

## Criblage virtuel par docking moléculaire

---

*Didier Rognan*

### *I. Introduction*

Les nombres grandissants de cibles génomiques d'intérêt thérapeutique ([Hopkins et Groom, 2002](#)) et de macromolécules (protéines, acides nucléiques) pour lesquelles une structure tridimensionnelle (3-D) est disponible ([Berman et al., 2000](#)) rendent les techniques de criblage virtuel de plus en plus attractives pour des projets d'identification de molécules bio-actives ([Walters et al., 1998; Lengauer et al., 2004](#)). Par **criblage virtuel**, on entend tout **procédé de recherche électronique dans des banques de données moléculaires permettant la sélection de molécules**. La requête peut être effectuée sous différents types de contraintes (descripteurs physicochimiques, pharmacophore, topologie d'un site actif) et doit aboutir à la sélection d'un faible pourcentage (1-2%) de molécules présentes dans la **chimiothèque (banque de données de ligands)** de départ. Nous allons ici aborder les diverses stratégies intégrées susceptibles d'aboutir à un criblage virtuel réussi à partir de la structure 3-D de la protéine cible.

### *II. Les 3 étapes d'un criblage virtuel*

Tout criblage virtuel se décompose en trois étapes d'égale importance:

- (1) la mise au point de la chimiothèque de départ,
- (2) le criblage proprement dit,
- (3) la sélection d'une liste de touches virtuelles.

Il est à noter que toute erreur à chacune de ces trois étapes aura des conséquences importantes se traduisant généralement par une augmentation du taux de faux positifs et de faux négatifs. Il convient donc d'être très attentif à chacune d'entre elles.

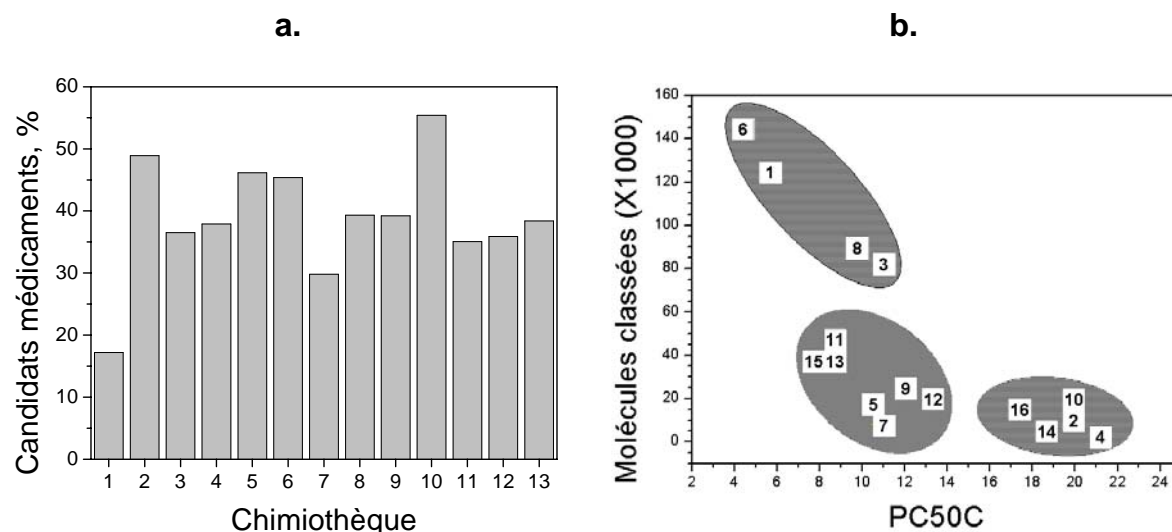
## a. Préparation de la chimiothèque

- **Choix de la chimiothèque.** Deux types de chimiothèques peuvent être utilisés à des fins de criblage virtuel : des collections de criblages (molécules disponibles) et des collections virtuelles (molécules à synthétiser). Toute compagnie pharmaceutique dispose maintenant sous forme physique (microplaques) et électronique de collections de criblage propriétaires pouvant aller jusqu'à 2 millions de molécules. Par ailleurs, diverses collections de criblage sont disponibles commercialement (Baurin et al., 2004) et constituent une source majeure de molécules bioactives pour le monde académique notamment.

Dans ce chapitre nous avons comparé des collections de molécules commerciales et publiques, en tentant de présenter un horizon le plus complet. Pour avoir un aperçu des collections de criblage commercialement disponibles, le lecteur peut se référer à une liste assez exhaustive décrite sur le site <http://www.warr.com/links.html#chemlib>. Le lecteur est aussi invité à consulter les documentations relatives à chacune de ces chimiothèques pour un descriptif de leur valorisation et de leur originalité.

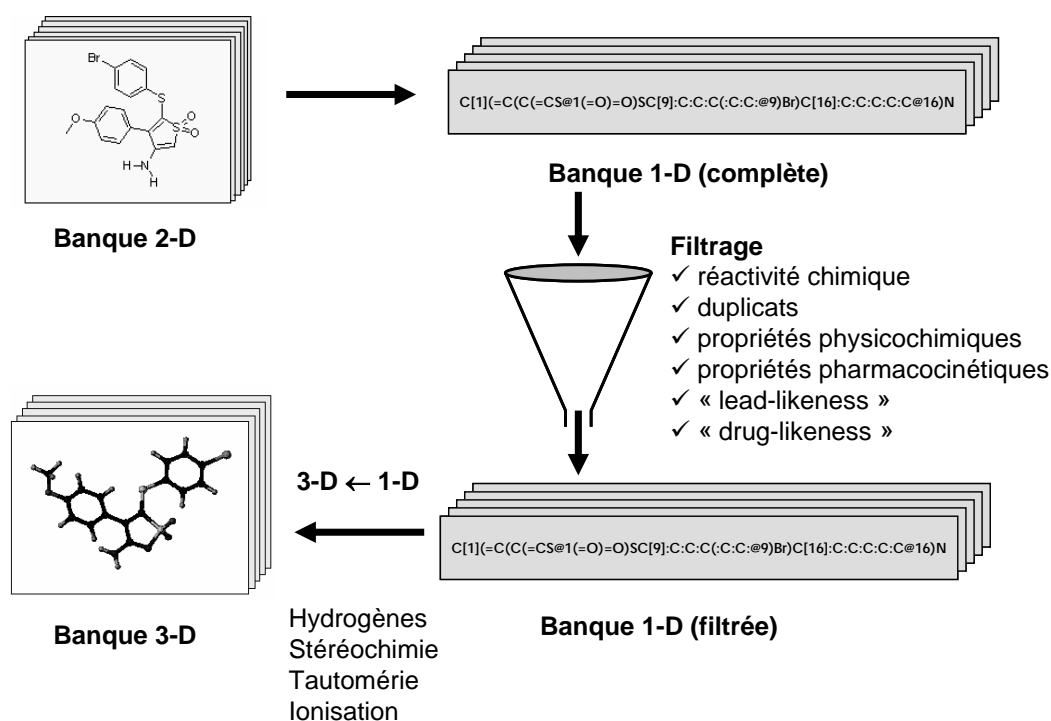
De façon générale, l'analyse de ces collections montre que le pourcentage de candidats médicaments pour chacune de ces sources de criblage reste modéré (Charifson et Walters, 2002) (Figure 16.1a).

Qu'en est-il de leur diversité moléculaire ? Elle est également assez faible mais très dépendante de leur origine. Les collections issues de chimie combinatoire sont vastes mais peu diverses en châssis moléculaires (Figure 16.1b). Des collections de moindre taille présentent également une faible diversité. Seul un petit nombre de collections (telles que la Chimiothèque Nationale ; voir Chapitre 2) présente un plus grand rapport taille/diversité. On remarquera que ce sont celles qui se rapprochent également le plus de la collection regroupant des candidats médicaments avérés (Sheridan et Shpungin, 2004). Le choix de la chimiothèque est donc crucial et dépendra du projet. Plutôt que de sélectionner une seule collection, il convient plutôt de choisir parmi les sources possibles les molécules les plus diverses en termes de châssis moléculaire et en évitant autant que possible les redondances.



**Fig. 16.1** Analyse de collections de criblage. **a)** Potentiel en candidats médicaments de collections de criblage commerciales (1 à 7 et 9 à 13) et publique (8 : Chimiothèque Nationale) **b)** Diversité en châssis moléculaire: Pourcentage de châssis couvrant 50% des candidats médicaments (métrique PC50C) de la chimiothèque (1 à 3 et 5 à 16, chimiothèques commerciales ; 4, Chimiothèque Nationale)

- **Filtrage et préparation finale de la chimiothèque.** Afin de ne sélectionner que des molécules d'intérêt, il est commun de filtrer la chimiothèque par un certain nombre de descripteurs (**Figure 16.2**) pour ne retenir que des candidats médicaments potentiels (**Charifson et Walters, 2002**).



**Fig. 16.2** Différentes phases de la préparation de la chimiothèque. Les termes « lead-likeness » et « drug-likeness » désignent les propriétés admises généralement (selon des modèles statistiques) pour être des têtes de séries (lead) ou des médicaments (drug).

Ces filtres sont destinés à éliminer les molécules chimiquement réactives, toxiques, métabolisables, présentant des propriétés physicochimiques inadéquates (ex: non conformes aux règles de Lipinski, voir **Chapitre 8**) susceptibles d'être des touches dites permissives, ou pouvant interférer avec le test de criblage expérimental (molécules fluorescentes, voir **Chapitre 3**). En tout état de cause, il est important d'adapter l'intensité du filtrage au cahier des charges. Ce filtrage sera intense si on veut identifier des touches pour une cible déjà très étudiée par le passé et pour laquelle de nombreux ligands existent. Il sera plus lâche si le but du criblage virtuel est de proposer les premiers ligands putatifs d'une cible orpheline. La dernière étape de préparation est donc de convertir ce format 1-D en format 3-D en y incluant une représentation atomistique complète. Nous proposons ci-dessous des liens vers divers outils de manipulation des chimiothèques (**Exemple 16.1**)

#### Exemple 16.1 : Principaux outils de conception / gestion de chimiothèques électroniques

Nom	Editeur	Site Internet	Fonction
ISIS/base	MDL	<a href="http://www.mdli.com">http://www.mdli.com</a>	archivage
ChemOffice	Cambridge Soft.	<a href="http://www.cambridgesoft.com">http://www.cambridgesoft.com</a>	archivage
Filter	Openeyes	<a href="http://www.eyesopen.com">http://www.eyesopen.com</a>	filtrage
Cliff	Molecular Network	<a href="http://www.mol-net.de">http://www.mol-net.de</a>	filtrage
Pipeline Pilot	SciTegic	<a href="http://www.scitegic.com">http://www.scitegic.com</a>	filtrage et automatisation de procédures
Marvin	Chemaxon	<a href="http://www.chemaxon.com">http://www.chemaxon.com</a>	archivage
Ligprep	Schrodinger	<a href="http://www.schrodinger.com">http://www.schrodinger.com</a>	filtrage

### b. Criblage par docking à haut débit

Le **docking** (ou **ancrage**) à **haut débit** consiste à prédire à la fois la conformation active et l'orientation relative de chacune des molécules de la chimiothèque sélectionnée par rapport à la cible d'intérêt. Très généralement, la recherche se focalise sur le **site actif** tel qu'il aura pu être déterminé expérimentalement (par mutagenèse dirigée par exemple). Il est important de prendre en compte à ce stade la notion de débit que l'on veut atteindre. Il faudra en effet considérer le meilleur compromis entre rapidité et précision. En général, le docking à haut débit nécessite une rapidité voisine de 1-2 minute/ligand. De nombreux programmes de docking sont disponibles ([Taylor et al., 2002](#)) (**Exemple 16.2**).

#### Exemple 16.2 : Principaux programmes de docking moléculaire

Nom	Editeur	Site Internet
AutoDock	Scripps	<a href="http://www.scripps.edu/mb/olson/doc/autodock/">http://www.scripps.edu/mb/olson/doc/autodock/</a>
Dock	UCSF	<a href="http://dock.compbio.ucsf.edu/">http://dock.compbio.ucsf.edu/</a>
FlexX	BioSolveIT	<a href="http://www.biosolveit.de/FlexX/">http://www.biosolveit.de/FlexX/</a>
Fred	OpenEyes	<a href="http://www.eyesopen.com/products/applications/fred.html">http://www.eyesopen.com/products/applications/fred.html</a>
Glide	Schrödinger	<a href="http://www.schrodinger.com/Products/glide.html">http://www.schrodinger.com/Products/glide.html</a>
Gold	CCDC	<a href="http://www.ccdc.cam.ac.uk/products/life_sciences/gold/">http://www.ccdc.cam.ac.uk/products/life_sciences/gold/</a>
ICM	Molsoft	<a href="http://www.molsoft.com/products.html">http://www.molsoft.com/products.html</a>
LigandFit	Accelrys	<a href="http://www.accelrys.com/cerius2/c2ligandfit.html">http://www.accelrys.com/cerius2/c2ligandfit.html</a>
Surflex	Biopharmics	<a href="http://www.biopharmics.com/products.html">http://www.biopharmics.com/products.html</a>

Ces méthodes utilisent toutes le principe de complémentarité stérique (Dock, Fred) ou d'interactions moléculaires (AutoDock, FlexX, Glide, Gold, ICM, LigandFit, Surflex) afin de placer un ligand dans le site actif d'une cible. Généralement, la protéine est considérée comme rigide alors que la flexibilité du ligand est prise en compte relativement bien jusqu'à une quinzaine de rotules.

Trois principes sont généralement utilisés dans le traitement de la flexibilité du ligand :

- (1) un ensemble de conformations du ligand est préalablement calculé et celles-ci sont dockées de manière rigide dans le site (ex: Fred),
- (2) le ligand est construit de manière incrémentale fragment après fragment (ex: Dock, FlexX, Glide, Surflex),
- (3) une analyse conformationnelle plus au moins exhaustive est conduite sur le ligand de manière à générer les conformations les plus propices au docking (ex: ICM, Gold, LigandFit).

Généralement, plusieurs poses du ligand sont générées et classées par ordre de probabilité décroissante selon une fonction d'évaluation (Gohlke et Klebe, 2001) qui au mieux essaiera d'approximer l'énergie libre de liaison du ligand à sa cible (c'est-à-dire son affinité) ou plus simplement classera les molécules de la chimiothèque par énergie d'interaction avec le site de la protéine cible. La précision d'outils de docking moléculaire peut être évaluée en comparant prédictions et solutions expérimentales sur un ensemble représentatif de complexes protéines-ligand issus de la **Protein Data Bank** (<http://www.rcsb.org/pdb/>). L'expérience montre que si l'on considère les 30 solutions les plus probables, un ligand est généralement bien ancré dans environ 75% des cas (Kellenberger et al., 2004). La « meilleure solution » (la plus proche de la solution expérimentale) n'est pas toujours celle prédite comme la plus probable par la fonction d'évaluation (seulement dans 50 % des cas environ) ce qui complique singulièrement l'analyse prédictive de solutions de docking (Kellenberger et al., 2004). De nombreuses raisons existent pour expliquer ces imperfections, dont le traitement peut s'avérer plus ou moins compliqué (Exemple 16.3).

#### Exemple 16.3 : Principales causes d'erreurs lors d'un docking moléculaire

Causes d'erreurs	Traitement
Site actif dénué de cavité	impossible ?
Flexibilité de la protéine	très difficile
Influence de l'eau	très difficile
Imprécision des fonctions d'évaluations	très difficile
Interactions non usuelles	difficile
Flexibilité du ligand	difficile
Pseudosymétrie du ligand	difficile
Mauvais jeux de coordonnées (protéine)	facile
Mauvais types atomiques (ligand, protéine)	facile

C'est la raison pour laquelle le docking moléculaire reste une technologie difficile à mettre en œuvre car elle doit être appliquée et constamment adaptée en fonction du contexte dans lequel le projet se place. Il est néanmoins possible de fournir des guides en fonction du type de protéine, site actif et ligand(s) à ancrer (Kellenberger et al., 2004).

Appliqué au criblage virtuel de chimiothèques, le docking moléculaire devra non seulement fournir des poses les plus exactes possibles pour chaque ligand de la chimiothèque, mais aussi être capable de classer les ligands par ordre décroissant d'affinité prédite de façon à sélectionner les touches d'intérêt. Ceci reste un des défis majeurs de la chimie théorique, car il convient de prédire avec une précision et une rapidité compatibles avec le débit du criblage, les composantes enthalpiques (une tâche facile) et entropiques (une tâche beaucoup plus difficile) de l'énergie libre de liaison de chaque ligand (Gohlke et Klebe, 2001). De nombreuses méthodes de prédiction d'énergie libre existent. Leur précision est toutefois liée au débit avec lequel elles peuvent être appliquées (Figure 16.3). Des méthodes

thermodynamiques permettent une précision relativement élevée mais ne seront applicables qu'à une paire de ligands (prédiction de différences d'énergie libre). A l'opposé, des fonctions empiriques pourront être utilisées à plus haut débit (> 100,000 molécules) mais avec une précision très moyenne (de l'ordre de 7 kJ/mol soit une unité et demie de pK en termes d'affinité)

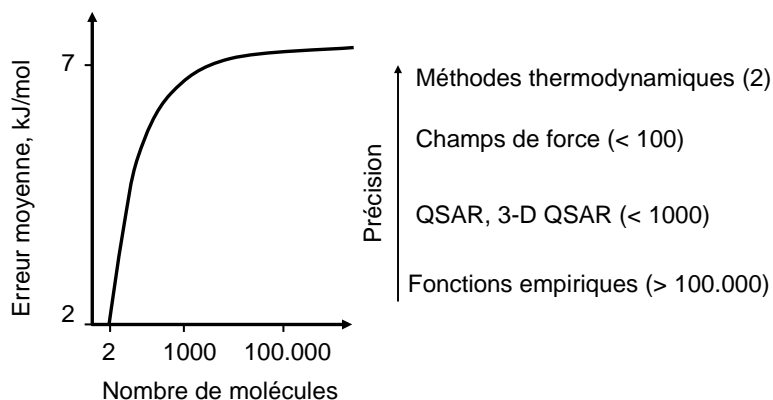
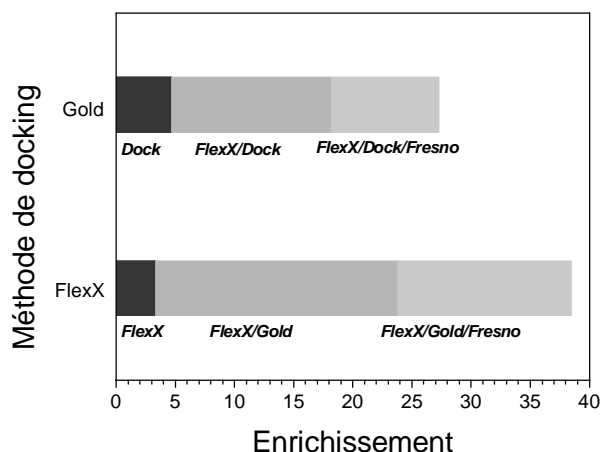


Fig. 16.3 Méthodes de prédiction d'énergie libre (affinité) de liaison d'un ligand

De nombreuses études montrent qu'il est impossible de prédire avec précision l'affinité de ligands chimiquement divers (Ferarra et al., 2004). **Il est raisonnable d'espérer discriminer ligands d'affinité nanomolaire, micromolaire et millimolaire**, ce qui est probablement suffisant pour identifier des touches dans une chimiothèque mais insuffisant pour les optimiser. A partir du moment où la sélection des touches se fait sur la base de scores de docking quels qu'ils soient, **le criblage virtuel par docking moléculaire va donc inmanquablement fournir de nombreux faux positifs et surtout de nombreux faux négatifs**, ce qui le différencie très distinctement du criblage expérimental à haut débit qui est plus exhaustif pour l'identification des vrais positifs.

### c. Post-traitement des données

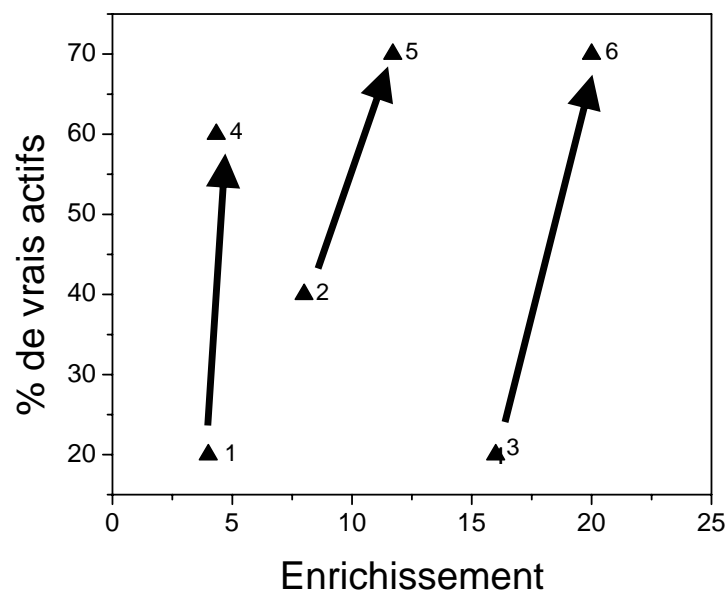
Acceptant le fait que les fonctions d'évaluation sont très imparfaites, la meilleure stratégie pour augmenter le taux de vrais positifs lors d'un criblage virtuel consiste à essayer de détecter les faux positifs. Ceci n'est possible qu'à condition d'analyser les sorties de criblage avec une méthode chémoinformatique additionnelle. Plusieurs solutions sont possibles. La plus simple consiste à scorer de nouveau les poses obtenues à l'aide de fonctions d'évaluations différentes de celle utilisée pendant le docking. Chaque fonction ayant des imperfections, l'analyse consensuelle (Charifson et al., 1999) permet de détecter les faux positifs comme les touches non communes à deux ou trois fonctions reposant sur des principes physicochimiques différents (Figure 16.4).



**Fig. 16.4** Influence de procédure d'évaluation consensus sur l'enrichissement en vrais actifs, par rapport à un criblage au hasard (une seule fonction: barre noire, deux fonctions: barre gris foncé, trois fonctions: barre gris clair). Les fonctions d'évaluations utilisées sont mentionnées en italiques (Bissantz et al., 2000).

La sélection de touches scorées parmi les premiers 5% de fonctions différentes permet d'enrichir la sélection finale en vrais positifs (Charifson et al., 1999; Bissantz et al., 2000). Cette méthode présente l'avantage d'ajuster une stratégie de criblage en fonction de données expérimentales connues. Il suffit de préparer une chimiothèque test ou un petit nombre de vrais actifs (de l'ordre d'une dizaine par exemple) et de la mélanger à un grand nombre de molécules supposées inactives (un millier par exemple), d'ancrer la chimiothèque avec divers outils de docking, et de rescorer les poses obtenues avec différentes fonctions de scoring. L'analyse systématique des enrichissements en vrais positifs se fait en calculant le nombre de vrais actifs dans diverses listes de sélection déterminées par scoring simple ou multiple. La stratégie de criblage (docking/scoring) donnant le meilleur enrichissement peut être ensuite appliquée au criblage en grandeur nature. Malgré ces avantages, cette technique ne peut être appliquée en absence de données expérimentales (connaissance de plusieurs vrais actifs chimiquement divers). Il faut dans ce cas mettre en œuvre des stratégies plus générales pour éliminer les faux positifs: détection de ligands insuffisamment enfouis (Stahl et Böhm, 1998) ; raffinement par minimisation énergétique des poses de docking (Taylor et al. 2003); docking consensus par divers outils (Paul et Rognan, 2002) ; docking sur des conformations multiples de la cible (Vigers et Rizzi, 2004) ; rescoring de poses multiples (Kontoyianni et al., 2004). Pour la plupart, ces approches sont assez compliquées à mettre en œuvre et ne garantissent pas une application large à divers projets de criblage.

Une stratégie de post-traitement plus simple mais efficace consiste à appliquer un traitement statistique aux molécules de la chimiothèque regroupées de façon « quasi phylogénétique » par châssis moléculaire (Nicolaou et al., 2002). Plutôt que de s'intéresser à des scores individuels, il suffit alors de regarder la distribution de ceux-ci au sein de classes chimiques homogènes. Ceci **permet de privilégier non plus des molécules mais des châssis moléculaires supposés suffisamment enrichis en touches virtuelles** (Figure 16.5) et ainsi d'identifier des faux négatifs (molécules actives mal ancrées et/ou scorées).



**Fig. 16.5** Influence de la stratégie d'analyse de données sur l'enrichissement en actifs, par rapport à un criblage au hasard, à partir d'un même jeu de données de docking (10 antagonistes du récepteur V1a de la vasopressine « enfouis » dans une base de 1.000 molécules; Bissantz et al., 2003) **1**: Sélection des 5 premiers pourcents ancrés et scorés par Gold; **2**: Sélection de 5 premiers pourcents ancrés et scorés par FlexX; **3**: Sélection des touches communes aux listes 1 et 2; **4**: sélection de châssis dont 60% des représentants présentent un score Gold supérieur à 37,5; **5**: sélection de châssis dont 60% des représentants présentent un score FlexX inférieur à -22; **6**: sélection de châssis dont 60% des représentants présentent un score Gold supérieur à 37,5 et inférieur à -22. Les châssis moléculaires ont été calculés par le logiciel ClassPharmer (Bioresaon, Santa Fe, USA). Les flèches indiquent le gain enregistré dans la sélection de vrais actifs par analyse des châssis moléculaires (singletons exclus)

En tout état de cause, la sélection finale des molécules à commander pour évaluation expérimentale passe par un examen individuel des interactions 3-D de chaque touche virtuelle avec le récepteur ainsi qu'une étude de disponibilité des molécules chez leurs fournisseurs respectifs si une collection commerciale a été criblée. Selon le laps de temps séparant la téléchargement du catalogue électronique de la chimiothèque et la commande, le pourcentage de molécules devenues indisponibles augmente significativement (environ 25% au bout de 3 mois). L'actualisation de ces chimiothèques commerciales est donc absolument nécessaire afin de garantir la disponibilité du maximum des ligands choisis par criblage virtuel.

### *III. Quelques succès du criblage virtuel par docking*

De nombreux exemples de criblages virtuels réussis ont été décrits dans les trois dernières années (Exemple 16.4). A partir de structures cristallographiques à haute résolution de protéines ou d'acides nucléiques, il est généralement possible d'obtenir des taux de touches validées expérimentalement de l'ordre de 20-30%, ceci à partir de chimiothèques de taille et diversité variables, mais toujours préalablement filtrées comme indiqué plus haut.



#### Exemple 16.4 : Exemples de réussites de criblage virtuel de chimiothèques

Cible moléculaire	Chimiothèque	Taille	Taux de touche	Référence
Bcl-2	NCI	207K	20 %	Enyedy et al., 2001
HCA-II	Maybridge/Leadquest	90K	61 %	Grüneberg et al., 2001
Erα	ACD-Screen	1500K	72 %	Shapira et al., 2001
GAPDH	Comb. Lib.	2K	17 %	Bressi et al., 2001
PTP1B	Pharmacia	230K	35 %	Doman et al., 2002
β-Lactamase	ACD	230K	5 %	Powers et al., 2002
BCR-ABL	Chemdiv	200K	26 %	Peng et al., 2003
XIAP	Chinese Nat. Lib.	8K	14 %	Nikolovska et al., 2004
Aldose reductase	ACD	260K	55 %	Kraemer et al., 2004
Chk-1 kinase	AstraZeneca	550K	35 %	Lyne et al., 2004
Ribosomal A-site	Collection Vernalis	900K	26 %	Foloppe et al., 2004

Le criblage virtuel de modèles par homologie reste plus difficile à cause de l'incertitude générée par le modèle lui-même (Oshiro et al., 2004). Des progrès notables ont toutefois été enregistrés, notamment dans le domaines des récepteurs couplés aux protéines G, ou plusieurs études rétrospectives (Bissantz et al., 2003; Gouldson et al., 2004) et prospectives (Becker et al., 2004; Evers et Klebe, 2004) ont montré qu'il était possible d'enrichir de manière significative des listes de touches en vrais actifs (Exemple 16.5). Il convient toutefois de bien ajuster le modèle 3-D du récepteur en fonction du type de ligand recherché (agoniste, agoniste inverse, antagoniste neutre).

#### Exemple 16.5 : État de l'art des possibilité du criblage virtuel par docking.

##### Ce qui est possible :

- Cribler environ 50,000 molécules/jour
- Discriminer les vrais actifs de molécules choisies au hasard
- Obtenir des taux de touches de 10-30%
- Identifier environ 50% des vrais actifs
- Prévoir un profil de sélectivité pour différentes cibles

##### Ce qui reste difficile :

- Prédire l'orientation exacte du ligand
- Prédire l'affinité exacte du ligand
- Discriminer les vrais actifs d'inactifs chimiquement voisins
- Identifier 100% des vrais actifs
- Tenir compte de la flexibilité de la cible

## IV. Conclusion

Le criblage virtuel de chimiothèques par docking est devenue une méthode chémoinformatique utilisée en routine de manière à identifier des ligands pour des cibles d'intérêt thérapeutique. Il faut garder en mémoire que cette technologie est très sensible aux coordonnées 3-D de la cible et génère malgré tout de nombreux faux négatifs. Aussi importantes que le criblage lui-même sont donc les phases de préparation de la chimiothèque et la fouille des résultats de façon à détecter de faux positifs éventuels et ainsi améliorer le taux de touches qui peut aller jusqu'à 30% dans des cas favorables. Plutôt que de se focaliser sur le taux de touches, il est plus intéressant de considérer le nombre de nouveaux chémotypes dans les ligands identifiés et validés par criblage. De ce point de vue, cet outil est un complément naturel du chimiste médicinal afin de proposer des châssis moléculaires susceptibles de conduire rapidement à des chimiothèques focalisées à haute valeur ajoutée.

Des progrès à venir dans le domaine de la prédiction de propriétés **ADME/Tox** (pour **Absorption, Distribution, Métabolisme, Excrétion** et **Toxicité**) devraient permettre d'enrichir significativement le potentiel de cet outil chémoinformatique.

#### IV Références

- Baurin N., Baker R., Richardson C., Chen I., Foloppe N., Potter A., Jordan A., Roughley S., Parratt M., Greaney P., Morley D., Hubbard R.E. (2004) Drug-like annotation and duplicate analysis of a 23-supplier chemical database totalling 2.7 million compounds. *J. Chem. Inf. Comput. Sci.* 44:643-651
- Becker O.M., Marantz Y., Shacham S., Inbal B., Heifetz A., Kalid O., Bar-Haim S., Warshaviak D., Fichman M., Noiman S. (2004) G protein-coupled receptors: in silico drug discovery in 3D. *Proc. Natl. Acad. Sci. U.S.A.* 101:11304-11309
- Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.* 28:235-242
- Bissantz C., Folkers G., Rognan D. (2000) Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* 43:4759-4767
- Bissantz C., Bernard P., Hibert M., Rognan D. (2003) Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets ? *Proteins* 50:5-25.
- Bressi, J.C., Verlinde C.L., Aronov A.M., Shaw M.L., Shin S.S., Nguyen L.N., Suresh S., Buckner F.S., Van Voorhis W.C., Kuntz I.D., Hol W.G., Gelb M.H. (2001) Adenosine analogues as selective inhibitors of glyceraldehyde-3-phosphate dehydrogenase of Trypanosomatidae via structure-based drug design. *J. Med. Chem.* 44:2080-2093
- Charifson P.S., Corkery J.J., Murcko M.A., Walters W.P. (1999) Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. *J. Med. Chem.* 42: 5100-5109
- Charifson P.S., Walters W.P. (2002) Filtering databases and chemical libraries. *J. Comput. Aided Mol. Des.* 16:311-323
- Doman T.N., McGovern S.L., Witherbee B.J., Kasten T.P., Kurumbail R., Stallings W.C., Connolly D.T., Shoichet B.K. (2002) Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem.* 45:2213-2221
- Enyedy I.J., Ling Y., Nacro K., Tomita Y., Wu X., Cao Y., Guo R., Li B., Zhu X., Huang Y., Long Y.Q., Roller P.P., Yang D., Wang S. (2001) Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *J. Med. Chem.* 44:4313-4324
- Evers A., Klebe G. (2004) Successful virtual screening for a submicromolar antagonist of the neurokinin-1 receptor based on a ligand-supported homology model. *J. Med. Chem.* 47:5381-5392
- Ferrara P., Gohlke H., Price D.J., Klebe G., Brooks C.L. (2004) Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.* 47:3032-3047
- Foloppe N., Chen I.J., Davis B., Hold A., Morley D., Howes R. (2004) A structure-base strategy to identify new molecular scaffolds targeting the bacterial ribosome A-site. *Bioorg. Med. Chem.* 12:935-947
- Gohlke H., Klebe G. (2001) Statistical potentials and scoring functions applied to protein-ligand binding. *Curr. Opin. Struct. Biol.* 11:231-235
- Gouldson P.R., Kidley N.J., Bywater R.P., Psaroudakis G., Brooks H.D., Diaz C., Shire D., Reynolds C.A. (2004) Toward the active conformations of rhodopsin and the beta2-adrenergic receptor. *Proteins* 56:67-84
- Grüneberg S., Wendt B., Klebe G. (2001) Subnanomolar inhibitors from computer screening : a model study using human carbonic anhydrase II. *Angew Chem. Int. Ed. Engl.* 40:389-393
- Hopkins A.L., Groom C.R. (2002) The druggable genome. *Nat. Rev. Drug Discov.* 1:727-30
- Kellenberger E., Rodrigo J., Muller P., Rognan D. (2004) Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* 57:225-242

- Kontoyianni M., McClellan L.M., Sokol G.S. (2004) Evaluation of docking performance: comparative data on docking algorithms. *J. Med. Chem.* 47:558-565
- Kraemer O., Hazmann I., Podjarny A.D., Klebe G. (2004) Virtual screening for inhibitors of human aldose reductase. *Proteins* 55:814-823
- Lengauer T., Lemmen C., Rarey M., Zimmermann M. (2004) Novel technologies for virtual screening. *Drug Discov. Today* 9:27-34
- Lyne P.D., Kenny P.W., Cosgrove D.A., Deng C., Zabudoff S., Vendoloski J.J., Ashwell S. (2004) Identification of compounds with nanomolar binding affinity for checkpoint kinase 1 using knowledge-based virtual screening *J. Med. Chem.* 47:1962-68
- Nicolaou C.A., Tamura S.Y., Kelley B.P., Bassett S.I., Nutt R.F. (2002) Analysis of large screening data sets via adaptively grown phylogenetic-like trees. *J. Chem. Inf. Comput. Sci.* 42:1069-1079
- Nikolovska-Coleska Z., Xu L., Hu Z., Tomita Y., Li P., Roller P.P., Wang R., Fang X., Guo R., Zhang M., Lippman M.E., Yang D., Wang S. (2004) Discovery of embelin as a cell-permeable, small-molecular weight inhibitor of XIAP through structure-based computational screening of a traditional herbal medicine three-dimensional structure database. *J. Med. Chem.* 47:2430-2440
- Oshiro C., Bradley E.K., Eksterowicz J., Evensen E., Lamb M.L., Lancot J.K., Putta S., Stanton R., Grootenhuys P.D. (2004) Performance of 3D-database molecular docking studies into homology models. *J. Med. Chem.* 47:764-767
- Paul N., Rognan D. (2002) ConsDock: a new program for the consensus analysis of protein-ligand interactions. *Proteins* 47:521-533
- Peng H., Huang N., Qi J., Xie P., Xu C., Wang J., Wang C. (2003) Identification of novel inhibitors of BCR-ABL tyrosine kinase via virtual screening. *Bioorg. Med. Chem. Lett.* 13:3693-3699
- Powers R.A., Morandi F., Shoichet B.K. (2002) Structure-based discovery of a novel, noncovalent inhibitor of AmpC beta-lactamase. *Structure* 10:1013-1023
- Sheridan R.P., Shpungin J. (2004) Calculating similarities between biological activities in the MDL Drug Data Report database. *J. Chem. Inf. Comput. Sci.* 44:727-740
- Stahl M., Böhm H.J. (1998) Development of filter functions for protein-ligand docking. *J. Mol. Graph. Model.* 16:121-132
- Taylor R.D., Jewsbury P.J., Essex J.W. (2002) A review of protein-small molecule docking methods. *J. Comput. Aided Mol. Des.* 16:151-166
- Taylor R.D., Jewsbury P.J., Essex J.W. (2003) FDS: flexible ligand and receptor docking with a continuum solvent model and soft-core energy function. *J. Comput. Chem.* 24:1637-1656
- Vigers G.P., Rizzi J.P. (2004) Multiple active site corrections for docking and virtual screening. *J. Med. Chem.* 47:80-89
- Walters W.P., Stahl M.T., Murcko M.A. (1998) Virtual screening - an overview. *Drug Discov. Today* 3:160-178
- Waszkowycz B., Perlins T.D.J., Sykes R.A., Li J. (2001) Large-scale virtual screening for discovery leads in the postgenomic era. *IBM Sys. J.* 40:361-376