur la structure du ligand (ligand-based) :**

### **VI.1. Méthodes basées sur la topologie du ligand (2D) :**

Ce sont des méthodes qui sont basées sur les informations topologiques et la table de connections des molécules chimiques.

#### **VI.1.1. Empreinte (fingerprint) topologique :**

Cette méthode permet de calculer de façon rapide et efficace l'empreinte topologique de toutes les molécules présentes dans une base de données, en ignorant les coordonnées des atomes. Ainsi, en utilisant des indicateurs de similarité tels que l'index de Tanimoto, on peut cribler les empreintes topologiques pré calculées pour une large base de données et en extraire les molécules les plus semblables au composé connu comme actif sur une cible biologique

#### **VI.1.2. Arbre de propriétés :**

C'est une méthode basée sur les descripteurs moléculaires qui utilisent les arbres de propriétés pour une analyse par similarité de larges bases de données de molécules. Toutes les molécules sont décrites par un arbre permettant de caractériser la composition et la hiérarchie de leur descripteur moléculaire. L'arbre décrit les groupements chimiques les plus importants dans la molécule en tenant compte de l'ensemble moléculaire

## **VII. Méthodes de CHD basées sur la structure de la cible (structure-based) :**

### **VII.1. Amarrages moléculaires :**

L'amarrage moléculaire des protéines et des ligands est considéré aujourd'hui comme une des méthodes principales de CHD basées sur la structure. Le processus

d'amarrage moléculaire entre protéine et ligand peut être divisé en deux étapes. En premier, il faut placer correctement le ligand dans le site de fixation de la protéine. En second, il faut calculer l'affinité du ligand au site de fixation de la protéine par une fonction de score. Au fil des années, on note l'apparition d'un grand nombre de programmes d'amarrage moléculaire avec une grande diversité des fonctions de score basées sur différents algorithmes et champs d'application. Ainsi, on distingue des algorithmes utilisant une construction incrémentale (FlexX), les algorithmes basés sur la forme (DOCK), les algorithmes génétiques (GOLD),

### **VIII. Profilage de l'activité et criblage parallèle :**

La recherche des composés tête de série, en plus de l'activité sur la cible biologique, il est aussi intéressant de connaître l'effet sur d'autres cibles des molécules à envoyer en test expérimental afin de minimiser les risques dans les phases suivantes du développement du médicament. Un certain nombre d'enzymes, de récepteurs et de canaux ioniques, aussi appelés « anti cible », ont été identifiés comme les principaux responsables des effets secondaires et ADME/Tox.

Ce qui fait qu'un composé ayant peu ou pas d'effet sur ces cibles avec des conséquences cardiovasculaires, toxiques ou métaboliques a plus de chance de passer les étapes suivantes du processus de conception de nouveaux médicaments. Ainsi, le profilage de l'activité des composés intéressants, au début du processus de découverte de nouveaux médicaments, peut réduire significativement le coût des méthodes expérimentales, ainsi que le risque d'échecs.

De plus, la technique de criblage parallèle a une grande valeur ajoutée dans la mise en évidence des modes de liaison inconnus et il permet aussi de cribler les médicaments approuvés sur des centaines de cibles afin d'identifier des interactions inconnues. Ces médicaments pourront ainsi être approuvés pour de nouveaux traitements avec un coût financier et expérimental considérablement faible.

## **Liste des tableaux :**

Tableau 1 : Caractéristiques de bibliothèques combinatoires.

# Introduction générale



## Introduction générale

---

Les découvertes thérapeutiques évoquent instantanément le souvenir de Pasteur et la première vaccination contre la rage de Joseph Meister ou celles d'Alexander Fleming dont l'oubli chanceux ouvrait en 1928 l'ère des antibiotiques. Cette réalité n'est que partielle car elle fait abstraction des équipes associées, des travaux lents et structurés des scientifiques, des échanges, des essais et des échecs qui constituent la véritable chaîne de la recherche thérapeutique.

Celle-ci n'est pas moins riche en progrès que par le passé mais elle est moins fournie en anecdotes héroïques, ceci étant certainement dû aux nouvelles méthodes de recherche qui nécessitent le maillage de multiples technologies, compétences et disciplines scientifiques.

La recherche de nouveaux médicaments s'inscrit dans le cadre plus général des sciences du Vivant. On entend par là un ensemble de disciplines scientifiques issues des progrès réalisés au XXe siècle autour de la biologie et dont les applications fondent désormais une très grande part des découvertes en pharmacologie et en médecine. Ces dernières années, les avancées conceptuelles initiées par la recherche en génétique ont considérablement élargi l'étendue de ces connaissances. Nous sommes ainsi passés d'une vision centrée sur le génome à celle reposant sur l'ensemble des protéines (le protéome), pour enfin considérer l'ensemble des interactions entre biomolécules (l'interactome). Protéome et interactome sont les concepts de représentation des organismes vivants

La découverte de nouveaux médicaments, de nouvelles formulations et mode de délivrance et la mise au point des analogues par modification moléculaire constituent un défi majeur pour l'industrie pharmaceutique autant aujourd'hui que par le passé

On mesure amplement à la vue de ce parcours de la molécule, de l'hypothèse scientifique à la disponibilité pour le malade, combien le chemin de recherche thérapeutique est long, en moyenne une douzaine d'années, et mobilisateur de ressources financières : entre 800 et 1 400 millions de dollars pour un médicament, d'après le TCSDD. La recherche thérapeutique présente donc à la fois un coût très élevé et un risque financier majeur. Les grandes entreprises du médicament y consacrent environ 18% de leur chiffre d'affaires

## Introduction générale

---

Ce coût [1] avait été estimé à 318 millions de dollars dans les années 80 et à 138 millions dans les années 70. En dix ans, l'investissement nécessaire à la mise au point d'une nouvelle entité moléculaire a doublé. Ces coûts élevés en recherche et développement peuvent être attribués à une variété de facteurs, tels que l'utilisation croissante dans l'industrie pharmaceutique de technologies de découverte de nouveaux médicaments qui sont coûteux, comme le criblage à haut débit, la chimie combinatoire et pharmaco-génomique.

Le travail que nous allons présenter porte sur l'introduction de l'outil informatique dans la recherche des nouvelles entités chimiques à visée thérapeutique, **le premier chapitre** représente l'état de l'art du processus complet de la conception des médicaments débuterons par l'apparition de la maladie, ses causalités, son identification et son traitement. Nous rappellerons aussi quelques définitions sur le médicament, sa composition, ses catégories passons par ses phases de recherche, découverte et développement jusqu'à sa fabrication.

**Le deuxième chapitre**, détaille la méthodologie développée afin d'améliorer les performances en temps de calcul dans le criblage, nous parlerons de l'identification des nouvelles molécules par criblage virtuel, nous exposerons aussi la notion de chimiothèque en explorant les différents espaces chimiques issues de la chimie combinatoire, chemins aléatoires, espace chimique des produits naturels et l'assemblage des fragments. Nous parlerons aussi des collections des ligands et des cibles avec la détermination des descripteurs moléculaires et indices de similarités.

**Le troisième chapitre**, concerne le travail sur la conception d'une application informatique qui génère une chimiothèque issue de la synthèse combinatoire à partir de quelques réactifs de départ à fin de produire de nouvelles molécules.

**Chapitre 1 :**

**La mise au point d'un médicament**

L'utilisation des plantes comme remèdes à certaines maladies, jusqu'à l'apparition récente des médicaments issus des biotechnologies, l'histoire du médicament est largement associée à la recherche scientifique, qui tente dès le XVI<sup>e</sup> siècle de trouver un remède spécifique pour chaque maladie[2].

Au Moyen-Age et jusqu'au début du XIX<sup>e</sup> siècle, la guérison des maladies garde un caractère « magico religieux » : on cherche par exemple à « extirper le mal » par le biais des saignées. Cependant, dès le XVI<sup>e</sup> siècle, Paracelse a l'intuition de la nécessité d'un médicament spécifique pour chaque maladie.

## **Généralités sur la mise au point d'un médicament :**

### **I. La maladie :**

Altération de la santé, des fonctions des êtres vivants (animaux et végétaux), en particulier quand la cause est connue (par opposition à syndrome). Toute maladie se définit par une cause, [3] des symptômes, des signes cliniques et para cliniques, une évolution, un pronostic et un traitement.

La cause principale de la maladie est l'état défaillant de notre corps( au moment de l'infection) lié à notre mode de vie incorrect (mauvaise alimentation, gavage, mauvaise hygiène de vie ,suralimentation, mauvaises combinaisons alimentaires, abus de toute sorte, intoxication physique et mentale, [4] manque d'activité physique, stress, effort excessif, surcharge , etc ) et les microbes ne sont que des causes secondaires .

Notre corps renferme en temps normal des milliards de microorganismes qui vivent en parfaite harmonie et qui nous protègent contre d'autres microbes. Mais quand le terrain devient défaillant et perd son équilibre, tout devient possible. D'où la nécessité de travailler en amont avec un mode de vie correct afin que notre corps reste sain et bien équilibré. Toutes les maladies sauf les traumatismes, ont leur cause dans notre mauvaise vie. Certains actes ont des effets immédiats et d'autres ont des effets à moyen ou à long terme. Si chaque malade réfléchit bien à son mode de vie passé il se retrouvera souvent dans ce qui est dit à quelques exceptions près. Parfois on croit qu'on vit bien ou qu'on mange bien alors que rien n'est évident .Tout cela à cause de notre ignorance des règles élémentaires de diététique ,d'hygiène et

d'alimentation saine et équilibrée. Nous devons manger équilibré et sans excès. Chaque effet a sa cause et ce n'est pas par hasard que l'on est bien portant, heureux, malade ou inefficace.

### **I.1.Causes des maladies :**

On peut classer les maladies en fonction de leurs causes. Une maladie est *simple* lorsque les symptômes observés peuvent tous se rapporter à une seule affection ; elle est *compliquée* quand les symptômes caractéristiques de deux ou de plusieurs affections existent simultanément [5] .

Les maladies infectieuses sont dues à une infection, l'entrée dans l'organisme d'un agent infectieux : une bactérie (maladie bactérienne), un virus (maladie virale), un prion, un parasite (maladie parasitaire). Si le parasite est un champignon, on parle de mycose.

Certaines maladies sont dues à des dérèglements du système immunitaire : les maladies allergiques (ou allergies), les maladies auto-immunes.

Les maladies de la nutrition sont liées aux apports de l'alimentation (il existe des maladies de l'excès et des maladies de carences).

Les maladies génétiques ont pour cause des anomalies dans le matériel génétique : les maladies géniques sont dues à des mutations dans des gènes ; les maladies dites chromosomiques sont la manifestation d'anomalies dans les chromosomes de l'individu. Une maladie est dite multifactorielle quand elle résulte de la conjonction de plusieurs facteurs. Elle est idiopathique si on ne connaît pas sa cause.

Une maladie est dite contagieuse lorsqu'elle est transmissible d'un individu à un autre, de façon directe (par le biais de la toux ou d'éternuements, par contact), ou indirecte (par le biais d'objet [poignée de porte, couverts...]).

Suivant leur mode de développement dans l'espace, on divise les maladies en sporadiques (lorsqu'elles se manifestent épisodiquement et par des cas isolés), épidémiques (lorsqu'elles s'étendent à un grand nombre d'individus), endémiques (lorsqu'elles persistent dans une région, sous forme de foyers permanents) et *intercurrentes* (lorsqu'elles surviennent au cours d'autres maladies

### **I.2. Identification et traitement :**

Une maladie se reconnaît à un ou à plusieurs critères réunis qui permettent son identification formelle. Ceux-ci sont déterminés par les sociétés savantes et les grands organismes sanitaires internationaux et peuvent être modifiés en fonction des progrès des connaissances.

Un exemple récent en est l'évolution de la définition du sida : la définition du stade intermédiaire entre séroconversion pure et sida confirmé s'est trouvée modifiée quand le sida a été défini comme stade clinique de la maladie, réunissant des critères cliniques et biologiques précis.

Le diagnostic formel d'une maladie peut reposer sur l'isolement d'un agent causal (le bacille tuberculeux, par exemple), sur la constatation et la localisation d'une lésion macroscopique (ulcère duodénal) ou microscopique (tissu cancéreux) ou encore sur la détection d'une anomalie biochimique (diabète sucré).

Dans certains cas, il n'existe pas de critères formels. Le diagnostic est alors porté d'après un ensemble d'anomalies cliniques, biologiques, morphologiques mais peut rester incertain, par exemple dans le cas d'un lupus (éruption localisée aux ailes du nez et aux joues) ou d'une maladie maniacodépressive. Dans de tels cas, les anomalies sont souvent classées en critères majeurs et mineurs de diagnostic, le malade devant en présenter un nombre minimal dans chaque catégorie pour que le diagnostic puisse être retenu. Certains patients ne réunissent en effet qu'une partie des critères de la maladie ; ils sont dits porteurs possibles ou probables de la maladie.

La *thérapeutique* est la section de la médecine s'occupant de l'étude des traitements. Les traitements consistent souvent, suivant le niveau évolutif de la société humaine concernée, en la prise de médicaments à base de molécules de synthèse ou bien de remèdes produits à partir de l'environnement naturel. Il existe toutefois de nombreuses autres thérapies, telles la radiothérapie ou la kinésithérapie, n'ayant pas recours à l'ingestion et à l'injection de substances extérieures.

## **II. Le médicament :**

D'après le code de la santé publique (1967), un médicament est « toute substance ou composition présentée comme possédant des propriétés curatives ou préventives à l'égard des maladies humaines ou animales, ainsi que tout produit

pouvant être administré à l'homme ou à l'animal, en vue d'établir un diagnostic médical ou de restaurer, corriger ou modifier leurs fonctions organiques ». La médication (du latin *médicatio* : emploi d'un remède - en anglais *médication*) consiste à employer systématiquement des composants médicaux, appelés également agents médicaux, de façon à améliorer un état pathologique (maladie) [6] .

Autrement dit, il s'agit d'une préparation utilisée dans le but de prévenir, diagnostiquer et soigner une pathologie ou une autre lésion à type de traumatisme entre autres. Le médicament permet également de corriger ou de modifier le fonctionnement d'un organisme.

On peut distinguer différents types de médicaments selon leur utilisation, leurs composants, leur mode d'enregistrement réglementaire, etc. :

**Médicament générique** : Un médicament générique est un médicament identique ou équivalent à celui d'une marque (appelé *médicament princeps*), mais produit et vendu sous sa dénomination commune internationale (DCI, nom chimique de la substance) ou sous un nouveau nom commercial. La substance active (ou principe actif du médicament) en est identique à celle du produit de marque, les seules différences possibles étant la présentation et les excipients.

**Médicament bio-similaire** : Les médicaments bio-similaires sont des médicaments produits par biotechnologie (biothérapies) dont le brevet a expiré. On peut citer ainsi certains ménotropines, constitués à base de gonadotrophine (hormones) ou l'enoxaparine (anticoagulant).

**Médicament orphelin** : On appelle médicament orphelin tout médicament développé pour le traitement de « maladies orphelines » (c'est-à-dire des maladies rares).

**Médicament biologique** : Un médicament biologique est un produit bio thérapeutique issu d'une substance biologique. Ceci inclut, par exemple, les vaccins ou des médicaments dérivés du sang et du sanguin humain.

**Médicament à base de plantes** : La Phytothérapie désigne la médecine fondée sur les extraits de plantes et des principes actifs naturels, une pratique traditionnelle à visée explicitement thérapeutique, parfois très ancienne fondée sur l'utilisation de

plantes selon les vertus découvertes empiriquement. Selon l'OMS, cette phytothérapie est considérée comme une médecine traditionnelle et encore massivement employée dans certains pays dont les pays en voie de développement. C'est le plus souvent une médecine non conventionnelle du fait de l'absence d'études cliniques systématiques

**Médicament essentiel :** Les médicaments essentiels, tels que définis par l'Organisation mondiale de la santé (OMS), sont ceux qui satisfont aux besoins de santé de la majorité de la population. Ils doivent être disponibles à tout moment dans des quantités adéquates et dans des formulations appropriées, à un prix abordable pour les individus et pour la communauté

### **II.1.Composition d'un médicament :**

Le médicament est composé de deux sortes de substances : d'une ou plusieurs substances actives (aussi désigné principe actif, c'est souvent la substance active qui est désignée dans le langage courant par médicament) et d'un ou plusieurs excipients.

La ou les substances actives sont constituées d'une quantité de produit active (dose) ayant un effet pharmacologique démontré et un intérêt thérapeutique également démontré cliniquement. Il est à remarquer que toute substance pharmaco- logiquement active ne constitue pas nécessairement la base d'un médicament et encore moins d'une thérapie médicamenteuse.

Les excipients sont des substances auxiliaires inertes servant à la formulation de la forme galénique ou destinée à créer une absorption par le corps. Ces excipients sont le plus souvent des substances inertes sur le plan pharmacologique.

### **II.2.Catégories thérapeutiques :**

Parmi les médicaments, des familles thérapeutiques sont notamment retrouvées :

- ✓ les anesthésiants pour obtenir anesthésie locale ou générale, en forme topique ou injectable ;
- ✓ les analgésiques (antalgiques), agissant contre la douleur ;
- ✓ les antibiotiques, antimicrobien ayant une activité bactériostatique et/ou bactéricide ;

- ✓ les antidépresseurs, qui traitent la dépression];
- ✓ les antidiurétiques, qui diminuent la sécrétion d'urine (diurèse) ;
- ✓ les anti-inflammatoires agissant contre l'inflammation ;
- ✓ les antihistaminiques agissant contre l'allergie ;
- ✓ les antihypertenseurs qui luttent contre l'hypertension ;
- ✓ les antipyrétiques agissant contre la fièvre ;
- ✓ les antiviraux agissant contre les virus ;
- ✓ les antirétroviraux agissant contre les rétrovirus ;
- ✓ les antitussifs qui luttent contre la toux ;
- ✓ les anxiolytiques, qui réduisent l'anxiété;
- ✓ les bronchodilatateurs, qui vont dilater les bronches ;
- ✓ les diurétiques, qui augmentent la sécrétion d'urine (diurèse) et diminue la tension ;
- ✓ les laxatifs, qui stimulent la défécation ;
- ✓ les psychotropes, pour le traitement des maladies psychiatriques (dont neuroleptiques, anxiolytiques, antidépresseurs, etc
- ✓ les sédatifs (calmants), qui diminuent l'activité d'un organe ;
- ✓ les vasopresseurs, qui font monter la pression artérielle .

### **II.3.Origines des médicaments :**

#### **II.3.1.Origine naturelle (5% de la pharmacopée) :**

Les médicaments peuvent être extraits de : minéraux ou végétaux (95%), ou d'animaux (y compris homme). En fait, le pourcentage de ces derniers a tendance à être réduit à 0, depuis les problèmes du prion (protéine pathogène), de la maladie de la vache folle, SIDA et autres maladies graves qui forcent l'industrie du médicament à se montrer plus prudente avec les échanges directs de substance animale à l'homme. Ils sont de plus en plus remplacés par la synthèse biotechnologique (on introduit la séquence de la protéine médicament dans des bactéries pour la reproduire en quantité suffisante)

#### **II.3.2 .Origine synthétique (95% de la pharmacopée) :**

Copie exacte (5%), hémicopie ou hémisynthèse (95%) représentent 60% des médicaments à origine synthétique (une hémisynthèse consiste à assembler des molécules déjà formées). La synthèse originale représente 40 % de la pharmacopée d'origine synthétique.

## II.4. La dénomination des médicaments :

Chacun possède trois noms :

- ✓ Nom chimique : traduction littéral de la formule développée, établie selon les règles de l'Union internationale de chimie (IUPAC) (exemple : phénylethylmalonylurée ; dicarbonate de méthylpropyl propanediol ; acide acétyl salicylique).
- ✓ Nom DCI : Dénomination commune internationale attribuée par l'OMS (phénobarbital ; meprobamate ; aspirine)
- ✓ Nom de marque (Gardinal® (barbiturique antiépileptique) ; Equanil® ; Aspro®)

Le nom DCI [7] possède plusieurs caractéristiques : c'est un nom de longueur limitée, utilisé dans le domaine public (médicament générique), qui possède un segment commun à tous les médicaments du même groupe pharmacologique. Tout médicament doit avoir sur son emballage : le nom de la spécialité (nom de marque), sous celui-ci le nom du principe actif, le nom et l'adresse de l'exploitant, le numéro AMM. Il est généralement indiqué : le mode et la voie d'administration, les éventuelles précautions, mises en garde et le mode de conservation (et la date de péremption).

## III. Identification et validation des cibles thérapeutiques :

la recherche s'initie à partir de l'identification d'une cible thérapeutique. Le mécanisme physiopathologique est disséqué au niveau le plus fin possible afin de discerner les éléments de dysfonctionnement. Longtemps, le processus de reconnaissance des cibles a reposé essentiellement sur des constats empiriques.

Aujourd'hui, génomique, bioinformatique et protéomique permettent d'identifier en amont gènes ou protéines impliqués dans les maladies et susceptibles de devenir des cibles thérapeutiques. Enzymes, protéines, récepteurs : leur action et leur lien de causalité avec la maladie une fois connus, on cherchera soit à bloquer ou à augmenter leur action, soit à combler leur déficit, soit à remplir la fonction qu'ils ne peuvent plus exécuter. Ces connaissances permettent de rationaliser très tôt le processus de recherche de nouveaux médicaments et d'élargir considérablement le champ de cette recherche. On estime en effet couramment aujourd'hui que l'ensemble des médicaments disponibles ne ciblerait, par leur action, que 500 produits génériques. Quand on sait que le génome humain comporte environ 35 000 gènes distincts et environ 200 000 protéines et que nombre d'entre elles sont très certainement impliquées dans le développement des pathologies, [9] on mesure pleinement l'immensité du champ exploratoire qui reste à couvrir.

### **III.1.Le site actif :**

L'effet d'un médicament est lié à l'interaction du médicament avec son site d'action, qui est généralement un récepteur mais qui peut aussi être une enzyme, une protéine de transport, un canal ionique ou un élément non encore identifié. L'interaction entre le médicament et son site d'action implique une reconnaissance mutuelle des 2 protagonistes, le médicament doit avoir une certaine affinité pour son site d'action.

Les différentes cibles des médicaments sont :

- Des récepteurs : ex : les antisécrétoires gastriques sont des antagonistes des récepteurs histaminiques H<sub>2</sub>,
- Des enzymes : ex inhibition de l'enzyme de conversion de l'angiotensine par les IEC,
- Des protéines de transport : elles permettent le transport des ions et petites molécules à travers les membranes
- Certains médicaments agissent par interaction physicochimique : par exemple action osmotique

- Des agents pathogènes tels que les virus, les champignons, les bactéries, parasites, les médicaments agissent sur des cibles spécifiques de ces agents tels que des enzymes ou des récepteurs.

L'interaction du médicament avec son site d'action va entraîner, via des mécanismes de signalisation intracellulaire, un effet pharmacologique quantifiable au niveau de la cellule, d'un organe isolé (ex : contraction d'artère isolée, ...) ou de l'organisme entier (ex : augmentation de la pression artérielle). Cet effet pharmacologique est suivi généralement d'un effet thérapeutique. Il est important de bien dissocier l'effet pharmacologique de l'effet thérapeutique. Par exemple, par définition, les antiagrégants plaquettaires ont, *in vitro*, un effet pharmacologique correspondant à l'inhibition de l'agrégation plaquettaire. L'effet thérapeutique qui résulte de cet effet pharmacologique est de diminuer le risque de thrombose et d'embolie artérielle. La mise en évidence de l'effet thérapeutique est du ressort de la pharmacologie clinique

#### IV. La genèse d'un médicament :

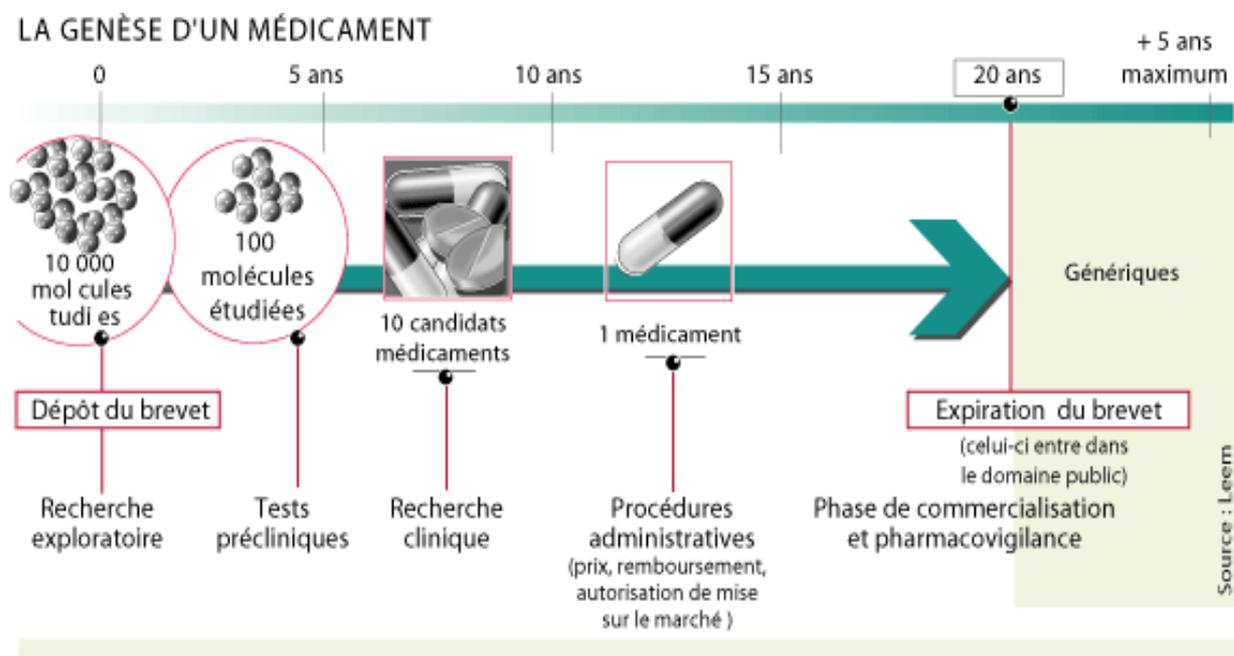


Figure 1 : La genèse d'un médicament

##### IV.1. Première étape :

La première étape est l'identification et le choix d'une cible biologique associée à la maladie que l'on veut soigner. Pour s'attaquer à une maladie donnée, différentes approches peuvent être suivies. Le choix d'en privilégier une par rapport à une autre se fait sur la base de nombreux critères tels que la spécificité de l'approche, les connaissances scientifiques, le caractère innovant, le type de médicament qui en découlera (anticorps, vaccins, peptides, petites molécules, etc.), et bien d'autres encore. Quel que soit le choix, il est fondé sur l'hypothèse d'un lien entre l'action sur la cible biologique et la maladie que l'on veut soigner. Cette cible peut être un gène avec lequel on veut interférer ou une protéine (enzyme, récepteur) que l'on voudra, par exemple, inhiber ou bloquer. [10]

#### **IV.2.Deuxième étape :**

Au cours de la deuxième étape, des tests biologiques spécifiques sont mis en place pour valider la cible thérapeutique, en d'autres termes établir un lien entre l'action sur la cible biologique choisie et une réponse associée à la maladie.

#### **IV.3.Troisième étape :**

La troisième étape consiste à chercher un ou plusieurs leads, ou têtes de séries, c'est-à-dire des molécules qui auraient une affinité détectable pour la cible choisie. L'identification de ces leads se fait par un criblage dit de haut débit qui consiste à tester un très grand nombre de composés sur la cible en question, par des techniques entièrement miniaturisées et robotisées. Ceci permet de tester plusieurs centaines de milliers de molécules par semaine. La principale source de composés est l'échantillothèque d'une compagnie pharmaceutique qui aura, au fil des ans, conservé un échantillon de la plupart des molécules synthétisées dans les différents projets ainsi que d'autres entités chimiques utiles à leur élaboration. La richesse, la diversité et la qualité de cette collection constitue un atout majeur pour l'identification de leads. D'autres sources sont utilisées, comme des banques de produits naturels ou disponibles commercialement. Si le criblage à haut débit présente un intérêt certain, il a également ses limites : il n'est pas accessible à tous les types de cibles, il génère un très grand nombre de données dont le traitement est souvent long et complexe.

Enfin, le rapport entre le nombre de molécules testées et le nombre de molécules bioactives identifiées peut s'avérer très faible. Pour augmenter leurs chances de succès, les chercheurs privilégient de plus en plus une approche de criblage ciblé qui ne porte que sur des collections restreintes, en relation avec la cible thérapeutique visée. Ces collections sont constituées de molécules ayant déjà montré une activité sur un membre de la famille de protéines dont fait partie la cible (par exemple les kinases ou les protéases) ou possédant des caractéristiques chimiques présentes dans ses ligands naturels ou inhibiteurs connus de cette famille. Le développement des méthodes de synthèse automatisée, qui permettent de fabriquer des centaines de molécules en parallèle, alliées aux techniques de modélisation moléculaire et de chimie informatique, contribue de façon majeure à augmenter la richesse et la diversité de ces collections, et par là même l'efficacité du criblage.

A ce stade, les leads identifiés n'ont encore rien d'un médicament. Ils possèdent cependant un certain nombre de propriétés qui font que l'on va s'intéresser à eux de plus près : ils ont certes un certain niveau d'activité in vitro et cellulaire sur la cible choisie, mais ils possèdent également un profil dit de drugability : leur structure chimique et leurs propriétés physico-chimiques (par exemple la solubilité et le poids moléculaire qui auront un impact sur l'absorption et la biodisponibilité, ou encore la structure chimique qui déterminera partiellement la stabilité, le métabolisme, la toxicité) sont potentiellement exploitables pour en faire un médicament (drug).

#### **IV.4. Quatrième étape :**

La quatrième étape sera donc consacrée à l'optimisation du lead, phase durant laquelle les chimistes thérapeutiques utiliseront tout l'arsenal de leurs connaissances et de leur expérience pour améliorer l'activité, les propriétés physico-chimiques, pharmacocinétiques et transformer un lead en médicament. Parallèlement, les biologistes auront mis au point une cascade de tests, de plus en plus sophistiqués, permettant d'évaluer l'activité intrinsèque des composés sur la cible isolée, de mettre en évidence l'activité cellulaire et la sélectivité, de déterminer le mode d'action, de mesurer les pharmacocinétiques et enfin l'efficacité dans des modèles in vivo. L'optimisation d'un lead est un processus itératif et rationnel. A partir de la structure initiale du lead, des modifications chimiques sont apportées, de nouvelles molécules sont synthétisées et leur activité biologique est évaluée.

Les résultats sont analysés, mettant en évidence les modifications bénéfiques ou au contraire néfastes au profil biologique. On établit ainsi des relations entre la structure des molécules et leur activité, relations qui vont guider le chimiste dans ses choix rationnels de nouvelles modifications. Ce travail aboutira après deux ou trois ans à une ou deux molécules qui répondront aux critères requis pour entrer dans la phase de développement clinique.

#### **IV.5.Cinquième étape :**

La cinquième étape est consacrée au développement clinique. Un élément-clé de cette nouvelle étape est de valider l'efficacité de la molécule chez l'homme. Si son efficacité est démontrée et sa tolérance bonne, la molécule poursuit alors les différentes phases du développement clinique, pour finalement être commercialisée, au minimum dix ans après le début de son histoire.

### **V. Processus de découverte :**

#### **V.1.Découverte fortuite :**

Entre hasard et nécessité... La pharmacologie et la fabrication de médicaments ne tarissent pas d'anecdotes parfois troublantes quant à la découverte d'un principe actif. Qui aurait cru que le caoutchouc pouvait soigner le mal de l'alcool? Que de puissants explosifs pouvaient guérir l'angine de poitrine? Que des animaux morts de façon mystérieuse nous apprendraient l'usage d'un médicament qui sauve de nombreuses vies désormais... [11]

Botanistes, ethnologues, éthologues... De nombreux professionnels ont pu observer des situations dans lesquelles une substance créait des conditions organiques particulières (soins, "vaccinations", morts subites, guérisons "miraculeuses", ...). Par l'observation plus rigoureuse, ces "fruits du hasard" ont donné naissance à de tout aussi nombreux médicaments.

\*Cobayes involontaires pour anecdotes médicales

30% des principes actifs utilisés ont été découvert par leurs effets chez l'homme.

Depuis plusieurs siècles, certaines médecines traditionnelles dont on ignore l'origine furent passées au crible de la science afin d'objectiver leurs effets, il pouvait s'agir :

- de substances naturelles ou médecine indigène : c'est le domaine de l'ethnopharmacologie
- de remèdes populaires
- d'effets secondaires de certains médicaments

La nitroglycérine...

...provoque des explosions, qui peuvent provoquer des crises cardiaques, bien entendu, et pourtant... on s'est rendu compte que les ouvriers qui la manipulaient, et avaient des problèmes cardiaques, voyaient ces derniers s'améliorer. De fait, on s'est aperçu que cette substance possède des propriétés vasodilatatrices très puissantes, on s'en est donc servi pour traiter l'hypertension.

La trinitrine, par exemple, est un des meilleurs médicaments contre l'angine de poitrine. D'ordinaire, les artères coronaires vieillissent et perdent leur souplesse de vasodilatation, ce qui provoque une douleur vive caractéristique, qui se propage dans le bras gauche, lorsque l'on fait un effort trop grand. On donne au patient, depuis cette découverte, des pastilles à mettre sous la langue, afin que le composé (proche de la nitroglycérine) passe directement dans le sang via les jugulaires, qui vont directement au cœur.

Désintox au caoutchouc

Le Disulfirame® : Découvert chez des ouvriers travaillant dans l'industrie du caoutchouc : certains d'entre eux (tout un groupe) ne supportaient plus l'alcool. On s'est rendu compte que le disulfirame, contenu dans le caoutchouc, agit sur le foie et bloque la dégradation de l'alcool en acide acétique, puis en composé psychotrope (blocage du 2ème processus). L'acide acétique s'accumule dans le foie et provoque des nausées et des vomissements. Une vaccination adéquate

pour l'absorption massive...

Sorcellerie d'autrefois, pharmacie d'aujourd'hui

L'Apomorphine. Elle servait à des tests d'épreuves dans l'antiquité, utilisée par les sorciers ; avec une dose spécifique, elle provoque des vomissements. Si elle est ingérée doucement toutefois, elle peut provoquer la mort par détresse respiratoire. Mais comment s'en servait les sorciers? Il s'en servait pour distinguer le menteur de l'honnête homme : lorsque deux congénères étaient soupçonnés d'un crime ou d'un délit, le sorcier leur faisait avaler la mixture contenant l'apomorphine, sans oublier de dire que "la substance tuait le menteur, et laissait l'homme honnête en vie". Connaissant alors le pouvoir du breuvage, l'honnête, sans crainte, absorbait d'un coup le poison, qu'il vomissait quelques minutes plus tard. Le menteur, sachant pertinemment que la substance serait dangereuse pour lui, buvait par petite gorgée... Se faisant, au lieu de vomir l'apomorphine, il lui permettait de se diffuser dans son organisme...

L'effet vomitif vient de son action sur les récepteurs dopaminergiques des noyaux vestibulaires (impliqués dans l'équilibre, par exemple, le mal de mer). Grâce à ces effets, on l'utilisa ensuite pour traiter le mal de mer, puis on s'est aperçu qu'elle avait un effet sur les parkinsoniens, en réduisant leurs tremblements, notamment. On s'est ensuite aperçu que ces mêmes parkinsoniens retrouvaient toute leur ardeur dans un tout autre domaine... et que l'apomorphine facilitait donc la vasodilatation périphérique. Un médicament qui n'en finit pas d'étonner!

D'autres exemples :

L'infusion de feuille de saule (acide acétylsalicylique - et donc aspirine) est un exemple de médicament emprunté aux remèdes de Grand-mère.

La peau de grenouille est un puissant hémostatique (elle contient de l'adrénaline, vasoconstrictrice)

La pénicilline a été découverte par hasard par Sir Alexander Fleming en 1928 lorsqu'il remarqua des moisissures aux propriétés étranges : autour d'elles, un

halo protecteur sur laquelle bactérie et autres moisissures ne parvenaient à croître...

#### Fléchettes anesthésiantes

Les flèches empoisonnées au curare des indiens d'Amazonie nous ont montré combien le curare est un puissant relaxant musculaire. Il est utilisé en anesthésie pour éviter notamment le réflexe Iléus (contraction des muscles abdominaux lors d'une incision), car il bloque les récepteurs cholinergiques des muscles, d'où un relâchement musculaire complet. Des précautions et une bonne connaissance de ce médicament sont nécessaires pour son utilisation : il peut provoquer un arrêt respiratoire (souvent, les curarisés sont placés sous assistance respiratoire). Il peut également y avoir une intolérance au curare, ou un dosage trop élevé... On ne sait pas évaluer l'allergie au curare.

#### Observations chez l'animal

Exemples : La coumarine (anticagulant, ex : Dicoumarol®). Dans les années 30, tous les bovins d'une vaste région des USA sont morts de manière incompréhensible. Printemps et été précédents avaient été particulièrement pluvieux, le foin sensé être séché a été stocké encore humide, et le trèfle qu'il contenait s'est décomposé en coumarine sous l'effet de l'humidité. Tous les animaux sont morts d'hémorragie...

La vinblastine (anticancéreux pour certaines leucémies. Utilisée au début en recherche pour traiter le diabète, on s'est aperçu que les rats diabétiques du laboratoire mourraient de septicémie. Or le système immunitaire des rats est un des meilleurs de toutes les espèces (on peut les opérer en milieu non-stériles, ça ne les affectera presque pas). La vinblastine provoque la mort des globules blancs et donc une immunodépression. Elle est donc utilisée en chimiothérapie contre certaines leucémies, dans lesquelles il y a hyperlymphocytose.

#### Observations du règne végétal

Exemple : l'indométacine, qui n'est pas de la famille de la cortisone, a été découverte à partir de l'observation de plantes.

**V.2.Découverte ciblée :**

Le plus fréquent. On va cibler une pathologie puis chercher des molécules qui vont agir sur le système. C'est trouver des molécules capables d'interagir avec un système physiopathologique connu c'est le cas du HTS [12] (= High Throughput Screening) le criblage à haut débit ou, plusieurs milliers de molécules appartenant à une « librairie chimique » sont choisies au hasard ou en fonction de certaines caractéristiques physico-chimiques leur permettant d'interagir avec une cible prédéfinie (récepteur ou enzyme). L'activité des molécules vis-à-vis de la cible choisie se fait in vitro (boîtes multi-puits). Les « librairies de molécules » sont approvisionnées à partir de molécules issues de la synthèse organique de synthèse ou isolées du monde vivant (végétaux, microorganismes terrestres ou marins,...). Le HTS va mettre en évidence quelques molécules ou « hits » qui possèdent l'activité recherchée (agoniste, antagoniste, inhibiteur,...) sur la cible étudiée.

**VI. Développement prés-clinique :**

Avant son administration à l'Homme, le candidat médicament devra subir toute une série de test sur des modèles cellulaires et animaux pour démontrer son efficacité et son innocuité. Les compagnies pharmaceutiques sont en effet obligées de démontrer la sécurité du médicament pour les humains et de prouver que les avantages thérapeutiques valent largement les risques associés aux effets indésirables causés par le composé. [13]

Les différentes études effectuées lors de la phase préclinique se font en parallèle.

**VI.1.Etude pharmacodynamique :**

Quel que soit le médicament que l'on désire sélectionner, un analgésique, un antibiotique etc., il est nécessaire :

- de déterminer de manière approfondie son mécanisme d'action,
- de préciser par une étude systématique tout autre effet concomitant éventuel sur les autres appareils : cardio-vasculaire, respiratoire, rénal, etc.

Ceci explique qu'une exploration systématique de tous les effets possibles d'un médicament soit nécessaire. Ces études, qu'il n'est pas possible de citer ici, se font sur l'animal entier, sur des organes isolés, des cellules isolées, des fractions cellulaires isolées, enzymes, récepteurs. Elles précisent les propriétés et les mécanismes d'action des médicaments.

### **VI.2. Etude pharmacocinétique chez l'animal :**

Elle évalue l'absorption, la distribution, le métabolisme et l'élimination du candidat-médicament. Elle est effectuée sur des modèles in vitro et chez l'animal.

### **VI.3. Etude toxicologique :**

Cette étude est menée partiellement in vitro et en partie chez l'animal. On distingue les études de toxicité aiguë (dose unique) et les études de toxicité subaiguë (doses répétées sur plusieurs semaines ou mois). S'y ajoutent les études de toxicité sur la reproduction, de tératogénèse, de mutagenèse, de carcinogénèse, de tolérance locale et de sensibilisation

La toxicité ne sera pas seulement jugée sur la survie ou la mort des animaux, mais aussi sur des critères cliniques (évolution pondérale, activité motrice, état tégumentaire,...). Les animaux seront sacrifiés à l'issue de cette période pour examen macroscopique des viscères. Au moindre doute à la nécropsie, un examen histopathologique devra être réalisé. Dans certains cas, des examens biologiques pourront être pratiqués. La dose maximale sans effet toxique sera la dose la plus élevée pour laquelle aucun effet toxique significatif clinique, biologique ou anatomopathologique n'aura été relevé par rapport au(x) lot(s) d'animaux témoins.

**VI.3.1. Toxicité aiguë** : administration unique, 2 espèces, recherche de la dose maximale tolérée (DMT) => permet de déterminer le seuil toxique

**VI.3.2. Toxicité subaiguë** : 4 semaines, 3 espèces, 2 voies d'administration, minimum 3 doses < DMT dont 1 dose thérapeutique.

**VI.3.3.Toxicité chronique** : 52 semaines, 2 espèces dont 1 non-rongeur, 1 voie d'administration, 3 doses

**VI.3.4.Toxicité de la reproduction :**

- traitement des géniteurs jusqu'au terme de la gestation
  - examen de la fertilité
  - examen de la progéniture (vision, ouïe...)
  - examen de la fertilité de la progéniture
- administration pendant l'organogenèse (6 → 17ème jour)
  - fœtotoxicité (léthalité)
  - tératogenèse (malformation)
- administration du 17ème jour de la gestation jusqu'au sevrage
  - parturition
  - galactogenèse
  - viabilité et développement post-natal des nouveau-nés

**VI.3.5.Mutagenèse** : étude d'un changement héréditairement transmissible de la séquence ou du nombre de nucléotides de l'ADN.

**VI.3.6.Cancérogenèse** : - traitement de longue durée (>2,5 ans)

- grand nombre d'animaux
- gavage impossible
- difficulté de différencier les tumeurs spontanées des néoformations imputables au traitement.

**VI.4.Etude technique (ou galénique):**

Cette phase inclut l'adaptation de la méthode de préparation du produit à la production de quantités élevées et le développement de formulations galéniques

adéquates. La galénique est la science relative à la mise en forme de médicaments. Cette phase aboutit à la conception de la spécialité pharmaceutique.

### **VI.5. Etude analytique :**

Cette phase développe et valide les méthodes d'analyse qui permettront de quantifier le principe actif et ses métabolites dans les fluides corporels (sang, plasma, urine) et les fèces.

### **VII. Développement clinique :**

Au terme de l'étude chez l'animal, on décidera si la molécule étudiée mérite d'être essayée chez l'homme ou doit être abandonnée. L'analyse complète des résultats permet d'évaluer le rapport efficacité/toxicité. La mise en route des essais chez l'homme dépendra de ce rapport et de sa comparaison avec ceux de principes actifs de même type, éventuellement déjà commercialisés. Ce dossier expérimental constitue le prérequis à un éventuel essai du candidat médicament chez l'homme.

La pharmacologie clinique évalue les propriétés des médicaments (efficacité, tolérance, pharmacocinétique...) très souvent chez l'homme sain et toujours chez le malade.

Cette évaluation s'effectue par des essais dont les protocoles doivent répondre à une rigueur suffisante pour aboutir au but fixé sans nuire à l'état de santé des volontaires sains ou des malades qui y participent, et être conformes à l'éthique

Le consentement éclairé, donné par écrit, des volontaires qui participent à l'étude nécessite qu'ils soient bien informés du déroulement de l'étude et des risques éventuels qu'elle comporte. Les études doivent être menées selon une méthodologie rigoureuse selon des normes sévères émises par des organismes de réglementation tels la Food and Drug Administration (FDA) aux Etats-Unis et

l'Agence Européenne du Médicament (EMA) en Europe. Avant toute mise en route, un comité d'éthique doit également se prononcer favorablement.

### **VII.1. Etude clinique Phase I :**

Le candidat-médicament est administré à l'homme pour la première fois. L'objectif premier est d'évaluer la sécurité du produit et de déterminer l'échelle des doses tolérées. En principe, le but de cette phase n'est pas de mettre en évidence l'efficacité du produit, bien sa tolérance. La phase I permettra aussi d'effectuer une première évaluation pharmacocinétique (recherche du changement subi par la substance active dans l'organisme) chez l'homme. Ces essais sont en général réalisés sur nombre restreint de volontaires et sains (100-200) répartis en petits groupes (10-20) et par des médecins ayant une compétence en pharmacologie clinique. Il n'est cependant pas licite d'administrer certains médicaments ayant une toxicité importante bien connue, tels que les anticancéreux actuels, à des volontaires sains et ils sont testés d'emblée chez des malades. Si le médicament se montre plus toxique que prévu, ou encore est mal toléré, les essais sont immédiatement arrêtés. Le composé est retiré du programme. En général 1-3 molécules sont retenues pour la phase II.

### **VII.2. Etude clinique de Phase II :**

Ces essais visent (i) à établir l'efficacité du candidat-médicament et à éliminer les produits trop peu actifs ; (ii) à optimiser la posologie et la durée du traitement pour les études ultérieures ; (iii) à étudier les effets indésirables et les risques associés à une utilisation à court terme. Cette étude est réalisée avec un nombre limité de patients sélectionnés (de 100 à 400). Cette phase dure de 1 à 2 ans.

### **VII.3. Etude clinique Phase III :**

Les essais de la phase III rassemblent un échantillon de plusieurs milliers de patients. Les essais cliniques de phase III ont plusieurs objectifs :

- faire la preuve de l'efficacité thérapeutique du composé dans les conditions réelles d'utilisation ;
- définir encore plus précisément le profil de sécurité du médicament, les

incidences d'effets indésirables, et sa toxicité à long terme ;

- démontrer que le composé est aussi sinon plus efficace que les médicaments ou traitements existants ;
- comparer rigoureusement des groupes de patients traités avec le candidat médicament, avec des groupes de contrôle à qui on administre un placebo ;
- ajuster la posologie, selon l'âge des patients, les pathologies dont ils souffrent et les autres médicaments qui leur sont administrés ;
- tester différentes formulations pour la prise de médicaments (comprimé, sirop, vaporisateur, etc.)

#### **VII.4.Types d'études cliniques :**

L'étude clinique doit être adaptée à chaque produit et à chaque classe de médicament. Une étude clinique peut-être : « contrôlée » - « randomisée » - « ouverte » - « en simple aveugle » - « en double aveugle » - « multicentrique ».

**Etude contrôlée** : comparaison du candidat médicament à un autre traitement (autre principe actif ou placebo). Le placebo est une forme pharmaceutique identique au candidat médicament (forme, couleur, goût,...) mais ne contenant pas de principe actif.

**Etude randomisée** : les groupes de patients soumis aux différents traitements sont constitués par tirage au sort.

**Etude ouverte** : le patient et le médecin connaissent la nature du traitement.

**Etude en simple aveugle** : seul le médecin connaît la nature du traitement.

**Etude en double aveugle** : ni le médecin, ni le patient ne connaissent le traitement.

**Etude multicentrique** : l'étude clinique est effectuée dans plusieurs centres hospitaliers généralement de pays différents.

**Etude séquentielle** : le traitement subi change au cours du temps.

**VII.5. Parcours administratif :**

En cas de succès au terme des études cliniques, le candidat-médicament doit encore parcourir plusieurs étapes administratives avant d'être mis sur le marché.

Tout nouveau médicament à usage humain doit parcourir au minimum deux procédures administratives:

- l'autorisation de mise sur le marché
- la procédure de fixation de prix.

Si l'entreprise souhaite obtenir le remboursement d'un médicament, celui-ci doit encore parcourir deux procédures supplémentaires:

- la procédure de transparence,
- la procédure de remboursement.

**VII.6. Pharmacovigilance ou phase clinique IV :**

Cette phase s'occupe du suivi du médicament après sa mise sur le marché, en ce compris et en particulier la détection de réactions indésirables au médicament. A cet effet, tout praticien (médecin, pharmacien) à l'opportunité de rapporter au Centre de Pharmacovigilance les effets indésirables qu'il suspecterait être occasionnés par un médicament ou une association de ceux-ci.

Le développement d'un médicament est donc est parcours long complexe, à haut risque et généralement rentable.

**VIII. Fabrication :**

Le risque est grand de ne pas pouvoir poursuivre le développement d'un médicament en raison d'effets indésirables observés au cours des études cliniques ou en raison d'une efficacité insuffisante. Sur 10 000 substances étudiées et contrôlées en laboratoire, 10 parviennent à la phase des études cliniques, et seule une satisfait finalement à tous les tests cliniques et parvient plus tard au stade de la commercialisation en tant que médicament. [14]

Dès qu'un médicament a été autorisé, on peut commencer à le fabriquer pour le vendre. On commence par isoler la substance active, puis on utilise des excipients (substances auxiliaires) pour produire une forme galénique déterminée telle que comprimé ou pommade. Enfin, le médicament est conditionné à la machine, c'est-à-dire placé dans un emballage spécialement conçu à cet effet. Les médicaments sont fabriqués par lots. Si jamais un lot de médicaments présente un défaut, le numéro de lot permet de rappeler les médicaments en cause.

Le consommateur doit avoir la garantie que la sécurité des médicaments est irréprochable

## **Chapitre 3 :**

**Conception d'une application informatique pour la  
génération d'une chimiothèque**

## **I. Introduction :**

La chimie combinatoire est apparue naturellement comme une option viable au problème de la diversité moléculaire. Aujourd'hui, c'est un moyen pratique pour prédire et synthétiser une grande quantité de molécules en chimie pharmaceutique.

Comme moteur de diversité, cet outil est devenu indispensable et a joué un rôle important dans le progrès de la synthèse automatique et parallèle survenu ces vingt dernières années. Cette méthode repose sur l'idée d'obtenir le plus grand nombre de produits possible d'une réaction particulière et ceci sous certaines conditions

L'organisation, l'analyse, la recherche et la gestion de cette grande quantité d'informations ouvrent de nouvelles possibilités aux techniques novatrices de chimie informatique.

## **II. Présentation du projet :**

### **II.1.Cahier des charges :**

#### **II.1.1.Constataion de départ :**

L'idée de créer une application informatique de génération d'une bibliothèque chimique est partie du constat que l'introduction de l'outil informatique est devenue une nécessité dans la recherche thérapeutique et la conception des médicaments.

L'avantage de ce genre d'application réalisée est aussi de réduire le temps de conception en phase de criblage et tests biologiques

#### **II.1.2.Définition de l'application à réaliser :**

Le projet consiste à concevoir une application informatique pour la génération d'une nouvelle bibliothèque de molécules « Chimiothèque » à l'aide de la synthèse combinatoire à partir de quelques réactifs de départ à fin de reproduire de nouvelles entités chimiques qui auront toutes les chances d'être le future médicament.

#### **II.1.3.Fonctions essentielles que devra remplir l'application :**

Elle doit remplir d'office certaines fonctions indispensables à la reproduction de nouveaux produits à la fin. En effet, n'oublions pas que son but premier est de faire gagner du temps tout en évitant de faire la reproduction manuellement

Avant toute chose, il faut constater que la saisie des éléments doit respecter quelques instructions :

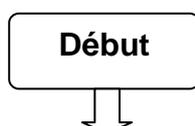
- Les réactifs de départ doivent être présentés sous forme semi développée
- La semi développée doit être présentée sous forme de chaîne linéaire non ramifiée
- Les ramifications doivent être introduites dans la chaîne linéaire tout en gardant l'ordre
- Le programme prend en compte que les réactifs avec simple liaison

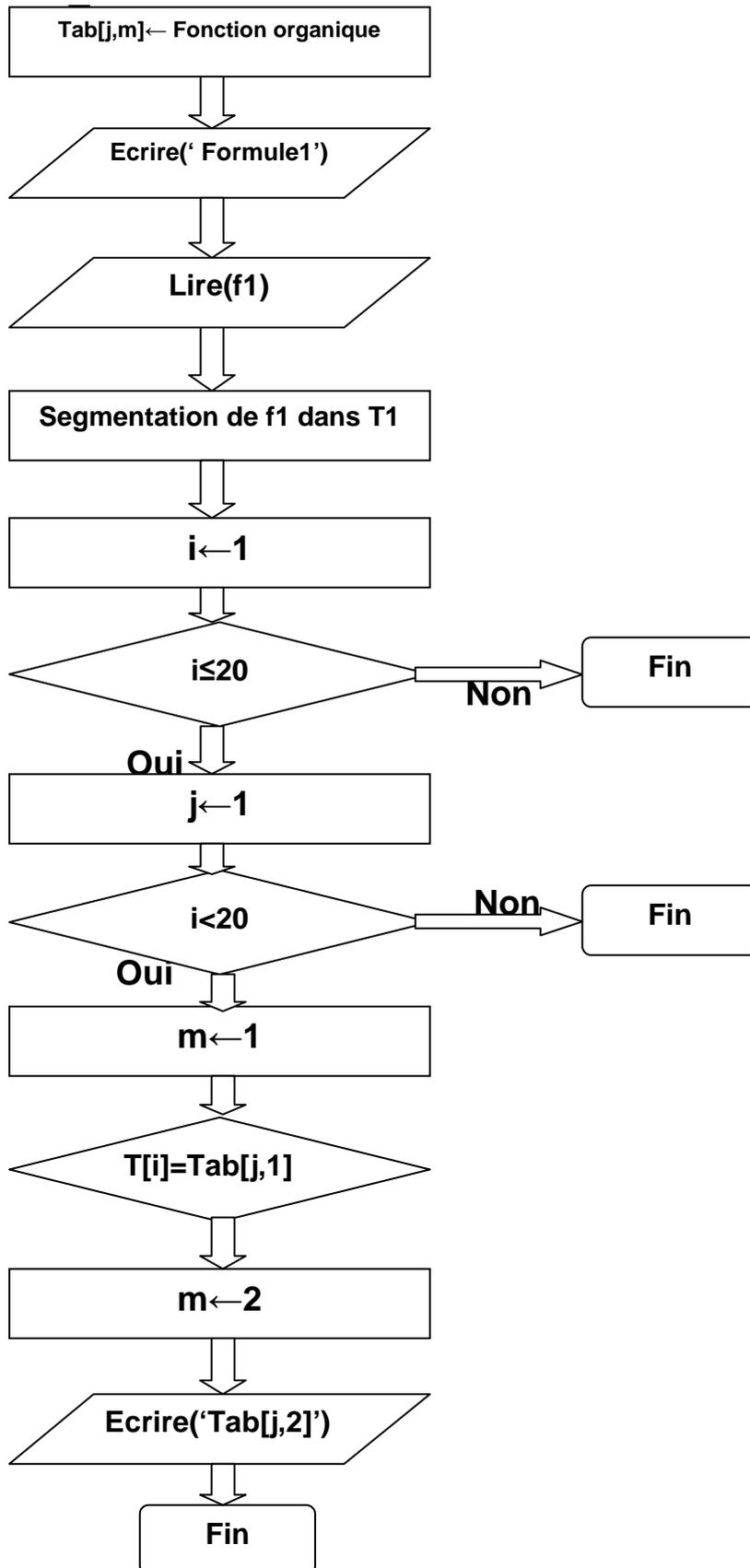
## **II.2. Conceptualisation de l'application :**

A présent, nous avons un cahier des charges précis de ce que nous souhaitons obtenir, il ne reste plus qu'à effectuer les choix techniques :

### **II.2.1. L'organigramme :**

On peut résumer le projet dans l'organigramme suivant :





**II.2.2.L'algorithme :**

Algorithme synthese\_combinatoire ;

Variables

Taborg[1..15 , 1..2] de chaine de caracteres ;

F1 ,F2 : chaine de caractere ;

T1[1..20] de chaine de caractere ;

T2[1..20] de chaine de caractere ;

l,j,k,m :entiers ;

Debut

Taborg[1,1]←COOH;

Taborg[1,2]←Acide carboxilique;

Taborg[2,1]←COCl;

Taborg[2,2]←Halogenure d'acide;

Taborg[3,1]←COOC;

Taborg[3,2]←Ester;

Taborg[4,1]←CONH2;

Taborg[4,2]←Amide;

Taborg[5,1]←CN;

Taborg[5,2]←Nitril;

Taborg[6,1]←CHO;

Taborg[6,2]←Aldehyde;

Taborg[7,1]←CO;

Taborg[7,2]←Cetone;

Taborg[8,1]← OH;

Taborg[8,2]←Alcool;

Taborg[9,1]←S H;

Taborg[9,2]←Thiol;

```
Taborg[10,1]←NH2;
Taborg[10,2]←Amine;
Ecrire('Donner la 1ere formule: ');
Lire('F1') ;
Ecrire('Donner la 2eme formule: ');
Lire('F2') ;
Seg('F1') à ['-'] // segmentation des 2 formules F1 et F2
// parcourir les deux tableaux T1 et T2 et les comparer avec Taborg
Pour i←1 à 20 faire
    Debut
        Pour j ←1 à 20 faire
            Pour m ← 1 à 2 faire
                Si (T[i] = Taborg[j,1] alors
                    Ecrire('Tab[j,2]') ;
            Fin ;
    Fin ;
//synthèse par croisement des deux tableaux T1 et T2 et reproduction des produits
S1←0 ;
S2←0 ;
Pour i ←1 à 20 faire S1←S1 + T1[i]
Pour k ←1 à 20 faire S2←S2 + T2[k]
S3←0 ;
S4←0 ;
Pour e ←1 à LONGUEUR T1[k] faire S3←S1 + T1[k]
Pour n ←LONGUEUR T1[i] à 0 faire S4←S2 + T1[i]
Fin.
```



## Conclusion

---

### **Conclusion :**

Le travail présenté était sur l'introduction de l'outil informatique dans la recherche des entités chimique à intérêt thérapeutique.

L'étude du processus de la conception nous a montré d'où la nécessité de l'ordinateur dans les différentes étapes de recherche

Cette étude nous a permis aussi de voir que les grandes industries pharmaceutiques ont constatés que l'introduction des systèmes robotisés permet de minimiser le temps de calcul dans de différentes phases de développement

Cela n'est pas suffisant pour bien comprendre les interactions ligand-protéine d'où l'introduction de la dynamique moléculaire dans le processus de criblage virtuel et donc pour bien compléter l'étude des interactions ligand-protéine nous pourrons faire appel aux méthodes de chimie quantique

## **Bibilographie :**

[1] : <http://www.interpharma.ch/de/index.asp>

[2] : <http://www.genopole.org/media/pdf/fr/comprendre/060606-medicaments-futur-PT.pdf>

[3] : <http://fr.wikipedia.org/wiki/Maladie>

[4] : <http://www.afriquebio.com/pages/des-remedes-par-maladie/origine-de-la-maladie.html>

[5] : <http://www.larousse.fr/encyclopedie/divers/maladie/67643>

[6] : <http://fr.wikipedia.org/wiki/M%C3%A9dicament>

[7] : <http://www.psychoweb.fr/articles/divers/216-un-medicament-c-est-quoi.html>

[8] : Lindsay MA. Target discovery. Nat Rev Drug Discov. 2003;2:831-8.

[9] : Lindsay MA. Finding new drug targets in the 21st century. Drug Discov Today. 2005;10:1683-7.

[10] : <http://www.jle.com/e-docs/00/04/06/18/article.phtml>

[11] : <http://www.psychoweb.fr/articles/divers/250-fabrication-de-medicaments-par-valorisation-d-informations-biolog-2.html>

[12] : <https://webcampus.fundp.ac.be/claroline/backends/download.php?url=L05PVEV TX0RFX0NPVVJTL0NoYXBpdHJILTA2LnBkZg%3D%3D&cidReset=true&cidReq=M PHAB100>

[13] : <https://webcampus.fundp.ac.be/claroline/backends/download.php?url=L05PVEV TX0RFX0NPVVJTL0NoYXBpdHJILTA2LnBkZg%3D%3D&cidReset=true&cidReq=M PHAB100>

[14] : <http://www.interpharma.ch/fr/medicaments/2642-medicaments-genese-et-fabrication>

- [15] : Hou T, Xu X. Recent development and application of virtual screening in drug discovery: an overview. *Curr Pharm Des.* 2004;10:1011-33.
- [16] : Dobson CM. Chemical space and biology. *Nature.* 2004;432:824-8.
- [17] : <http://bibliothek.fzk.de/zb/berichte/FZKA6558Dateien/FZKA6558.pdf>
- [18] :  
<http://bibliothek.fzk.de/zb/berichte/FZKA6558Dateien/FZKA6558%20Conclusion.pdf>
- [19] : Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov.* 2007;6:881-90.
- [20] : Wermuth CG. Selective optimization of side activities: another way for drug discovery. *J Med Chem.* 2004;47:1303-14.
- [21] : Newman DJ, Cragg GM. Natural Products as Sources of New Drugs over the Last 25 Years. *Journal of Natural Products.* 2007;70:461-477
- [22] : Schreiber SL. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science.* 2000;287:1964-9.
- [23] : Petitjean M. Three-Dimensional Pattern Recognition from Molecular Distance Minimization. *Journal of Chemical Information and Computer Sciences.* 1996;36:10381049.
- [24] : Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr. et al. CHARMM: the biomolecular simulation program. *J Comput Chem.* 2009;30:1545-614.
- [25] : Hert J, Willett P, Wilton DJ et al. Comparison of topological descriptors for similarity-based virtual screening using multiple bioactive reference structures. *Org Biomol Chem.* 2004;2:3256-66.
- [26] : Rarey M, Dixon JS. Feature trees: a new molecular similarity measure based on tree matching. *J Comput Aided Mol Des.* 1998;12:471-90.
- [27] : Hartmann C, Antes I, Lengauer T. Docking and scoring with alternative side-chain conformations. *Proteins.* 2009;74:712-26.

[28] : McGaughey GB, Sheridan RP, Bayly Cl et al. Comparison of topological, shape, and docking methods in virtual screening. *J Chem Inf Model.* 2007;47:1504-19.

[29] : Renner S, Noeske T, Parsons CG, Schneider P, Weil T, Schneider G. New allosteric modulators of metabotropic glutamate receptor 5 (mGluR5) found by ligand-based virtual screening. *Chembiochem.* 2005;6:620-5.

[30] : Renner S, Schwab CH, Gasteiger J, Schneider G. Impact of conformational flexibility on three-dimensional similarity searching using correlation vectors. *J Chem Inf Model.* 2006;46:2324-32.