

Cour 08 : Modélisation de structures de protéines par docking Moléculaire

1. Le docking moléculaire

Le docking moléculaire (ancrage, amarrage ou arrimage moléculaire en français) est une approche in silico consistant à prédire et à simuler la position la plus favorable d'un ligand au sein de son récepteur (cible souvent protéique). A l'heure actuelle, il existe plus que 40 programmes (logiciels) de docking moléculaire. Bien que ces programmes reposent le plus souvent sur des algorithmes spécifiques, leur protocole est composé de 2 étapes complémentaires :

La première étape dite docking permet au ligand d'adopter plusieurs conformations et plusieurs positions au niveau de son récepteur afin de retenir celle la plus favorable.

La deuxième étape, dite scoring, permet d'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.

2. Applications

Il y a au moins deux applications principales au programme de docking. La plus ancienne est la prédiction du mode d'interaction qui consiste à déterminer le positionnement correct du ligand par rapport à son récepteur.

La deuxième utilisation réside à rechercher par criblage virtuel de nouvelles molécules actives envers une cible thérapeutique donnée.

Le criblage virtuel in silico est une approche permettant de simuler et prédire le mode d'interaction d'un très grand nombre de molécules (collectées en chimiothèques) vis-à-vis un site actif ciblé. Autrement dit, le criblage virtuel est un docking moléculaire de plusieurs milliers de molécules envers un seul récepteur.

3. Type de docking moléculaires

- **Docking rigide**

Dans le cas des méthodes de docking rigide, la recherche de la pose optimale se limite au positionnement. Cette opération consiste en la recherche exhaustive dans l'espace discrétisé des 6 degrés de liberté. Certains programmes, s'ils n'appartiennent pas à la famille des techniques de docking rigide, utilisent plusieurs étapes successives d'optimisation dont les premières peuvent s'apparenter à du docking rigide. Par exemple, le programme Glide, utilise initialement, dans son approche multi-étapes, une recherche systématique pour positionner le ligand de façon approchée au sein du site actif de la protéine.

- **Docking flexible**

Lorsque les méthodes de docking prennent en compte la flexibilité du ligand, deux étapes sont effectuées successivement pendant toute la durée du docking. La première étape correspond à une exploration de l'espace conformationnel de manière à retrouver, parmi les conformations proposées, la conformation bioactive. Pendant la deuxième étape une fonction de score évalue ces conformations. Il existe plusieurs types d'algorithmes pour le traitement de la flexibilité du ligand : les méthodes systématiques (fragmentation/reconstruction), les méthodes aléatoires, et les méthodes de simulation (dynamique moléculaire).

- **Docking semi-flexible**

Lorsque l'espace conformationnel des ligands est exploré, le nombre de degrés de liberté de l'espace de recherche peut être conséquent dans le cas de molécules très flexibles. Dans un tel contexte, l'emploi de méthodes de recherche exhaustives apparaît souvent inapproprié car nécessitant des simplifications importantes au niveau de l'échantillonnage. D'autres algorithmes, dits de fragmentation, sont employés pour construire de façon incrémentielle le ligand au sein du site actif de la protéine. L'espace des conformations du ligand est alors restreint au voisinage d'un ensemble initial d'états simplifiés. Cette stratégie de recherche par construction, qui se présente sous diverses variantes, est notamment adoptée par les programmes DOCK, FLEXX, et Hammerhead.

4. Les outils du docking moléculaires

➤ Le récepteur

La première grande voie d'étude et de conception de molécules bioactives par modélisation moléculaire est celle qui se fonde sur la structure des récepteurs. Cette approche est basée sur l'exploitation de la structure moléculaire tridimensionnelle de la protéine cible. Trois méthodes expérimentales permettent aujourd'hui de déterminer la structure des protéines : la résonance magnétique nucléaire (RMN), la microscopie électronique et la cristallographie par rayons X. Cette dernière technique est responsable de la majorité des structures issues d'une base de données de structures appelée la Protéine Data Bank (PDB).

La PDB est un répertoire mondial de dépôt d'informations sur la structure tridimensionnelle des protéines et des acides nucléiques. Ces molécules proviennent de l'ensemble des règnes biologiques. La PDB est gratuitement accessible par Internet (<http://www.rcsb.org/pdb/>). Elle contient plusieurs de milliers de structures protéiques obtenues soit par cristallographie aux rayons X, soit par RMN. Si la cible n'est pas encore déposée au niveau de la banque, et cette dernière contient une protéine avec des séquences similaires, la modélisation par homologie intervient afin de construire la structure 3D de la cible souhaitée.

➤ Le ligand

En docking moléculaire, le choix du ligand est une étape très importante. Ce choix doit être pertinent en raison de la spécificité du site actif de la cible évitant de tester inutilement des molécules.

Pour un docking moléculaire, le ligand doit être également sous forme 3D. A présent, il existe deux moyens pour obtenir la structure chimique d'un ligand donné : La première souvent d'aspect commercial, est constituée de bases de données de structures chimiques appelés chimiothèques ou espaces chimique .Le second moyen consiste à utiliser des ligands de la PDB ou 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.

➤ Un programme de docking moléculaire

Le docking moléculaire s'accomplit en deux étapes complémentaires. La première consiste à rechercher les conformations du ligand aptes à établir des interactions idéales avec le récepteur.

La deuxième est une fonction de score qui permet d'évaluer ces conformations par un calcul rapide de leur énergie d'interaction avec ce récepteur.

5. Filtrage par ADME –TOX

L'abréviation ADMET représente les concepts pharmacocinétiques suivants : l'absorption, la distribution, la métabolisation et l'excrétion. Ce sont des critères qui décrivent la disposition d'une molécule bioactive dans un organisme. Ils expriment les concentrations du produit dans les différents tissus et système circulatoire de l'organisme. Les résultats obtenus par ces critères permettent d'analyser la performance et l'efficacité d'un produit afin de juger sa capacité à devenir un médicament et son intérêt à poursuivre son développement pour des études cliniques.

Les filtres de type ADME-Tox (Absorption, Distribution, Métabolisme, Elimination et Toxicité) sont rapidement devenus populaires. Ils reposent sur plusieurs critères déterminant les propriétés pharmacocinétiques potentielles des molécules et sont désormais largement utilisés pour réduire le nombre de composés d'une chimiothèque en sélectionnant les plus adaptés à devenir des candidats médicaments, avant tout processus de criblage. L'utilisation de ces filtres a donné de très bons résultats.

Absorption : Cela consiste à analyser la faculté d'une molécule à pénétrer au sein de l'organisme après administration. Une faible solubilité (par exemple due à une trop grande hydrophobicité) ou une forte polarité ont un impact drastique sur l'absorption intestinale d'un composé.

Distribution : Ce critère mesure la capacité d'une molécule à diffuser, par exemple via le flux sanguin, à travers l'organisme. En effet, une molécule doit pouvoir passer d'un compartiment à un autre, afin de pouvoir arriver in fine à l'endroit où sa cible doit être atteinte. La forte liaison à des protéines plasmatiques a un impact négatif sur la distribution d'une molécule.

Métabolisme : Le filtre métabolisme vise à détecter (a) la stabilité de la molécule dans l'organisme qui impacte sur son temps d'action et (b) les métabolites de la molécule initiale, à savoir les composés résultant de sa dégradation ou de modifications enzymatiques ayant lieu au sein de l'organisme. Chez l'homme, les cytochromes P450 du foie sont les principales enzymes modifiant les xénobiotiques.

Ces derniers sont notamment rendus plus hydrophiles par l'introduction d'atomes d'oxygène. Les métabolites peuvent être inactifs, plus actifs que le composé original et bien entendu potentiellement toxiques, d'où la nécessité de les caractériser et de les étudier.

Excrétion (élimination) : Afin d'éviter les phénomènes d'accumulation, souvent synonyme de toxicité, il faut veiller à ce que les composés administrés, ainsi que leurs métabolites, soient bien excrétés de l'organisme, par exemple via l'urine ou les selles.

Toxicité : Comme son nom l'indique, ce filtre sert à mesurer la toxicité d'un composé et de ses métabolites. Désormais, la toxicité et le manque d'efficacité des candidats médicaments sont les deux plus grandes causes d'échecs dans le développement d'un médicament. Différents types de toxicité sont évalués, entre autre la cancérogénicité.

Un contributeur majeur dans le domaine permettant d'identifier rapidement et à grande échelle des molécules à caractère « drug-like » est communément appelées « règles de Lipinski » ou « la règle de 5 » permettant d'estimer la biodisponibilité d'un composé par voie orale à partir de sa structure bidimensionnelle (2D). Ces règles concernant les propriétés physico-chimiques ont été définies après l'analyse de 2245 médicaments commercialisés ou en phase finale de développement.

- Le poids moléculaire du composé ne doit pas être supérieur à 500 daltons (Da).
- Le logarithme décimal du coefficient de partage eau / 1-octanol, noté logP, doit être inférieur à 5.
- Le nombre de donneurs de liaisons hydrogène doit être inférieur à 5.
- Le nombre d'accepteurs de liaisons hydrogène doit être inférieur à 10.

D'autres critères ont été mis en place pour compléter et ajuster les règles de Lipinski dans la sélection de composés « drug-like ».

Ainsi, Veber Choisissent d'utiliser critères suivants la surface polaire (PSA, polar surface area) du composé doit être inférieur à 140 Å² et le nombre de liaisons de rotation ("rotatable bonds" en anglais) doit être inférieur à 10 Sont souvent employés en complément de la « règle des 5 » Ces critères ont été établis par l'étude de la biodisponibilité orale candidats médicaments.