

Cour 07 : Le criblage virtuel

Introduction :

Le criblage virtuel, également appelé criblage in silico, est une méthode utilisée en chimie pharmaceutique et en biologie structurale pour identifier des composés potentiellement actifs contre une cible biologique.

La conception d'un nouveau médicament est un processus très long et coûteux, allant de l'identification d'une cible biologique d'intérêt thérapeutique jusqu'au patient, dans lequel les essais cliniques succèdent au développement préclinique. Par le passé, la méthode classique pour concevoir un nouveau médicament est basée sur une méthode quasiment expérimentale « le criblage à haut débit » (« HTS » en anglais).

À ce titre, le criblage virtuel par docking moléculaire est une nouvelle approche visant à simuler et à prédire l'affinité d'un très grand nombre de ligands pour le site actif d'une cible thérapeutique donnée, ce qui est assez facile à mettre en place, plus rapide, pour un coût bien moindre que les criblages expérimentaux.

L'intérêt fondamental du criblage (screening) virtuel est d'aboutir à une liste très réduite de molécules présentant l'activité souhaitée. Seules, ces molécules seront testées expérimentalement. Une telle optimisation doit permettre de raccourcir le temps entre la définition d'un besoin thérapeutique et la délivrance du médicament correspondant au patient, en gagnant du temps et d'argent.

Le criblage virtuel, analogue in silico de l'HTS, peut schématiquement être assimilé à un entonnoir dans lequel on verse un grand nombre de composés, constituant la chimiothèques à cribler, pour obtenir, à l'aide d'un algorithme de criblage, un plus faible nombre de composés qui seront ensuite testés expérimentalement.

1. Contexte et Nécessité du Criblage Virtuel :

- Avec l'explosion de la disponibilité des données biologiques et des avancées technologiques, le criblage virtuel s'est imposé comme un moyen efficace d'accélérer le processus de découverte de médicaments.
- Il répond au besoin pressant de réduire les coûts et les délais associés à la recherche de nouveaux médicaments.

2. Différences entre Criblage Virtuel et Criblage Expérimental :

- Alors que le criblage expérimental implique des tests réels en laboratoire, le criblage virtuel utilise des simulations informatiques pour prédire l'interaction entre les molécules.
- Le criblage virtuel permet d'explorer un grand nombre de composés avant même leur synthèse ou leur acquisition physique.

3. Objectifs et Avantages du Criblage Virtuel :

- Les objectifs principaux incluent l'identification rapide de candidats médicamenteux potentiels et la réduction des échecs en phase préclinique.
- Les avantages comprennent la rapidité, l'économie de coûts, et la possibilité de cibler des interactions moléculaires spécifiques.

4. Historique et Évolution du Criblage Virtuel :

- L'émergence de l'informatique et des techniques de modélisation moléculaire a ouvert la voie au criblage virtuel dans les années 1970.
- Les progrès récents dans les technologies informatiques et la disponibilité croissante de bases de données moléculaires ont considérablement renforcé cette approche.

5. Le rôle du Criblage Virtuel :

Le rôle des méthodes de criblage est donc d'éliminer les composés supposés inactifs ou les molécules indésirables tout en priorisant les composés les plus susceptibles d'être actifs.

Tout criblage virtuel se décompose en trois étapes importance :

- La mise au point de la chimiothèque de départ, celle-ci peut être soit une collection de criblage (molécules disponibles) soit une collection virtuel (molécule à synthétisé)
- Le criblage proprement dit, qui consiste à prédire à la fois la conformation et l'orientation relative de chacune des molécules de la chimiothèque, par rapport à la cible d'intérêt
- La sélection d'une liste de touches virtuelles. Après différents traitements permettant d'éliminer les « faux positifs », la sélection finale des molécules à évaluer passe par un examen individuel des interaction 3D de chaque hit virtuel avec le récepteur.

6. Les outils du criblage virtuel

6.1 Cible

Les structures 3D des différentes protéines sont disponibles et accessibles gratuitement dans la banque de données PDB. (« Protein Data Bank », en anglais). Il s'agit de la plus grande archive de données structurales de macromolécules biologiques (protéines, ADN, ARN ...etc.) avec la mise à disposition de plus de 160000 structures 3D. Elles ont été résolues pour la grande majorité par cristallographie aux rayons X.

6.2 Ligand

À l'heure actuelle, il existe deux moyens pour obtenir la structure chimique d'un ligand donné : **Le premier moyen** est souvent d'aspect commercial, est constituée de bases de données de structures chimiques appelées chimiothèques ou espaces chimique. **Le second moyen** consiste à utiliser des ligands issus de la littérature qu'on peut dessiner, optimiser et enregistrer dans différents formats (pdb, mol, mol2...etc.) grâce à des logiciels de construction moléculaire tels que ChemDraw, Arguslab, Titan ou Sybyl ...etc. Il est à souligner que les programmes Titan et ChemDraw ont été utilisés dans le présent travail en vue de construire les ligands d'intérêts.

En ce qui concerne les chimiothèques, il en existe deux grands types : les chimiothèques réelles et les chimiothèques virtuelles.

Les chimiothèques réelles sont souvent sous forme de plaques de puits contenant chacun un produit différent. Ces plaques sont donc prêtes à être testées.

Les chimiothèques virtuelles représentent un ensemble d'informations plus ou moins organisées et hiérarchisées, regroupant des données concernant un ensemble de composés pouvant aller jusqu'à 2 millions. Dans ce type de chimiothèques, les coordonnées de chaque composé sont enregistrées dans des fichiers avec les formats SMI, SDF, MOL2 et PDB. Parmi les chimiothèques virtuelles nous citons : La chimiothèque Nationale Française, la chimiothèque de l'institut Curie, ZINC, PubChem...etc.

6.3 Programmes

Au cours des deux dernières décennies, une grande variété de plus de 60 programmes de docking différents ont été proposés (DOCK, AutoDock, FlexX, Surflex, OR, ICM, Glide, Cdocker, LigandFit, MCDock, et plein d'autres). Bien que ces programmes reposent le plus souvent sur des algorithmes spécifiques, leur protocole est composé de 2 étapes essentielles :

La 1ère étape : dite de docking proprement dit est l'étape de sélection, consistant à placer le ligand dans le site actif de la protéine et à échantillonner les conformations, positions et orientations (poses) possibles, en ne retenant que celle qui représente les modes d'interactions les plus favorables.

La 2ème étape : le scoring est l'étape de classement, qui consiste à évaluer l'affinité entre le ligand et la protéine, et de donner un score aux poses obtenues lors de la phase de docking. Ce score permettra de retenir la meilleure pose parmi toutes celles proposées.

On regroupe les méthodes *in silico* en deux grandes familles, le criblage virtuel «structure-based» et le criblage virtuel «ligand-based».

➤ **Criblage virtuel «ligand-based» :**

Les méthodes de criblage «ligand-based» reposent sur la connaissance préalable de ligands ayant une activité sur la cible thérapeutique. Il sera ainsi possible d'utiliser ces ligands comme une première base de «hits» afin d'identifier d'autres composés similaires, présentant des caractéristiques d'activité communs aux ligands connus de la cible. Différents types de descripteurs moléculaires pourront être calculés pour quantifier la similitude entre composés. Suivant le nombre de ligands connus de la cible thérapeutique, plusieurs méthodes peuvent être employées: la recherche de similarité, le criblage de pharmacophores, ou les approches QSAR.

➤ **Criblage virtuel «structure-based» :**

Le criblage virtuel « structure-based » est considéré comme un équivalent *in silico* d'un test expérimental étudiant la liaison ligand-cible biomoléculaire. Cependant, ce criblage dépend essentiellement de la disponibilité de la structure 3D de la cible biologique qui est obtenue, soit par méthodes expérimentales (RX et RMN) dans des bases de données (tel que PDB: Protein Data base), soit par des méthodes de prédiction de la structure 3D par homologie de séquence. Différentes approches peuvent être employées pour réaliser ce criblage: la construction de modèles pharmacophoriques 3D, l'établissement de modèles 3D-QSAR), la conception de novo « de novo design » et les méthodes de docking moléculaire qui sont les plus populaires.

7. Les différentes étapes du criblage virtuel

Le criblage virtuel est largement utilisé pour identifier de nouvelles substances bio-actives. Cette approche est souvent composée de plusieurs étapes essentielles :

a- Préparation de la cible

La préparation de la structure 3D d'une cible thérapeutique commence par choisir la meilleure structure cristallographique disponible dans la PDB. Ce choix s'effectue selon la valeur de la résolution définissant la qualité de la structure. Il est à souligner que les structures 3D des protéines provenant de la PDB nécessitent une préparation préalable avant leur utilisation dans des expériences de docking moléculaire. En effet, une seule chaîne doit être gardée dans le cas des protéines contenant plusieurs sous-unités identiques et ce pour une utilisation plus aisée lors du docking. L'inhibiteur de référence et d'autres éléments appelés hétéroatomes (alpha-l-fucose, le 1,2-Ethandiol et le N-Acetyl-d-glucosamine) ayant servi à la cristallisation doivent également être supprimés pour ne conserver que la structure seule de la chaîne A de la protéine. Par la suite, les hydrogènes manquants doivent être rajoutés tandis que toutes les molécules d'eau doivent être supprimées sauf ceux faisant partie de la cavité catalytique de l'enzyme. Les états de protonation de chaque résidu dans la cavité de l'enzyme doivent être vérifiés et l'énergie d'interaction intramoléculaire de la protéine doit être minimisée.

b- Préparation de la chimiothèque

Dans certains cas, il est préférable de procéder à un nettoyage de la chimiothèque avant de lancer les calculs de criblage virtuel. Cette étape consiste à ne garder que les composés présentant une structure adéquate avec un bon potentiel de candidat médicament. Pour ce faire, il est nécessaire d'appliquer, sur la chimiothèque en question, des filtres dits « ADMET ». Les composés de la chimiothèque nettoyé (ou pas) doivent être préparés en générant les différentes formes ionisées et les tautomères car dans les conditions physiologiques, une molécule peut prendre différentes formes en subissant des phénomènes de tautomérie ou d'ionisation.

c- Criblage virtuel proprement dit

Le criblage virtuel doit être effectué en utilisant un ou plusieurs logiciels. Malgré l'existence d'un grand nombre de programmes de docking, aucun ne peut être appliquée de façon générale sur la totalité des systèmes biologiques car chacun présente des avantages mais également des imperfections. Ainsi, l'utilisation de plusieurs programmes de docking suivie par l'application d'une méthode consensus est recommandée pour combiner et tirer le meilleur profit des

logiciels de docking utilisés dans le cadre d'une même étude. Autrement dit, cette approche permet de combiner les informations issues des différentes fonctions de score afin de compenser leurs imperfections individuelles ce qui améliore la qualité des résultats obtenus. L'hypothèse sous-jacente est que la probabilité qu'une molécule soit active doit augmenter si cette dernière est associée à de bons scores d'affinité selon plusieurs fonctions de score. De même, la probabilité qu'une molécule fautive/positive soit bien classée doit diminuer car, avec leurs fonctions de score différentes, les logiciels ont peu de chance de commettre la même erreur.

d- Analyse visuelle et sélection des composés à tester expérimentalement

Bien que l'application d'un protocole de criblage virtuel permet de supprimer automatiquement un grand nombre de composés indésirables, on ne peut pas substituer l'intervention humaine. Un biais bien connu, est par exemple le bon classement des ligands de haute masse moléculaire, qui en raison d'un nombre d'atomes plus important, sont tout simplement capables de créer un plus grand nombre d'interactions même-ci au niveau l'entrée voir à l'extérieur de la cavité étudiée. Désormais, l'inspection visuelle du mode d'interaction des composés ayant obtenus les meilleurs scores est une étape primordiale afin de faire ressortir une liste réduite de molécules prometteuses à tester expérimentalement. Lors de cette inspection, plusieurs critères de sélection doivent être pris en considération :

-Positionnement des composés dans la cavité étudiée

-Le mode d'interaction (nombre optimal de liaisons H, interactions ioniques, hydrophobes, Pi-Pi...etc.).

-Interaction avec les résidus clés de la cible.

- Diversité structurale

-Disponibilité des molécules réelles dans la chimiothèque.