



HAL
open science

Evaluation et application de méthodes de criblage in silico

Hélène Guillemain

► **To cite this version:**

Hélène Guillemain. Evaluation et application de méthodes de criblage in silico. Pharmacologie. Conservatoire national des arts et métiers - CNAM, 2012. Français. NNT : 2012CNAM0849 . tel-00814270

HAL Id: tel-00814270

<https://theses.hal.science/tel-00814270>

Submitted on 16 Apr 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE ARTS ET METIERS

Génomique, Bioinformatique et applications

THÈSE présentée par :

Hélène GUILLEMAIN

soutenue le : **25 octobre 2012**

pour obtenir le grade de : **Docteur du Conservatoire National des Arts et Métiers**

Discipline/ Spécialité : Bioinformatique

**Evaluation et application de méthodes de
criblage *in silico***

THÈSE dirigée par :

M. ZAGURY Jean-François
M. MONTES Matthieu

Professeur, Directeur de thèse, Cnam
Docteur, Co-directeur de thèse Cnam

RAPPORTEURS :

M. GASTREICH Marcus
M. LANGER Thierry

Docteur, BioSolveIT GmbH
Professeur, Prestwick Chemical SAS

JURY :

M. LATOUCHE Aurélien

Professeur, Cnam

Si vous cachez votre ignorance,
vous ne recevrez pas de coups
et vous n'apprendrez rien.

Fahrenheit 451 – Ray Bradbury

Remerciements

Dans un premier temps, je tiens tout particulièrement à remercier le Pr. Jean-François Zagury de m'avoir accordé sa confiance pendant toutes ces années. Et un grand merci à Matthieu de m'avoir encadrée et soutenue pendant cette thèse.

Je suis très reconnaissante au Pr. Thierry Langer et au Dr. Marcus Gastreich d'avoir accepté d'examiner et de juger cette thèse. Je remercie également le Pr. Aurélien Latouche de participer à ce jury.

Je remercie aussi l'Agence Nationale de la Recherche et de la Technologie pour avoir financé cette thèse.

De profonds remerciements également à toute l'équipe pour leur folie et leur bonne humeur : Nathalie et Nesrine, qui seront dans la même galère que moi dans pas si longtemps ;) ; les "génomiques" Sigrid (alias "j'ai perdu mon badge"), Lieng (le geek psychopathe), Cédric (monsieur débats), Vincent (et son quart d'heure de folie journalier et extensible), Pierre (voisin de bureau à temps partiel), Damien (le petit nouveau), Taoufik (le tueur insoupçonné) et Jean-Louis (l'indépendantiste corse) ; les "cytokines" Hadley (ma soeur de thèse :)), Lucille (la batgirl), Gabi (le chanteur pour enfants) et Julie (l'autre petite nouvelle). Egalement un très grand merci à Hervé pour avoir pris le temps de corriger mon résumé ;) et à Christiane, la maman de l'équipe :). Et bien sûr, je n'oublie pas les "expats" Sophie (ancienne voisine de bureau et tigrou addict), Olivier et Rojo ainsi qu'à tous les stagiaires qui ont marqué cette équipe d'une manière ou d'une autre.

Enfin, un immense merci à ma famille et mes amis qui m'ont supportée pendant cette thèse (c'était déjà pas forcément facile avant) et m'ont toujours soutenue.

Glossaire

ADME-tox : Absorption Distribution Métabolisme Excrétion Toxicité

ASASA : Apolar Solvent Accessible Surface Area

AUC : Area Under the Curve

BEDROC : Boltzmann-Enhanced Discrimination of ROC

DUD : Directory of Useful Decoys

FE : Facteur d'Enrichissement

HTS : High Throughput Screening

IBAC : Interactions-Based Accuracy Classification

IC50 : Half Maximal Inhibitory Concentration

LBVLS : Ligand Based Virtual Ligand Screening

PDB : Protein Data Bank

PMF : Potential of Mean Force

RDE : Relative Displacement Error

RIE : Robust Initial Enhancement

RMSD : Root Mean Square Deviation

ROC : Receiver Operating Characteristic

SASA : Solvent Accessible Surface Area

SBVLS : Structure Based Virtual Ligand Screening

SDF : Structure Data File

SMILES : Simplified Molecular-Input Line-Entry System

TNF α : Tumor Necrosis Factor alpha

Résumé

Lors de la conception de médicaments, le criblage *in silico* est de plus en plus utilisé et les méthodes disponibles nécessitent d'être évaluées. L'évaluation de 8 méthodes a mis en évidence l'efficacité des méthodes de criblage *in silico* et des problèmes de construction de la banque d'évaluation de référence (DUD), la conformation choisie pour les sites de liaison n'étant pas toujours adaptée à tous les actifs. La puissance informatique actuelle le permettant, plusieurs structures expérimentales ont été choisies pour tenter de mimer la flexibilité des sites de liaison. Un autre problème a été mis en évidence : les métriques d'évaluation des méthodes souffrent de biais. De nouvelles métriques ont donc été proposées, telles que BEDROC et RIE. Une autre alternative est proposée ici, mesurant la capacité prédictive d'une méthode en actifs. Enfin, une petite molécule active sur le TNF α *in vitro* et *in vivo* sur souris a été identifiée par un protocole de criblage *in silico*. Ainsi, malgré le besoin d'amélioration des méthodes, le criblage *in silico* peut être d'un important soutien à l'identification de nouvelles molécules à visée thérapeutique.

Mots clés : criblage *in silico*, docking, similarité 3D, métriques d'évaluation, courbe de prédictibilité, ensemble docking, TNF α

Résumé en anglais

Since the introduction of virtual screening in the drug discovery process, the number of virtual screening methods has been increasing and available methods have to be evaluated.

In this work, eight virtual screening methods were evaluated in the DUD database, showing adequate efficiency. This also revealed some shortcomings of the DUD database as the binding site conformation used in the DUD was not relevant for all the actives.

As computational power now permits to address this issue, classical docking runs have been performed on several X-ray structures, used to represent the binding site flexibility. This also revealed that evaluation metrics show some biases. New evaluation metrics have thus been proposed, e.g. BEDROC and RIE. An alternative method was also proposed using predictiveness curves, based on compound activity probability

Finally, a virtual screening procedure has been applied to TNF α . A small molecule inhibitor, showing *in vitro* and *in vivo* activity in mice, has been identified. This demonstrated the value of virtual screening for the drug discovery process, although virtual screening methods need to be improved.

Keywords: virtual screening, docking, 3D similarity, evaluation metrics, predictiveness curve, ensemble docking, TNF α

Table des matières

<u>Remerciements.....</u>	<u>5</u>
<u>Glossaire.....</u>	<u>7</u>
<u>Résumé.....</u>	<u>8</u>
<u>Résumé en anglais.....</u>	<u>9</u>
<u>Table des matières.....</u>	<u>10</u>
<u>Liste des tableaux.....</u>	<u>11</u>
<u>Liste des figures.....</u>	<u>12</u>
<u>Première</u>	<u>partie :</u>
<u>Introduction.....</u>	<u>15</u>
<u>I.1. Introduction générale.....</u>	<u>17</u>
<u>I.2. Criblage de petites molécules basé sur la structure d'un ligand de référence (LBVLS).....</u>	<u>28</u>
<u>I.3. Criblage de petites molécules basé sur la structure de la protéine cible (SBVLS).....</u>	<u>32</u>
<u>I.4. Evaluation de méthodes.....</u>	<u>43</u>
<u>I.5. Objectifs de thèse.....</u>	<u>51</u>
<u>Deuxième</u>	<u>partie :</u>
<u>Résultats.....</u>	<u>52</u>
<u>II.1. Evaluation de méthodes.....</u>	<u>54</u>
<u>II.2 Application : Petites molécules inhibitrices du TNFα oralement actives in vivo.....</u>	<u>111</u>
<u>Conclusion.....</u>	<u>131</u>
<u>Résumé.....</u>	<u>148</u>
<u>Résumé en anglais.....</u>	<u>148</u>

Liste des tableaux

Table 1 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen inférieur à 350Å ³ (petits) et un pourcentage moyen d'hydrophobicité inférieur à 70% (polaires)	82
Table 2 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen inférieur à 350Å ³ (petits) et un pourcentage moyen d'hydrophobicité supérieur à 70% (hydrophobes)	83
Table 3 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen supérieur à 350Å ³ (grands) et un pourcentage moyen d'hydrophobicité inférieur à 70% (polaires)	84
Table 4 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen supérieur à 350Å ³ (grands) et un pourcentage moyen d'hydrophobicité supérieur à 70% (hydrophobes)	85
Table 5 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surfex-dock et ICM pour chaque structure des systèmes petits et polaires.	86
Table 6 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surfex-dock et ICM pour chaque structure des systèmes petits et hydrophobes.....	87
Table 7 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surfex-dock et ICM pour chaque structure des systèmes grands et polaires.....	88
Table 8 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surfex-dock et ICM pour chaque structure des systèmes grands et hydrophobes.....	89

Liste des figures

Figure 1 : Usage de méthodes in silico dans le processus de développement d'un médicament	19
Figure 2 : Exemples de formats de fichiers pour le composé ZINC4517771.....	23
Figure 3 : Représentation du pharmacophore d'une petite molécule à l'aide de ICM57. Les accepteurs de liaison hydrogène sont représentés en vert, les cycles aromatiques en orange, les groupements hydrophobes en turquoise et la charge négative en bleu.....	28
Figure 4 : Représentation des interactions entre le site de liaison du facteur Xa et son ligand cocrystallisé (code PDB : 1NFY). Représentation faite avec Pymol (gauche) et Poseview ⁹⁹ (droite).....	34
Figure 5 : Représentation de la sélection de n molécules dans une chimiothèque de N composés ¹⁵⁶	45
Figure 6 : Relations entre les vrais positifs (VP), les faux positifs (FP), les vrais négatifs (VN) et les faux négatifs (FN). Ces valeurs représentent la concordance entre les résultats attendus et les résultats observés pour un seuil donné. Elles permettent de calculer la sensibilité (Se), la spécificité (Sp), la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN)..	46
Figure 7 : Exemples de distribution des actifs dans les n premières molécules d'une banque classée à partir d'une fonction de score.....	47
Figure 8 : Représentation de 3 courbes de ROC (courbe idéale, courbe d'une distribution homogène et courbe intermédiaire). La courbe de la distribution homogène est donnée par l'équation $Se=1-Sp$ ¹⁵⁶	48
Figure 9 : Exemples de ROC AUC non informatives ¹⁵⁶ . (a) la courbe 1 est meilleure que la courbe 2 pour une spécificité élevée, mais les AUC sont identiques, (b) la courbe 2 est meilleure que la courbe 1 pour une spécificité élevée, mais l'AUC de la courbe 1 est meilleure.....	49
Figure 10 : Courbes de ROC globales. Les méthodes SBVLS sont représentées par un trait plein et les méthodes LBVLS par des pointillés. Les méthodes Surflex-dock/Surflex-sim, FRED/ROCS, FlexX/FlexS et ICM/ICMsim sont représentées en bleu, rouge, jaune et vert, respectivement.....	71
Figure 11 : Courbes de densité des distribution de scores attribués aux actifs (bleus) et aux inactifs (rouges) par ICM pour les 39 systèmes de la DUD. Pour chaque distribution, la partie	

hachurée représente les scores se trouvant au-delà de l'écart-type à la moyenne.	96
Figure 12 : Courbes de densité des distribution de scores attribués aux actifs (bleus) et aux inactifs (rouges) par Surfex-dock pour les 39 systèmes de la DUD. Pour chaque distribution, la partie hachurée représente les scores se trouvant au-delà de l'écart-type à la moyenne.....	99
Figure 13 : Courbes d'enrichissement en semi-log pour Surfex-dock (trait plein noir) et ICM (pointillés rouges) sur les 39 systèmes de la DUD. La distribution homogène des actifs est représentée en gris.	102
Figure 14 : Courbes de ROC pour Surfex-dock (trait plein noir) et ICM (pointillés rouges) sur les 39 systèmes de la DUD. La distribution homogène des actifs est représentée en gris.....	105
Figure 15 : Courbes de prédictibilité pour Surfex-dock (trait plein noir) et ICM (pointillés rouges) en semi-log, sur les 39 systèmes de la DUD. Les mêmes graphes avec une échelle linéaire sont représentés en encadré. Les pointillés gris au niveau de la prévalence représentent une distribution homogène des actifs dans la banque.....	108

Première partie :

Introduction

I.1. Introduction générale

La conception de médicaments est un processus coûteux, tant au niveau du temps que de l'argent investi. La mise au point d'un médicament coûte ainsi entre 230 millions et 1,43 milliards d'euros capitalisés, en fonction des méthodes et données utilisées, ainsi que de la période de l'étude¹. De même, selon différentes études²⁻⁴, entre la phase préliminaire de recherche pour l'identification de têtes de série et la mise sur le marché d'une molécule, de 7 à 17 ans peuvent s'écouler. Lors de la phase préliminaire de recherche, des méthodes de criblage à haut débit (ou HTS pour High Throughput Screening) peuvent être utilisées. Ces méthodes permettent l'identification de touches parmi les composés des banques de petites molécules chimiques (chimiothèques) disponibles. Les meilleures touches sont ensuite optimisées afin d'obtenir des molécules ayant une meilleure activité envers la cible d'intérêt (têtes de séries)⁵. La molécule la plus prometteuse est sélectionnée pour entrer en phases pré-clinique puis cliniques.

La phase pré-clinique consiste à tester l'activité et la toxicité du candidat-médicament sur cellules et tissus puis sa biodisponibilité, son activité et sa toxicité sur animaux (propriétés ADME-tox : Absorption, Distribution, Métabolisme, Excrétion et Toxicité). Ces tests permettent de déterminer les effets de la molécule sur les récepteurs cellulaires et les organes, la marge thérapeutique (dose toxique / dose efficace), ainsi que sa toxicité, aiguë et chronique. Le composé présentant une activité significative sans effets secondaires majeurs passe en étude clinique chez l'homme. La phase I clinique est menée sur de faibles effectifs d'individus sains, sélectionnés pour recevoir la molécule candidate. Des doses croissantes du composé sont administrées pour déterminer le seuil d'intolérance. Les essais cliniques de phase II se font sur de plus grands effectifs, cette fois de patients. Les conditions optimales de prescription (posologie, rythme, durée) sont déterminées et les effets thérapeutiques et indésirables sont définis précisément. La phase III clinique va permettre la validation des résultats des deux étapes précédentes à une large échelle. Le composé est comparé à un placebo et/ou au traitement de référence, s'il existe. Le dossier d'enregistrement peut alors être monté pour obtenir l'autorisation de mise sur le marché. Au moment de sa commercialisation, le médicament entre dans la phase IV clinique, phase de pharmacovigilance à long terme,

répertorient les effets secondaires rares et permettant de mettre en place de nouvelles indications thérapeutiques et de nouvelles formulations.

La biochimie et les méthodes pharmacologiques traditionnelles ont permis de tester *in vitro* les composés sur les cibles d'intérêt. Cependant, à cause des limitations technologiques, de 20 à 50 composés pouvaient être testés de manière hebdomadaire⁶. Actuellement, la conception de médicament est largement assistée par ordinateur dans l'industrie pharmaceutique pour accélérer le processus⁷, permettant de tester quelques millions de composés, selon les méthodes et la puissance informatique disponibles⁸. Ces aspects seront détaillés plus tard dans l'introduction.

Cependant, moins d'un candidat sur 10 entrant en phase clinique pourra recevoir l'autorisation de mise sur le marché^{5; 9}. Ce haut taux d'échec est en partie dû à une mauvaise compréhension des mécanismes pathologiques et le manque d'anticipation du profil ADME-tox pendant les premières étapes de la conception de médicament. Par exemple, une faible solubilité ou une faible perméabilité du composé expliquent une grande partie des échecs dus à la pharmacocinétique, qui représentent environ 30% des arrêts en essai clinique^{9; 10}. Concernant le criblage à haut débit, tant traditionnel que virtuel, le choix de la chimiothèque, dont sont issues les molécules qui seront criblées sur la cible d'intérêt, a aussi une grande importance. Cet aspect est donc détaillé dans le chapitre suivant.

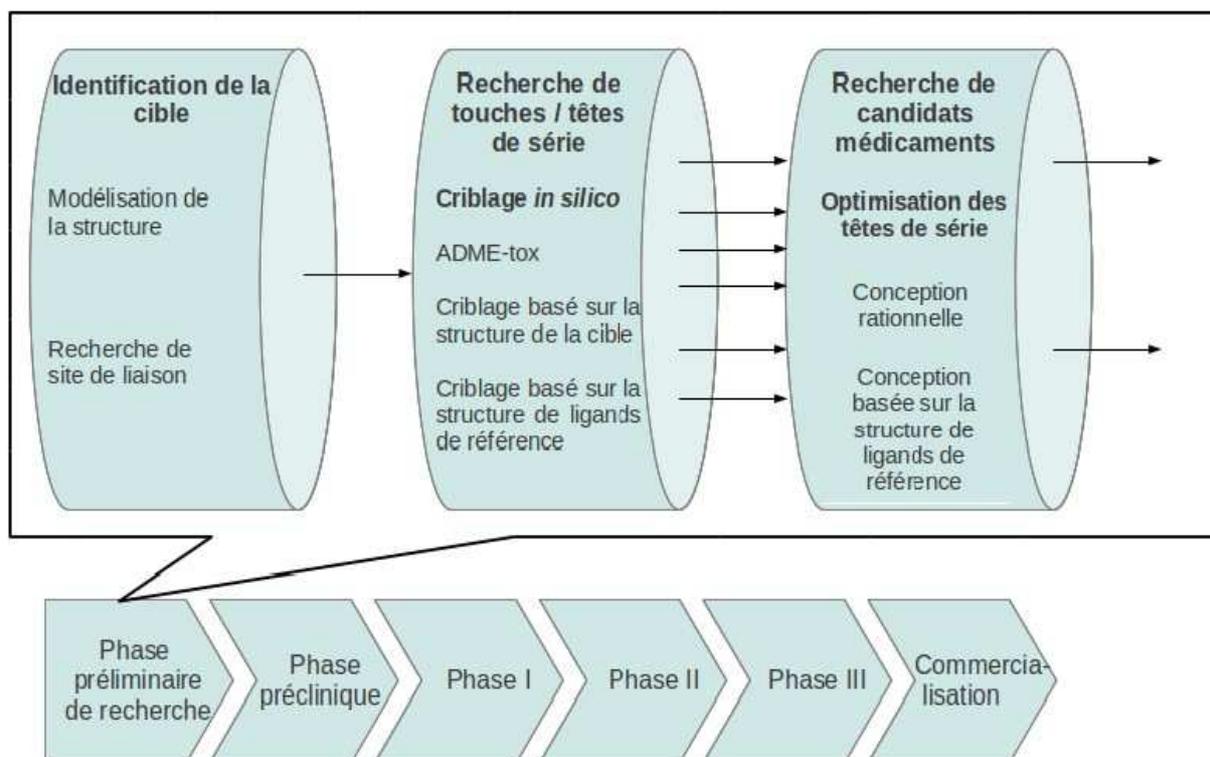


Figure 1: Usage de méthodes *in silico* dans le processus de développement d'un médicament

I.1.1. Les chimiothèques

Le nombre de molécules potentiellement synthétisables est immense, de l'ordre de 10^{60} composés¹¹. Cependant, on estimait en 2006 qu'environ 80 millions de molécules existaient physiquement¹². Selon la classification de Bradley¹³, il existe différents types de chimiothèques : les bases de données de structures chimiques, les chimiothèques généralistes de produits utilisables en synthèse, celles contenant exclusivement des produits biologiquement actifs et enfin les chimiothèques de criblage.

La majorité des chimiothèques fournissent les structures 1D et 2D des composés. Des programmes générant les structures tridimensionnelles des molécules peuvent donc être utilisés (OMEGA¹⁴, CORINA¹⁵, Concord¹⁶, ICM¹⁷). Mais la structure 3D ainsi générée ne représente pas forcément une conformation active de la molécule.

I.1.1.1. Chimiothèques principales

I.1.1.1.1. Bases de structures chimiques

Les bases de structures chimiques sont des bases de référencement de la plupart des composés du monde académique, voire du monde industriel lorsque l'information est disponible dans un brevet. Le Chemical Abstract Service¹⁸ contient différentes chimiothèques. La CAS Registry est constituée des substances identifiées dans les journaux, brevets, catalogues chimiques et sources internet fiables. Elle répertorie ainsi plus de 67 millions de substances¹⁹. La CAplus contient des substances provenant de journaux, brevets, rapports techniques et de comptes-rendus de conférences. Plus de 35 millions de substances sont ainsi répertoriées²⁰. D'autres bases de données existent, telles que SPRESIweb²¹ qui contient actuellement 11,8 millions de composés, ou Beilstein (maintenant distribuée sous le nom de Reaxys). Elles ne sont pas utilisables directement en chemoinformatique car elles ne contiennent pas de fichiers regroupant tous les produits.

I.1.1.1.2. Chimiothèques généralistes

Les chimiothèques généralistes contiennent des molécules directement achetables auprès de différents fournisseurs. La quantité, la qualité, la disponibilité et le prix de chaque produit sont disponibles, permettant une commande groupée des différents produits souhaités. A cela s'ajoutent généralement des fichiers informatiques regroupant les structures des différents composés et facilitant leur utilisation par les bioinformaticiens. Certaines de ces molécules ont une activité biologique connue sur des cibles d'intérêt. Mais ces chimiothèques n'ayant pour but que de fournir des composés, cette activité n'est pas toujours indiquée. La chimiothèque ACD Chemicals²² est une des principales banques de petites molécules, contenant plus de 3,8 millions de composés synthétiques. D'autres chimiothèques sont également disponibles, contenant le même type d'informations, bien que plus petites, telles que Sigma-Aldrich²³ (plus de 212.000 molécules) ou encore Specs²⁴ (240.000 composés). Sigma-Aldrich permet aussi de faire des recherches de composés par similarité.

I.1.1.1.3. Chimiothèques de produits biologiquement actifs

Ces chimiothèques contiennent des composés pour lesquelles des informations d'activité biologique sont connues. Les molécules présentes dans ces banques ne sont pas forcément disponibles physiquement. Mais la présence de fichiers informatiques regroupant ces molécules rend ces chimiothèques utiles et utilisables en chemoinformatique. Certaines banques contiennent des médicaments en développement ou commercialisés, comme la MDL Drug Data Report (MDDR²⁵), qui contient pour l'instant plus de 150.000 composés biologiquement actifs avec la structure chimique, des informations concernant le brevet quand il existe, des propriétés calculées à partir de la structure chimique, l'activité biologique exprimée qualitativement²⁶, une description de son action thérapeutique ou encore son niveau de développement en R&D. Des informations similaires sont disponibles à partir du Derwent World Drug Index (WDI²⁷), une chimiothèque qui regroupe 80.000 médicaments ou molécules pharmacologiquement actives issus de journaux ou conférences. Des données d'interactions, d'usage médical, d'indication ou encore de mécanisme d'action sont également disponibles pour chaque composé. Des chimiothèques plus spécialisées sur une maladie ou un type de maladie précis sont également disponibles, comme celle fournie par le National Cancer Institute qui comprend quatre chimiothèques publiquement disponibles dont deux centrées sur le SIDA et le cancer²⁸. La chimiothèque ChEMBLdb²⁹ regroupe quasiment 1.213.300 composés bioactifs, fournissant leur structure 2D, des propriétés calculées (par exemple le logP, le poids moléculaire...) et les informations sur sa bioactivité (les constantes de liaison, les données ADME-tox...).

I.1.1.1.4. Chimiothèques de criblage

Les chimiothèques de criblage sont directement destinées au criblage à haut débit et peuvent être utilisées en chemoinformatique. Il en existe plusieurs formes, répondant à des besoins spécifiques :

- Banques de châssis moléculaires utilisables en synthèse chimique
- Banques de produits naturels. En effet, entre 1994 et 2007, près de la moitié des médicaments approuvés par la Food and Drug Administration étaient basés sur des produits naturels^{30; 31}.

- Banques de diversité chimique comprenant des composés sélectionnés pour couvrir l'espace chimique comme Tripos Leadscreens³², Chemdiv CombiLab³³ ou Chembridge diverset³⁴.
- Banques focalisées proposant des composés potentiellement actifs sur une cible particulière ou une famille de cibles. Par exemple, ChemBridge³⁵ propose plusieurs chimiothèques spécialisées sur les kinases (KinaSet et KinaCore), les récepteurs couplés aux protéines G (GPCR Library), les canaux ioniques (ION Channel Set et IONCore Library) ou encore les récepteurs à hormones nucléaires (NHRCORE Library).

Plutôt que de proposer des composés entiers, certaines chimiothèques contiennent des fragments moléculaires.

Le concept sous-jacent du criblage de fragments a été décrit par Jencks³⁶ et Ariens³⁷ qui ont montré que les molécules « drug like » peuvent être considérées comme étant la combinaison de 2 ou plusieurs fragments. Les fragments moléculaires ont une complexité moléculaire plus faible que les composés entiers, ainsi, la probabilité que leurs caractéristiques (donneurs/accepteurs de liaisons hydrogène, forme, hydrophobicité...) correspondent à celles du site de liaison est plus importante³⁸. Ces fragments ont une affinité de liaison plus faibles que la plupart des touches trouvées par criblage de composés entiers. Une plus grande proportion de leurs atomes sont impliqués dans les interactions d'intérêt avec la cible, ce qui entraîne une plus grande énergie de liaison par unité de masse moléculaire³⁹. ChemBridge⁴⁰ et ChemDiv⁴¹ proposent, par exemple, des chimiothèques de fragments.

Il existe ainsi un grand nombre de chimiothèques et il convient de sélectionner la chimiothèque de travail selon le projet. En effet, une chimiothèque spécialisée sur une cible particulière ne conviendra pas obligatoirement si, bien que travaillant sur la cible en question, de nouveaux châssis moléculaires sont recherchés.

I.1.1.2. Formats de chimiothèques

Il existe plusieurs formats électroniques pour représenter les petites molécules. Les plus utilisés pour les chimiothèques sont SMILES, SDF, MOL2 ou PDB.

imparfaite. Deux améliorations lui ont été apportées : les notations SMILES canoniques et SMILES isomériques.

La notation SMILES canonique a été introduite dès 1989⁴³ en convertissant la structure d'une molécule à l'aide de deux algorithmes séparés : CANON et GENES. CANON permet d'appliquer à chaque atome de la molécule un nombre canonique, c'est-à-dire qu'il sera le même, quelque soit l'atome de départ. L'algorithme GENES part ensuite de l'atome ayant le nombre canonique le plus faible pour construire le code SMILES. Cette règle implique qu'un atome terminal est choisi s'il existe. De la même manière, à chaque branche, l'atome voisin ayant le plus petit nombre est choisi pour l'exploration d'une nouvelle ramification. Cette manière d'écrire un code SMILES évite d'avoir une redondance trop importante dans une chimiothèque. En revanche, cette notation ne permet pas de faire la discrimination entre deux molécules stéréo-isomères.

Deux molécules sont des stéréo-isomères lorsque, bien que leurs formules semi-développées soient identiques, elles divergent l'une de l'autre par l'agencement spatial de leurs atomes. Ces composés ayant des propriétés physico-chimiques différentes, il était nécessaire d'introduire des précisions de chiralité ou de configuration cis-trans dans la notation SMILES⁴⁴. La spécification chirale générale se fait en trois parties :

- L'indicateur de la chiralité : @
- La classe de la chiralité : par exemple TH pour un centre tétraédral
- L'ordre d'écriture des atomes voisins : par exemple 2 représentant un ordre anti-horaire des atomes autour d'un centre tétraédral.

Le degré de chiralité d'un atome reflète le nombre d'atomes auxquels il est lié. Pour chacun de ces degrés, plusieurs classes peuvent exister, la classe considérée par défaut pouvant alors être implicite. Par exemple, un centre chiral tétraédral, présent lorsque ses 4 atomes voisins sont différents, va être représenté dans la notation SMILES par un @ suivi des atomes voisins dans l'ordre horaire, ou par @@ (équivalent à @2) suivi des atomes voisins dans l'ordre anti-horaire. Le code TH peut ne pas être indiqué, cette classe étant la classe par défaut pour les centres chiraux de degré 4.

les règles de notation isomérique permettent de spécifier, complètement ou non, la chiralité de n'importe quelle structure. Ainsi, les liaisons directionnelles (\ et /) peuvent être utilisées pour indiquer la direction relative des atomes voisins d'une double liaison.

En raison de sa particularité de n'utiliser qu'une ligne par molécule, ce format est très utile pour le stockage de très grandes banques de petites molécules.

I.1.1.2.2. Les formats de fichier 3D

Différents formats existent donnant des informations détaillées (type d'atome, coordonnées spatiales, charge partielle pour le MOL2) sur une molécule, à raison d'un atome par ligne (là où le format SMILES utilise une ligne par molécule). La connectivité entre atomes est aussi détaillée avec le type de liaison (liaisons simples, doubles, triples, aromatiques) et les atomes impliqués. Parmi les formats les plus couramment utilisés, on trouve le format PDB (utilisé pour construire la Protein Data Bank^{45; 46}), le format SDF (Structure Data Format⁴⁷) ou le format MOL2⁴⁸. Le format PDB est surtout utilisé pour les macromolécules, telles que les protéines.

De nombreux programmes de conversion permettent le passage, parfois nécessaire, d'un format à un autre, lors d'un protocole de criblage. La conversion entre deux formats 3D est relativement aisée, cependant, lors du passage d'un format SMILES à un format 3D, l'information 3D doit être générée par un programme de conversion (OMEGA¹⁴, CORINA¹⁵, Concord¹⁶, ICM¹⁷). La qualité de ces fichiers influe directement sur les résultats des calculs les utilisant.

Les différents états d'ionisation des groupements titrables peuvent influencer sur leurs propriétés physico-chimiques et tous les programmes de criblage ne les prennent pas en compte dynamiquement. Chaque état d'ionisation peut alors être présent dans le fichier de structure. Ligprep⁴⁹ ou QuacPac⁵⁰ permettent cet échantillonnage.

Les composés issus des chimiothèques peuvent être filtrés selon leur profil ADME-tox (détaillé dans le point suivant) pour ne pas tester, par criblage in silico, des composés qui pourraient ensuite être écartés à cause d'un profil ADME-tox inadéquat.

I.1.2 Profil ADME-tox

Le profil ADME-tox d'une molécule est l'ensemble des paramètres caractérisant sa biodisponibilité dans l'organisme, c'est à dire, son absorption, sa distribution, son métabolisme, son excrétion et sa toxicité. Lorsque la voie d'administration orale est choisie, l'absorption du médicament se fera par la muqueuse intestinale, soit par diffusion passive, soit par transport actif. Une fois dans l'organisme, la molécule va être dégradée en métabolites. Ce processus est déterminant sur plusieurs points, lors de l'administration du médicament⁵¹ :

- La posologie, la vitesse de dégradation du composé déterminant en partie sa fréquence d'administration
- Les contre-indications, dues aux interférences entre les métabolismes de deux médicaments
- Sa toxicité, le médicament pouvant être rendu plus, ou moins, toxique en étant métabolisé.

La connaissance de la toxicité d'un médicament est cruciale. Il est en effet préférable qu'un médicament ait un index thérapeutique élevé, c'est à dire que la concentration plasmatique toxique soit significativement plus élevée que la concentration plasmatique thérapeutique⁵¹.

La démocratisation du criblage *in silico* a entraîné la nécessité de modèles ADME-tox pour éliminer rapidement les composés ayant les propriétés physico-chimiques les moins similaires avec les médicaments disponibles sur le marché (qui ne sont pas « drug like »⁵²). Typiquement, une molécule est considérée drug-like si elle répond à au moins 3 des règles de Lipinski⁵³ :

- Masse moléculaire inférieure ou égale à 500 Da
- Coefficient de partage entre l'eau et un solvant organique (logP) inférieur ou égal à 5
- Moins de 5 groupes donneurs de liaisons hydrogène
- Moins de 10 groupes accepteurs de liaisons hydrogène

Depuis, des modifications ont été apportées, notamment par Ghose et al⁵⁴ qui préconisent par exemple un logP entre -0,4 et 5,6, une masse moléculaire comprise entre 160 et 480 Da et un nombre d'atomes entre 20 et 70.

La prédiction de l'ADME-tox permet aussi d'identifier les faiblesses potentielles des séries sélectionnées par criblage, ce qui pourra servir pendant leur optimisation. Au moment de

l'identification de têtes de série, le profil ADME-tox peut être déterminé en utilisant les méthodes expérimentales et *in silico* de manière itérative³.

Les molécules dont le profil ADME-tox est satisfaisant pourront être soumises au criblage *in silico*. Si la structure de la protéine cible est connue, une approche basée sur la structure de cette dernière est généralement utilisée. Si cette structure est inconnue mais qu'une ou plusieurs molécules sont connues pour être actives sur cette protéine, une approche basée sur leurs structures pourra être envisagée⁵⁵. Cependant, ces deux types d'approches peuvent être utilisées dans un protocole hiérarchique, un certain nombre de molécules étant sélectionnées grâce à une approche basée sur la structure d'un ligand de référence avant d'être criblées dans le site de liaison d'intérêt⁵⁶. Ces deux types d'approches sont détaillées dans les deux prochaines parties, en commençant par les méthodes basées sur la structure d'un ou de plusieurs ligands de référence.

I.2. Criblage de petites molécules basé sur la structure d'un ligand de référence (LBVLS)

I.2.1 Pharmacophores

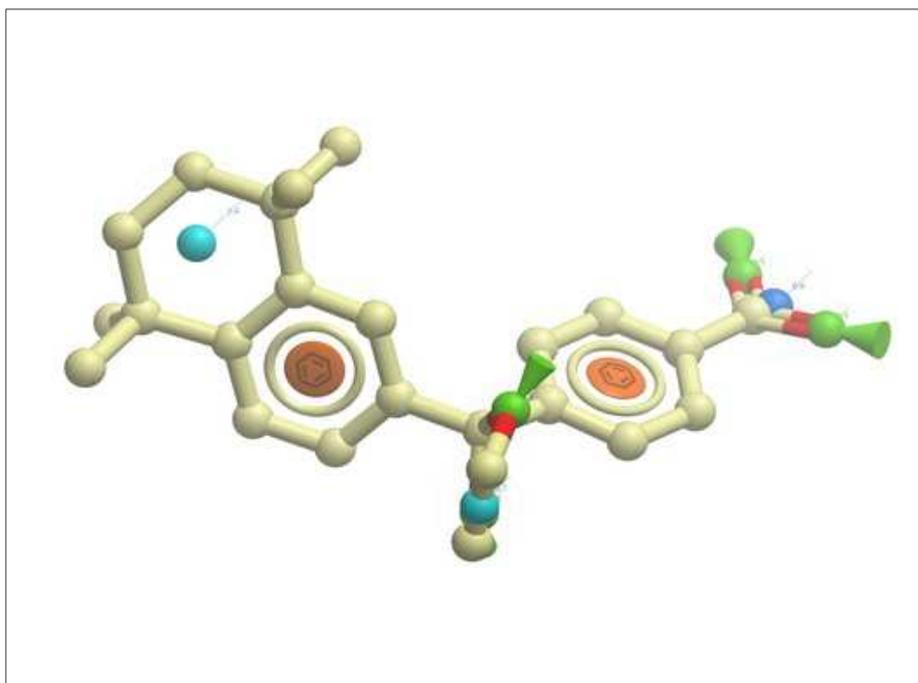


Figure 3 : Représentation du pharmacophore d'une petite molécule à l'aide de ICM⁵⁷. Les accepteurs de liaison hydrogène sont représentés en vert, les cycles aromatiques en orange, les groupements hydrophobes en turquoise et la charge négative en bleu.

La première utilisation du terme pharmacophore est trouvée dans un article de Lemont Kier en 1971⁵⁸. Depuis, ces caractéristiques seules n'étant pas suffisantes pour expliquer l'activité biologique d'une molécule, leur disposition spatiale a été ajoutée. Actuellement, le pharmacophore est défini comme étant l'agencement tridimensionnel des caractéristiques permettant à une molécule d'avoir une activité biologique spécifique⁵⁹. En d'autres termes, un pharmacophore est l'ensemble des points fonctionnels (donneurs/accepteurs de liaison

hydrogène, hydrophobes...) d'une molécule selon la conformation spatiale existant pendant sa fixation dans le site de liaison de la cible. Cette définition est à la base de son utilisation en criblage *in silico*, en s'appuyant sur le postulat selon lequel les molécules ayant le même pharmacophore auront une activité similaire sur la cible, pouvant permettre de s'affranchir du châssis moléculaire de la molécule servant de référence, ce qui n'est pas le cas, par exemple, des méthodes de similarité 2D⁵⁵. Quand plusieurs molécules sont connues comme étant actives sur la cible d'intérêt, les points fonctionnels communs à l'ensemble (ou à la majorité) d'entre elles sont retenus pour constituer le pharmacophore sur lequel seront criblées les molécules de test⁶⁰.

La majorité des méthodes de construction de pharmacophores fonctionnent sur le même principe : 1. les composés connus pour se fixer dans le même site de liaison de la cible seront utilisés pour construire le modèle, 2. une recherche conformationnelle est faite sur ces composés, afin de les aligner de la meilleure manière qui soit. Cette étape peut aussi se faire plus tard, pendant l'identification du motif pharmacophorique, 3. les points fonctionnels communs sont extraits, 4. les éléments pharmacophoriques de chaque composé sont représentés dans l'espace par des points, 5. des motifs pharmacophoriques sont identifiés à partir des localisations relatives de chaque point pharmacophorique, et 6. un score est attribué à chaque motif⁵⁵. L'identification du motif peut être faite soit en utilisant le maximum de points pharmacophoriques communs à tous les ligands du modèle, soit en utilisant cette règle de manière plus souple, c'est à dire que certains points pharmacophoriques peuvent être conservés même s'ils ne sont pas partagés par toutes les molécules du modèle⁵⁵.

La construction d'un pharmacophore nécessite beaucoup d'informations à la fois sur la structure de ligands de référence et sur leur activité, afin de déterminer quelles sont les parties impliquées dans cette activité et qui serviront donc pour identifier de nouvelles molécules actives⁶¹.

I.2.2 Similarité

Ces méthodes s'intéressent directement à la structure des molécules de référence et des composés à cribler. Comme pour les pharmacophores, l'idée sous-jacente est que les molécules qui ont une structure similaire vont vraisemblablement partager des propriétés

similaires⁶². L'étude de la similarité entre différentes molécules se base sur différents descripteurs, bi ou tridimensionnels.

Les méthodes de similarité 2D sont basées sur le partage de descripteurs 2D entre les références et les composés criblés. Ces descripteurs sont généralement des séquences binaires dénotant la présence (1) ou l'absence (0) de sous-structures^{63;64}.

Les méthodes de similarité 3D sont non seulement basées sur les structures 2D des molécules, mais aussi sur leurs conformations. Cependant, la conformation bioactive d'un ligand est, le plus souvent, connue seulement si la structure expérimentale de la protéine cible est connue également⁶⁵. Une ou plusieurs conformations tridimensionnelles de cette molécule peuvent être générées (avec OMEGA⁶⁶, par exemple), servant d'hypothèses pour le criblage des molécules testées. Lorsque plusieurs ligands sont connus, une première recherche conformationnelle est faite afin d'avoir un alignement unique conduisant à un modèle prédictif d'une activité biologique⁶⁷. La recherche de similarité 3D a deux composantes principales : 1. la similarité de conformation des deux molécules, qui peut être prise en compte en tentant de minimiser les différences entre les deux surfaces moléculaires (Surflex-sim⁶⁷) ou de maximiser le recouvrement du volume des deux molécules (ROCS⁶⁸), 2. la similarité des points pharmacophoriques qui peut se retrouver dans l'utilisation de « couleurs » par ROCS⁶⁹, par exemple.

Une alternative à ces méthodes est la similarité de l'activité biologique. Les méthodes de recherche de similarité précédentes s'appuient sur le principe selon lequel des molécules de structures similaires auront plus de chances d'avoir des propriétés similaires⁶² et de se lier au même groupe de protéines⁷⁰. Cependant, la similarité chimique seule ne peut pas expliquer l'activité d'un composé sur une voie de signalisation ou un groupe de voies de signalisation. En effet, un effet secondaire de cette hypothèse est la mise de côté de molécules entraînant le même phénotype mais ayant des structures diverses, ces méthodes ne prenant pas en compte l'activité biologique⁷¹. Kauvar et al⁷² se sont basés sur les interactions observées entre un certain nombre de composés et plusieurs protéines pour construire des « empreintes d'affinité » et prédire ainsi les affinités de nouveaux composés. De la même manière, Fliri et al^{73; 74} ont prédit des interactions simultanées de molécules avec le protéome à l'aide d'empreintes appelées « biospectra ».

Une autre alternative, décrite plus récemment, est l'utilisation des réseaux bayésiens pour le criblage de petites molécules⁷⁵, alors qu'ils ont d'abord été développés pour la recherche de documents⁷⁶.

Les alignements entre les différentes molécules ont ensuite besoin d'être évalués. Ceci est fait à l'aide d'une métrique de similarité, qui possède deux composants principaux, selon Holliday et al⁷⁷ : 1. ce qui est utilisé pour la comparaison (représentation des molécules, descripteurs) et 2. le coefficient de similarité qui doit quantifier le degré de ressemblance entre deux représentations. Cependant, bien que le nombre de coefficients de similarité et de distance soit grand, beaucoup d'entre eux sont reliés les uns aux autres⁶¹. Ce qui explique la présence de redondance parmi ces différentes métriques, permettant de les regrouper en différentes classes⁷⁷⁻⁷⁹.

Il existe trois principaux types de métriques de similarité : les coefficients de distance (distance de Hamming), d'association (Jaccard/Tanimoto, Simpson...) et de corrélation (Pearson, Stiles...).

Les coefficients de distance vont permettre de quantifier le degré de différence entre deux composés. Par exemple, la distance de Hamming⁸⁰ : $D_{A,B} = a + b - 2c$ représente le nombre de différences entre les composés A et B, avec a , le nombre d'empreintes dans le composé A ; b , le nombre d'empreintes dans le composé B et c , le nombre d'empreintes qui apparaissent dans les deux composés⁶¹. Ces coefficients sont généralement les complémentaires de coefficients d'association, plus couramment utilisés⁷⁷.

Les coefficients d'association vont, quant à eux, quantifier le degré de similarité entre deux composés. Le coefficient de Tanimoto est le plus souvent utilisé dans les études cherchant à

évaluer la similarité entre différents composés⁸¹⁻⁸³ : $S_{A,B} = \frac{c}{a+b-c}$.

Le troisième type de coefficients comprend les coefficients de corrélation. Ceux-ci mesurent le degré de corrélation entre différentes valeurs, chacune d'entre-elles caractérisant une paire

de composés⁷⁷. Par exemple, le coefficient de Pearson : $C_{A,B} = \frac{cd - ba}{\sqrt{(c+a)(c+b)(a+d)(b+d)}}$ avec d , le nombre d'empreintes qui n'apparaissent chez aucun des deux composés.

I.3. Criblage de petites molécules basé sur la structure de la protéine cible (SBVLS)

I.3.1 Identification du site de liaison

L'identification du site de liaison est une étape cruciale des protocoles de criblage *in silico* basés sur la structure d'une macromolécule cible. Bork et al estiment que pour chaque protéine, de 3 à 10 partenaires d'interaction existent⁸⁴, n'utilisant pas nécessairement les mêmes sites de liaison. Il existe deux possibilités pour déterminer le site de liaison à cibler, en fonction des informations disponibles : 1. si une structure de la protéine en complexe avec une molécule active existe, son site de liaison peut être utilisé, 2. si la structure existante n'est pas complexée avec un ligand, certaines méthodes peuvent aider à la recherche de sites de liaisons potentiels à la surface de la protéine. Prymula et al⁸⁵ ont classé ces méthodes en trois groupes : 1. les méthodes basées sur la géométrie pure, 2. les méthodes énergétiques, 3. les méthodes dérivées de connaissances biochimiques.

Les méthodes géométriques recherchent des cavités ou des poches à la surface de la protéine, en ne se focalisant que sur la forme de la protéine en laissant parfois de côté les propriétés physico-chimiques des résidus. Ces méthodes peuvent utiliser une grille (PocketPicker⁸⁶, POCKET⁸⁷, LIGSITE⁸⁸) ou des sphères (PASS⁸⁹) pour déterminer les cavités présentes à la surface de la protéine. CASTp est un programme disponible en ligne, calculant les volumes et surfaces des poches de liaison à l'aide de modèles de surface accessible au solvant (surface de Richard) et de surface moléculaire (surface de Connolly)⁹⁰. La poche la plus grande est généralement considérée comme ayant le plus de chance d'être la poche de liaison⁹¹.

Les méthodes énergétiques recouvrent la surface protéique de sondes et calculent ensuite les interactions de Van der Waals et électrostatiques entre la protéine et ces sondes. Les zones dans lesquelles les interactions entre les sondes et la protéine sont énergétiquement favorables représentent les poches identifiées. Ces poches sont ensuite classées en fonction de la somme des énergies d'interaction de chaque sonde. La méthode Q-SiteFinder⁹¹ fait partie de cette

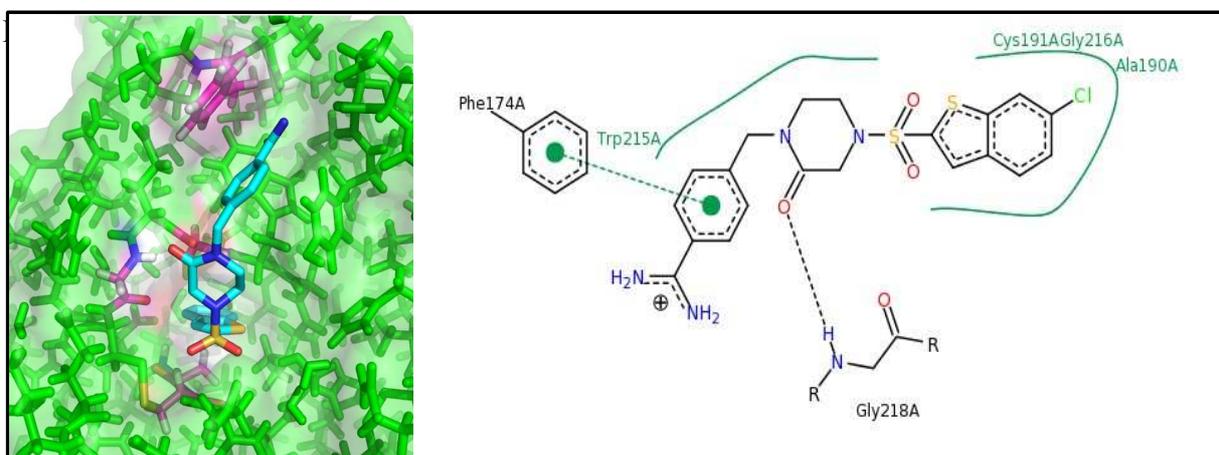
catégorie et utilise des groupements méthyle en guise de sondes. Il est à noter que la poche présentant l'énergie la plus favorable n'est pas forcément la plus grande, géométriquement⁹¹.

Les méthodes dérivées des connaissances biochimiques comprennent les méthodes statistiques⁹² (les réseaux neuronaux⁹³, les machines à vecteurs de support⁹⁴ ou la classification naïve bayésienne⁹⁵) et les méthodes de similarité avec des complexes connus⁹⁶. Ces dernières peuvent prendre en compte l'existence de similarité de structure locale entre la protéine étudiée et des sites de liaison connus⁹⁷ ou la conservation de résidus dans les sites de liaison⁹⁸.

Des méthodes hybrides peuvent également être utilisées, par exemple, Dai et al⁹⁶ utilisent d'abord un algorithme géométrique pour rechercher les différentes poches potentielles à la surface de la protéine, puis se basent sur la conservation de séquence de ces poches pour déterminer la poche la plus prometteuse.

I.3.2 Docking

Pendant la liaison, le ligand et le site de liaison de la macromolécule s'adaptent l'un à l'autre (induced fit). Pour mimer le plus justement possible cette liaison, les méthodes de docking devraient donc explorer l'espace conformationnel de ces deux entités. Cependant, explorer toutes les combinaisons possibles des conformations des résidus de la protéine, ne serait-ce que ceux appartenant au site de liaison, est pour l'instant très coûteux en temps de calcul. Une majorité de programmes négligent donc cette information et n'explorent que l'espace conformationnel de la petite molécule à tester.



I.3.2.1. Docking avec ligand rigide

Le ligand et la protéine sont tous les deux considérés comme étant rigides. Seuls les 6 degrés de liberté de translation et rotation sont explorés pour le ligand, c'est à dire que le ligand peut subir des translations et des rotations dans chacune des trois dimensions. Par exemple, le programme FRED¹⁰⁰ délimite le site de liaison à l'aide d'une boîte et positionne le ligand à l'intérieur de celle-ci. La recherche conformationnelle commence en énumérant toutes les translations et rotations possibles, avec au minimum 1 Å et 1,5 Å de RMSD, respectivement, entre chaque changement. Chaque pose sans clash est notée à l'aide d'une fonction de score. Les poses ayant les meilleurs scores sont affinées par le programme, qui teste les rotations et translations proches (RMSD supérieur à 0,5 Å et 0,75 Å, respectivement)⁸. D'autres programmes, tels que DOCK¹⁰¹ font aussi du docking rigide.

La précision de ces méthodes peut être améliorée en générant plusieurs conformères pour chaque ligand. Ces conformères peuvent être générés au préalable à l'aide de programmes comme OMEGA^{14; 66}, ROTATE¹⁰² ou FLOG (utilisation de bibliothèques de conformations pré-générées pour faire le docking¹⁰³).

Bien que moins précises que les méthodes de docking flexibles, leur rapidité et leur relative efficacité les rendent intéressantes pour les protocoles de criblage hiérarchique. Elles peuvent constituer un premier filtre, éliminant de la chimiothèque les molécules trop grandes ou présentant une mauvaise complémentarité avec le site de liaison¹⁰⁴⁻¹⁰⁶.

I.3.2.2. Docking avec ligand 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 conformations¹⁰⁷. Il existe plusieurs types d'algorithmes de recherche conformationnelle, classifiés par Brooijmans et Kuntz en 2003¹⁰⁸ : recherche systématique (fragmentation/reconstruction), recherche aléatoire (Monte Carlo, algorithmes génétiques), méthodes de simulation (dynamique moléculaire).

I.3.2.2.1. Recherche systématique

Les algorithmes de recherche systématique essaient d'explorer tous les degrés de liberté d'une molécule (translations, rotations, angles de torsion). Mais, très rapidement, se pose le problème de l'explosion combinatoire¹⁰⁹. En effet, le nombre de conformations pour un ligand est fonction du nombre d'angles dièdres de ce ligand et de leur valeur :

$$N_{conf} = \prod_{i=1}^N \prod_{j=1}^{n_{inc}} \frac{360}{\theta_{i,j}} \quad (1)$$

avec N le nombre d'angles dièdres et θ_{ij} la valeur de l'angle dièdre j et i la liaison appartenant à un des plans impliqués¹⁰⁷. Pour palier à ce problème, les méthodes de recherche systématique utilisent donc souvent un algorithme de fragmentation/reconstruction qui consiste à diviser le ligand en plusieurs fragments pour le reconstruire dans le site de liaison^{110; 111}. Par exemple le ligand est divisé en fragments rigides et flexibles. Tous les fragments rigides sont placés dans le site de liaison et les fragments flexibles sont ajoutés de manière incrémentée (FlexX¹¹¹, DOCK 4.0¹¹², Surfex-dock¹¹³).

I.3.2.2.2. Recherche aléatoire

Ces méthodes font des changements aléatoires, sur la conformation ou la position (translation/rotation) d'un ligand ou d'une population de ligands. Le nouveau ligand est ensuite évalué par une fonction de score. Deux approches populaires sont les algorithmes de Monte Carlo et génétiques.

– L'algorithme de Monte Carlo génère une conformation initiale d'un ligand, dans le site de liaison, de manière aléatoire. Cette conformation est évaluée par une fonction de score et devient la conformation de référence. Une nouvelle conformation est ensuite générée et évaluée. Si cette conformation a un score plus favorable que la précédente, elle devient la nouvelle conformation de référence. Dans le cas contraire, elle peut devenir la conformation de référence selon un critère d'acceptation, par exemple le critère de Métropolis défini par la probabilité de Boltzmann : $P = e^{(E_{new} - E_{old})/kT}$, avec E_{new} et E_{old} les énergies de la nouvelle conformation et de celle de référence, respectivement, k la constante de Boltzmann et T la température de la simulation. Ces étapes sont répétées un certain nombre de fois, ce qui

correspond au nombre de pas de Monte Carlo. La première version d'AutoDock¹¹⁴ et ICM¹¹⁵ sont deux méthodes basées sur l'algorithme de Monte Carlo.

– Les algorithmes génétiques adaptent les principes de compétition biologique et de dynamique des populations. Plusieurs conformations (correspondant à une population d'individus) sont générées, chacune correspondant à un ensemble de variables (par exemple angles dièdres, donneurs de liaisons hydrogène, accepteurs de liaisons hydrogène¹¹⁶), ou chromosomes. Ces conformations sont évaluées par une fonction de score (fonction de survie) et les meilleures sont conservées. Leurs variables sont alors modifiées aléatoirement, soit de manière ponctuelle (mutation), soit en échangeant deux groupes de variables (cross over), formant une nouvelle génération de conformations¹⁰⁷. La méthode GOLD¹¹⁶ et les versions plus récentes d'AutoDock¹¹⁷ utilisent les algorithmes génétiques.

I.3.2.2.3. Dynamique moléculaire

Contrairement aux méthodes stochastiques, la dynamique moléculaire ne peut pas passer les hautes barrières d'énergie pour des simulations d'une durée raisonnable. Ainsi, lorsque le paysage d'énergie potentielle est accidenté, la simulation peut rester prisonnière d'un minimum local d'énergie¹⁰⁸.

Différentes approches ont été développées pour permettre une meilleure exploration du paysage d'énergie potentielle :

- Augmenter la température de la simulation pour la baisser par la suite (recuit simulé). Par exemple : SDOCKER¹¹⁸.
- Manipuler le paysage énergétique pour le rendre plus lisse et rendre possible son exploration complète¹¹⁹
- Coupler les degrés de liberté de la molécule à différentes températures. Pendant la simulation, les températures associées aux degrés de liberté responsables de la flexibilité du ligand sont changées. Cette étape aide l'algorithme à ne pas rester prisonnier d'un minimum local^{120, 121}
- Commencer à partir de plusieurs conformations de départ, chacune aboutissant à un minimum local différent¹⁰⁷

I.3.2.3. Prise en compte de la flexibilité du site de liaison

Peu de méthodes de docking prennent en compte la flexibilité du site de liaison. Pourtant, celle-ci joue un rôle majeur dans la formation du complexe protéine/ligand, les chaînes latérales des résidus du site de liaison changeant de conformation pour s'adapter au ligand (induced fit)¹²². Plusieurs approches ont été mises au point pour prendre en compte la flexibilité d'au moins une partie de la cible¹²³ : 1. les algorithmes de Monte Carlo^{117; 124}, 2. les bibliothèques de rotamères^{125; 126}, 3. les grilles d'ensembles de protéines¹²⁷.

Les bibliothèques de rotamères modélisent l'espace conformationnel de la protéine sur la base des conformations expérimentales des chaînes latérales les plus fréquemment observées¹²⁵. Elles sont couplées à un algorithme d'élimination des impasses (dead-end elimination) qui élimine les conformations des chaînes latérales qui n'aboutissent pas à un minimum d'énergie¹²⁶. Ceci permet de limiter le nombre de conformations prises en compte.

Les grilles d'ensembles de protéines peuvent aussi être utilisées pour prendre en compte la flexibilité du site. Ces ensembles peuvent être représentés sur une grille de différentes façons. La première consiste à générer une grille d'énergie potentielle moyenne des conformations du récepteur, par exemple utilisée avec DOCK 3.5¹²⁷. Une autre possibilité consiste à associer à chaque nœud de la grille les énergies potentielles de plusieurs conformations. Un score est ensuite attribué aux ligands par rapport à chaque jeu de potentiels énergétiques du récepteur¹²⁸.

I.3.2.4. Fonctions de score

Dans un solvant, la liaison d'une petite molécule avec sa cible est conduite par une combinaison d'effets enthalpiques et entropiques sur la petite molécule, la cible et les molécules du solvant. Tous ces paramètres doivent être pris en compte pour s'approcher le plus possible de l'énergie libre de liaison¹⁰⁷.

Les bases de la reconnaissance moléculaire ont été établies par Fischer en 1894 dans son modèle clé/serrure¹²⁹. Dans ce modèle, le ligand se fixe à sa cible si leurs formes sont complémentaires. En 1940, Pauling et Delbrück affinent ce modèle en suggérant la stabilisation des complexes moléculaires par des interactions de type Van der Waals,

électrostatiques et par la formation de liaisons hydrogène¹³⁰. En 1959, Kauzmann démontre le rôle des interactions hydrophobes dans ces interactions¹³¹.

Avant la liaison, les solutés sont en interaction avec le solvant. Lors de la formation du complexe, il y a désolvatation des parties des deux molécules qui seront en contact. Un coût élevé en énergie est nécessaire à la désolvatation des parties polaires et chargées de ces molécules. Ce coût est en partie contrebalancé par les interactions électrostatiques et les liaisons hydrogène formées dans le complexe, mais pas en totalité car elles sont moins favorables que les interactions natives faites avec les molécules d'eau. La désolvatation des parties apolaires libère les molécules d'eau, ce qui augmente l'entropie du solvant. Ce phénomène est appelé effet hydrophobe et est considéré comme la force majeure qui stabilise les complexes. En revanche, les interactions électrostatiques et les liaisons hydrogène seraient plutôt responsables de la spécificité des interactions entre les solutés. Les rôles de ces interactions (hydrophobes et électrostatiques) ont été décrits par Davis et al en 1998¹³². Pendant que l'entropie du solvant augmente, l'entropie des solutés diminue, en raison des contraintes imposées par la formation du complexe et de la perte de degrés de liberté.

Il faut donc aussi prendre en compte l'énergie interne des deux solutés, qui peut être plus élevée lorsqu'ils sont dans un complexe qu'à l'état non lié dans le solvant. Cette énergie déstabilise le complexe.

La stabilité du complexe dépend de la constante d'équilibre K_{eq} , qui représente le rapport entre la constante d'association k_a et la constante de dissociation k_d . La constante d'équilibre est directement impliquée dans l'énergie libre de liaison (ΔG_{bind}) :

$$\Delta G_{bind} = -RT \ln K_{eq} = -RT \ln\left(\frac{k_a}{k_d}\right) \quad (2)$$

L'énergie libre de liaison est aussi la somme des contributions entropiques et enthalpiques du complexe et des deux solutés :

$$\Delta G_{bind} = \Delta H - T \Delta S \quad (3)$$

Cette énergie est dépendante de nombreuses conditions : température, pression, force ionique, pH, concentration des solutés. Si des différences existent dans ces conditions entre deux

expérimentations, les résultats seront difficilement comparables. L'énergie libre de liaison est donc toujours déterminée par rapport aux conditions standards (concentration 1M, température 298K).

Bien qu'il soit possible de calculer directement l'énergie libre de liaison, ces méthodes sont trop lourdes pour des méthodes basées sur la structure de la cible. Une estimation est donc faite par l'évaluation des contributions entropiques et enthalpiques à la stabilité du complexe. Les énergies libres de chaque soluté et du complexe sont calculées séparément avant d'en faire la différence :

$$\Delta G_{bind} = \Delta H - T \Delta S = G_{complexe} - (G_{récepteur} + G_{ligand}) \quad (4)$$

Pour évaluer l'énergie libre d'un système, les approches les plus rigoureuses sont les méthodes de perturbation de l'énergie libre et d'intégration thermodynamique¹³³. Dans les deux cas, elles passent par une simulation d'une très lente évolution du système d'un état vers un autre, par exemple, en rapprochant le ligand de sa cible lors d'une très longue simulation de dynamique moléculaire¹³⁴. D'autres méthodes utilisent les champs de force de mécanique moléculaire afin d'estimer la contribution enthalpique à l'énergie de liaison. Les effets entropiques sont calculés séparément pour le solvant et les solutés.

Les champs de force de mécanique moléculaire sont généralement de la forme suivante :

$$E_{MM} = \sum_{bonds} K_r (r - r_{ref})^2 + \sum_{angles} K_\vartheta (\vartheta - \vartheta_{ref})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)]^2 + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (5)$$

Les trois premiers termes de l'équation 5 pénalisent les écarts de longueur et d'angle des liaisons par rapport aux valeurs de référence. Le dernier terme estime les contributions enthalpiques à l'énergie potentielle du système entre les atomes du ligand et ceux du récepteur par le potentiel de Lennard-Jones, pour les interactions de Van der Waals, et le potentiel de Coulomb, pour la force des interactions électrostatiques.

L'énergie libre de solvation est divisée en deux termes : un terme apolaire et un terme polaire. Le terme apolaire représente le coût énergétique associé à la création d'une cavité pouvant accueillir le soluté et l'énergie des interactions de Van der Waals entre le solvant et le

soluté. Ce terme intervient dans la stabilité du complexe, la différence entre le complexe et ses composants étant souvent favorable. Le terme polaire correspond à la différence d'énergie électrostatique entre le soluté dans un solvant peu polarisable et le soluté dans un solvant fortement polarisable. L'effet général de solvatation sur la stabilité du complexe est évalué, mais ne rend pas compte des interactions spécifiques que peuvent faire les molécules d'eau avec le soluté.

Ces forces physiques permettent d'écrire l'équation de l'énergie libre d'interaction suivante¹⁰⁸ :

$$\Delta G_{bind} = E_{MM} - T \Delta S_{soluté} + \Delta G_{solvant} \quad (6)$$

L'évaluation des poses, aboutissant à leur classement, prédites par les méthodes de docking est une étape cruciale des protocoles de criblage^{107; 108}. Même lorsque le bon mode de liaison est prédit lors de la première étape du docking, la fonction de score doit ensuite différencier les poses correctes des poses incorrectes. La pose ayant le meilleur score est choisie pour représenter chaque molécule et les « vrais » ligands doivent avoir un meilleur score que les autres composés. Les méthodes calculant l'énergie libre de liaison le plus rigoureusement sont trop coûteuses en calcul pour le criblage de grandes chimiothèques. Les fonctions de scores utilisées par les méthodes de docking font des simplifications dans l'évaluation des complexes et ne prennent pas nécessairement en compte tous les paramètres de la reconnaissance moléculaire. Trois classes de fonctions de score peuvent être décrites : 1. les fonctions empiriques, 2. les fonctions basées sur les champs de force de mécanique moléculaire et 3. les fonctions « knowledge-based ».

I.3.2.4.1. Fonctions empiriques

Les données expérimentales d'affinité et de contacts non covalents entre un ligand et son récepteur sont utilisées pour mettre au point les termes pondérés de ces fonctions de score, dans le but d'évaluer au mieux l'énergie libre du complexe, comme décrit pour la première fois par Böhm avec LUDI^{135; 136}. Un problème de ces fonctions est leur dépendance aux jeux de données moléculaires utilisés pour leur construction¹⁰⁷. Plusieurs fonctions de score empiriques, comme LUDI, ChemScore ou Surfex⁶⁷, contiennent des termes prenant en compte le nombre de liaisons rotatives pour approximer les pénalités dues à la perte d'entropie lors de la fixation. D'autres fonctions de score, telles que Fresno¹³⁷, mais aussi Surfex,

contiennent aussi des termes pour évaluer la désolvatation de ligand et du récepteur lors de la formation du complexe.

I.3.2.4.2. Fonctions basées sur les champs de force

Un champ de force de mécanique moléculaire exprime l'énergie d'un système en tant que somme de différents termes, tels que les écarts par rapport aux valeurs de références pour les termes liés et les interactions hydrophobes et électrostatiques entre une protéine et son ligand pour les termes non liés. G-Score est basé sur le champ de force Tripos^{116, 138} et AutoDock¹¹⁷ sur AMBER¹³⁹.

Les interactions entre le ligand et le récepteur sont souvent représentées par les énergies de Van der Waals et électrostatiques (équation 7). L'énergie de Van der Waals est estimée par un potentiel de Lennard-Jones. Les paramètres de ce potentiel varient selon sa « dureté ». D-Score¹⁴⁰ utilise un potentiel standard (12-6), tandis que G-Score utilise un potentiel plus atténué (8-4)¹¹⁶. les termes électrostatiques sont estimés par un potentiel coulombien.

$$\Delta G_{bind} = \sum_{i=1}^{ligand} \sum_{j=1}^{protéine} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (7)$$

Avec A et B fonction du type d'atomes en interaction, R la distance entre les centres des atomes, q la charge partielle associée à chaque atome et ϵ la constante diélectrique du milieu. La forme fonctionnelle de l'énergie interne du ligand est similaire à celle de l'énergie d'interaction protéine/ligand, incluant aussi les énergies de Van der Waals et électrostatiques¹⁰⁷.

Les fonctions standards basées sur les champs de force ont été formulées originellement pour modéliser les contributions enthalpiques sur la structure et les énergies. Elles ne contenaient donc pas de termes entropiques et de solvatation¹⁰⁷. Des fonctions avancées ont ensuite intégré des termes empiriques d'entropie pour les angles dièdres (G-Score¹¹⁶) ou des termes empiriques évaluant de manière explicite l'énergie des liaisons hydrogène (GOLD¹⁴¹ et AutoDock¹¹⁷). Ce dernier terme permet d'augmenter le poids accordé à la spécificité des interactions ligand/protéine¹⁰⁷.

I.3.2.4.3. Fonctions « knowledge-based »

Ces fonctions sont construites sous l'hypothèse qu'un complexe cristallographique observé représente le placement optimal des atomes du ligand, relativement aux atomes de la protéine. Ainsi, le calcul des potentiels des paires d'atomes ligand / protéine se fait à partir des données structurales uniquement. Par exemple, les potentiels de force moyenne (PMF)¹⁴² sont construits à partir des complexes de la PDB et les potentiels statistiques sont dérivés par types similaires d'atomes. SMOG¹⁴³ et DrugScore¹⁴⁴ sont d'autres fonctions « knowledge-based » populaires.

I.4. Evaluation de méthodes

I.4.1 Précision du positionnement

Le criblage *in silico* est un domaine en pleine expansion et le nombre de méthodes disponibles, depuis DOCK¹⁰¹, augmente également¹⁴⁵⁻¹⁴⁸. Les performances de ces méthodes dépendent en grande partie des complexes ligand / récepteur utilisés pour les mettre au point. Par exemple, les fonctions de score empiriques ne fonctionnent bien que si les molécules font des interactions similaires à ce qui est observé dans les complexes utilisés en apprentissage¹⁰⁸. Il est donc nécessaire de les évaluer pour savoir quelles méthodes utiliser en fonction de l'objet de l'étude.

L'évaluation ultime est de prédire de manière prospective de nouveaux ligands. C'est à dire d'évaluer la capacité d'une méthode de docking à prédire l'affinité de liaison d'un ligand pour son récepteur¹⁴⁹. Les méthodes actuelles ne le permettant pas, l'évaluation se fait de manière rétrospective¹⁴⁹. Deux critères sont étudiés : 1. la capacité de la méthode à retrouver le mode de liaison correct du ligand, 2. la capacité à discriminer les actifs des inactifs.

Leur capacité à retrouver le mode de liaison expérimental est généralement mesuré par l'écart quadratique moyen (root mean square deviation : RMSD), observé entre la structure expérimentale du ligand et la pose prédite¹⁰⁸. Le RMSD mesure la distance entre les paires d'atomes correspondants pour deux conformations différentes. Ces deux conformations sont considérées identiques lorsqu'elles sont séparées par un RMSD inférieur ou égal à 0,3 Å¹⁵⁰ et la pose n'est pas considérée comme correcte lorsqu'elles sont séparées par un RMSD supérieur à 2 Å^{65; 151; 152}.

Le RMSD est utilisé en raison de son objectivité et sa simplicité d'automatisation, mais ne reflète pas toujours la réalité de manière satisfaisante¹⁵³. Selon Kroemer et al, il ne rend pas compte des interactions faites entre le ligand et le récepteur, ce qui entraîne :

– une trop grande sensibilité aux variations des poses prédites, même lorsque le mode de liaison n'est pas différent

- une augmentation du RMSD alors que le mode de liaison reste correct, si une partie de la molécule, ne faisant pas d'interactions clés avec le récepteur, est très flexible
- une augmentation du RMSD avec un mode de liaison presque correct, lorsqu'une grande molécule est presque symétrique

Un RMSD plus faible est également plus facilement observé pour des petites molécules¹⁵⁴.

D'autres métriques sont disponibles, telles que l'erreur de déplacement relative (RDE¹⁵⁵) ou la classification basée sur la précision des interactions (IBAC¹⁵³).

RDE permet de réduire l'impact des grandes divergences sur sa valeur moyenne mais les autres problèmes du RMSD restent présents¹⁵⁶. IBAC est un critère dépendant du mode de liaison inspecté. Le ligand est divisé en deux parties majeures : 1. un noyau comprenant les caractéristiques essentielles à sa liaison avec la protéine, 2. une aire périphérique moins importante pour la liaison. Si le noyau est représenté précisément mais que l'aire périphérique diffère modérément, la pose est considérée comme étant correcte¹⁵³.

I.4.2 Enrichissement d'une chimiothèque

Le deuxième critère d'évaluation est la capacité d'une fonction de score à discriminer les actifs des inactifs. C'est essentiel car le criblage *in silico* doit permettre de tester expérimentalement certaines molécules (parmi celles ayant les meilleurs scores) et d'en laisser d'autres de côté (les moins bien classées). Plusieurs banques d'évaluation ont été mises en places pour évaluer ce critère. Bissantz et al¹⁵⁷ ont constitué leur banque avec 10 actifs sur la thymidine kinase et 10 actifs sur le récepteur aux oestrogènes, dans sa conformation antagoniste. 990 molécules choisies aléatoirement ont servi d'inactifs supposés. Par la suite, Pham et Jain¹⁵⁸ ont agrandi la banque précédente en y ajoutant 1000 composés drug-like et 252 actifs pour 27 cibles. Les 100000 molécules de la MDL Drug Data Report (MDDR²⁵) ont aussi été utilisées comme source d'actifs et d'inactifs¹⁵⁹. Les inactifs n'ayant pas été sélectionnés en fonction de leurs similarité physiques avec les actifs, la discrimination peut être faite sur ce simple critère. Ce biais est moins important pour les banques utilisant la MDDR, mais celui-ci reste tout de même significatif¹⁶⁰. Huang et al¹⁶⁰ ont créé la DUD (Directory of Useful Decoys) comprenant 40 systèmes. Chaque système est composé de la structure d'une protéine cible, d'actifs (entre

12 et 416) sur cette cible et d'inactifs, 36 par actifs. Ces inactifs ont été sélectionnés de sorte qu'ils aient des propriétés physicochimiques similaires aux actifs, tout en étant structuralement différents. La DUD a depuis été mise à jour (DUD-E¹⁶¹) et comprend maintenant 102 protéines et des inactifs définis expérimentalement, en plus des inactifs supposés.

La majorité des descripteurs de l'enrichissement sont basés sur deux valeurs : 1. la sensibilité (équation 8), 2. la spécificité (équation 9). Si les n premières molécules d'une chimiothèque, classée par une fonction de score, sont sélectionnées, tous les composés peuvent être classés de cette manière (**figure 5**) :

- les actifs appartenant aux n molécules sélectionnées : vrais positifs (VP)
- les actifs non sélectionnés : faux négatifs (FN)
- les inactifs sélectionnés : faux positifs (FP)
- les inactifs non sélectionnés : vrais négatifs (VN)

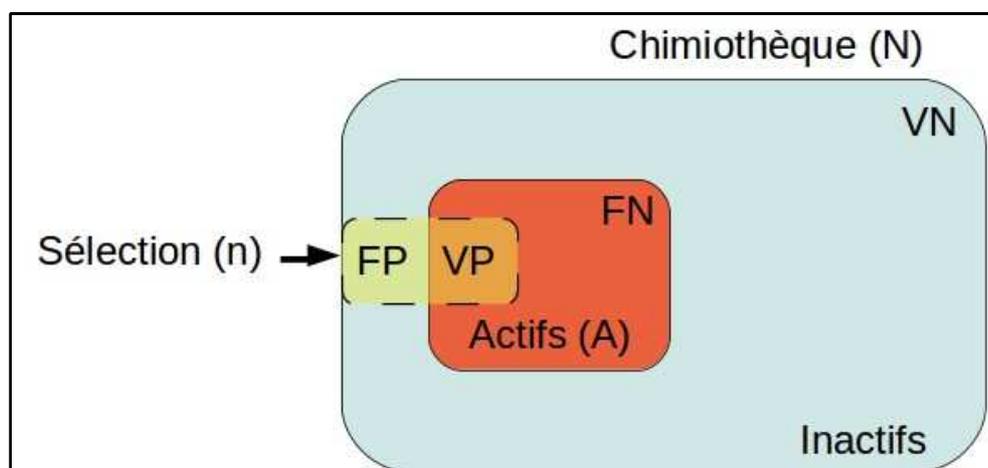


Figure 5: Représentation de la sélection de n molécules dans une chimiothèque de N composés¹⁵⁶

La sensibilité (Se) décrit le ratio du nombre d'actifs retrouvés contre le nombre d'actifs présents dans la chimiothèque¹⁶².

$$Se = \frac{N_{\text{actifs sélectionnés}}}{N_{\text{actifs totaux}}} = \frac{VP}{VP + FN} \quad (8)$$

La spécificité (Sp) décrit le ratio du nombre d'inactifs non sélectionnés contre le nombre d'inactifs présents dans la chimiothèque¹⁶².

$$Sp = \frac{N_{\text{inactifs non sélectionnés}}}{N_{\text{inactifs totaux}}} = \frac{VN}{VN + FP} \quad (9)$$

Après un protocole de criblage in silico, une fraction allant de 0,1 à 10% des composés seront testés expérimentalement¹⁶³. Une méthode doit donc non seulement donner un bon score aux actifs, mais doit aussi leur donner un meilleur score qu'aux inactifs, et le descripteur de l'enrichissement doit tenir compte de l'identification précoce des actifs¹⁶⁴.

		Activité réelle		
		Actifs	Inactifs	
Activité prédite	Positifs	VP	FP	$VPP = \frac{VP}{VP + FP}$
	Négatifs	FN	VN	$VPN = \frac{VN}{FN + VN}$
		$Se = \frac{VP}{VP + FN}$	$Sp = \frac{VN}{FP + VN}$	

Figure 6: Relations entre les vrais positifs (VP), les faux positifs (FP), les vrais négatifs (VN) et les faux négatifs (FN). Ces valeurs représentent la concordance entre les résultats attendus et les résultats observés pour un seuil donné. Elles permettent de calculer la sensibilité (Se), la spécificité (Sp), la valeur prédictive positive (VPP) et la valeur prédictive négative (VPN).

I.4.2.1. Facteur d'enrichissement

Le facteur d'enrichissement (FE) rend compte de l'amélioration du taux d'actifs dans le sous ensemble de composés sélectionnés par rapport à une sélection aléatoire¹⁶⁵.

$$FE = \frac{VP}{n} \times \frac{N}{A} \quad (10)$$

Il est encore fréquemment utilisé¹⁶⁶⁻¹⁶⁸, bien que controversé^{162; 164}. Un des inconvénients décrits par ces deux études est sa grande dépendance au ratio d'actifs présents dans la chimiothèque. Il est donc important d'utiliser cette métrique lorsque deux méthodes sont comparées sur une même chimiothèque (ou deux chimiothèques ayant un même ratio d'actifs). Un autre problème est la contribution égale de chaque actif à la valeur du FE. Ainsi, si deux méthodes (A et B) retrouvent le même nombre d'actifs dans les n premiers composés de la banque, mais avec des distributions différentes (**voir figure 7**), elles sont considérées comme aussi efficaces, alors que la méthode A est plus efficace que la méthode B. Cette métrique ne permet donc pas de prendre en compte l'identification précoce des actifs.

Le facteur d'enrichissement donne un aperçu ponctuel de la performance d'une méthode. Les courbes d'enrichissement représentent le pourcentage d'actifs retrouvés pour n'importe quel pourcentage de la banque, classée par une fonction de score. Elles permettent ainsi d'avoir une vue globale de la performance de la méthode de criblage¹⁶⁴.

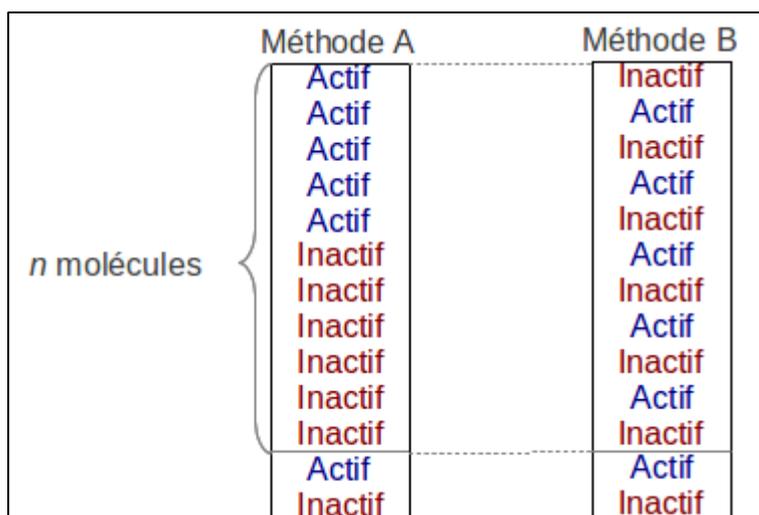


Figure 7: Exemples de distribution des actifs dans les n premières molécules d'une banque classée à partir d'une fonction de score

I.4.2.2. Receiver Operating Characteristic (ROC) curve

La courbe de ROC représente l'évolution de la sensibilité (Se) pour n'importe quel changement possible de $(1-Sp)$. Elle peut aussi être exprimée comme étant le taux de vrais positifs pour n'importe quelle valeur du taux de faux positifs. Ainsi, pour une distribution idéale, sans chevauchement entre les scores des actifs et des inactifs, la sensibilité et la

spécificité prennent toutes les deux la valeur de 1. Ensuite, seuls les inactifs peuvent être retrouvés et la courbe suit une ligne horizontale, la spécificité prenant la valeur de 0¹⁶². La courbe de ROC, représentant une distribution homogène des actifs parmi les inactifs, est la diagonale reliant l'origine et le point représentant une sensibilité de 1 et une spécificité de 0 (courbe aléatoire)¹⁶². Pour les méthodes capables de donner un score moyen supérieur aux actifs, la courbe de ROC se situe entre les courbes idéale et aléatoire.

Si deux courbes de ROC ne se croisent pas, la courbe la plus proche du point ($Se=1$; $Sp=1$) représente une méthode réussissant mieux à discriminer les actifs des inactifs, ce qui permet une comparaison visuelle entre les méthodes¹⁶⁴.

Ce résultat peut aussi être interprété à l'aide de l'aire sous la courbe de ROC (AUC). Pour une distribution idéale, AUC prend la valeur de 1, et 0,5 pour une distribution aléatoire. Pour les criblages fonctionnant mieux que l'aléatoire, AUC prend donc une valeur comprise entre 0,5 et 1. Si l'AUC est inférieure à 0,5, un inactif sélectionné au hasard a plus d'une chance sur deux d'être mieux classé qu'un actif choisi au hasard. Par exemple, pour une AUC égale à 0,3, un inactif choisi au hasard aura une score supérieur à un actif choisi au hasard 7 fois sur 10¹⁶².

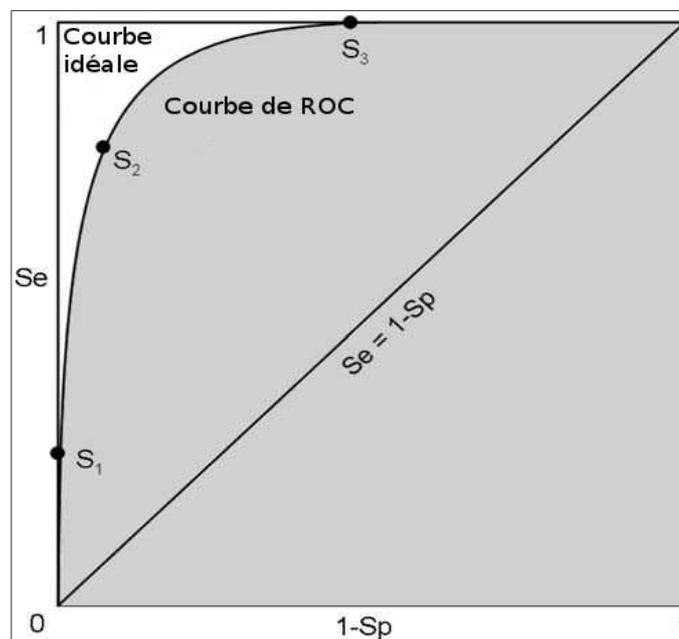


Figure 8: Représentation de 3 courbes de ROC (courbe idéale, courbe d'une distribution homogène et courbe intermédiaire). La courbe de la distribution homogène est donnée par l'équation $Se=1-Sp$ ¹⁵⁶

L'avantage de cette métrique est son indépendance par rapport au ratio d'actifs, contrairement au facteur d'enrichissement. Cependant, l'AUC ne permet pas la discrimination de deux méthodes en fonction de leur capacité à retrouver les actifs au début d'une chimiothèque classée en fonction d'un score. Ainsi, deux distributions différentes d'actifs parmi les inactifs peuvent avoir la même AUC (**figure 9a**). Pourtant, la courbe 1 retrouve plus d'actifs au début de la chimiothèque que la courbe 2¹⁶⁹. Le fait d'avoir une AUC plus grande qu'une autre ne renseigne pas non plus sur ce problème (**figure 9b**).

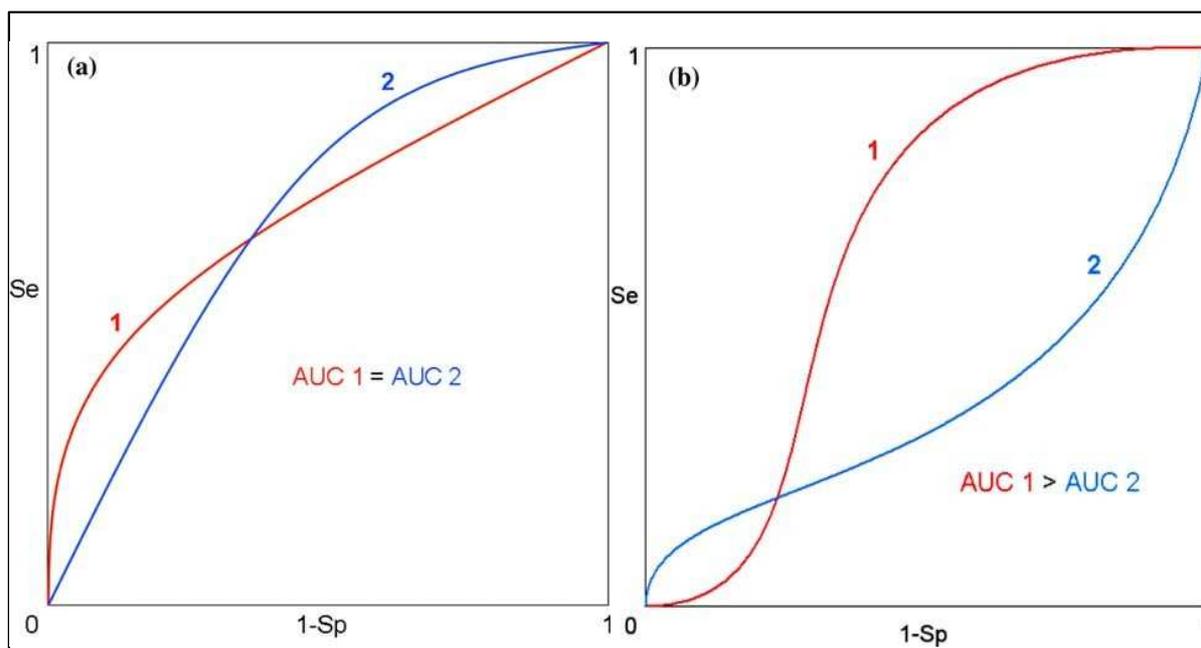


Figure 9: Exemples de ROC AUC non informatives¹⁵⁶. (a) la courbe 1 est meilleure que la courbe 2 pour une spécificité élevée, mais les AUC sont identiques, (b) la courbe 2 est meilleure que la courbe 1 pour une spécificité élevée, mais l'AUC de la courbe 1 est meilleure.

I.4.2.3. Robust Initial Enhancement (RIE)

Cette métrique a été développée par Sheridan et al en 2001¹⁷⁰, dans le but d'avoir une métrique qui ne souffre pas de grandes variations lorsque peu d'actifs sont considérés et qui permet de favoriser les méthodes identifiant de manière précoce les actifs.

La métrique RIE est représentée par la fonction ci-dessous :

$$RIE = \frac{S}{\langle S \rangle} \quad (11)$$

$$\text{avec : } S = \sum_{i=1}^{\text{actives}} \exp\left(\frac{-\text{rank}(i)}{a}\right) \quad (12)$$

S exprime la relation entre le rang du $i^{\text{ème}}$ actif et le nombre de composés sélectionnés. Une fonction exponentielle est utilisée pour avoir un poids de 1 lorsque l'actif est le composé le mieux classé et une pondération décroissante lors du parcours de la banque. $\langle S \rangle$ est la somme moyenne, lorsque les actifs sont classés aléatoirement.

RIE quantifie l'amélioration de la distribution des actifs par rapport à une distribution aléatoire. RIE est supérieur à 1 lorsque la méthode de criblage est capable de donner un meilleur score à plus d'actifs que l'aléatoire. Lorsque la distribution des actifs est aléatoire, RIE est égal à 1. Cependant, comme le facteur d'enrichissement, cette métrique est dépendante du ratio d'actifs. La comparaison entre deux valeurs de RIE ne peut donc être faite que si les chimiothèques possèdent un ratio d'actifs similaires^{164; 170}.

I.4.2.4. Boltzmann-enhanced discrimination of ROC (BEDROC)

Dans le but de combiner les avantages de RIE (identification précoce des actifs) et de l'AUC (valeurs comprises entre 0 et 1, métrique globale), Truchon et al¹⁶⁴ ont développé BEDROC. Cette métrique est dérivée de l'AUC et inclut une fonction de pondération décroissante, donnant ainsi plus de poids aux composés les mieux notés par la fonction de score. Il représente la probabilité pour un actif sélectionné aléatoirement d'être retrouvé avant un composé sélectionné aléatoirement à partir d'une distribution de probabilité hypothétique. BEDROC compare donc la distribution étudiée avec une distribution hypothétique. Lorsque ces deux distributions sont identiques, BEDROC prend la valeur de 0,5 et une valeur de BEDROC entre 0,5 et 1 signifie que la distribution étudiée est meilleure que la distribution de référence¹⁵⁶.

I.5. Objectifs de thèse

Ce travail a été effectué autour de deux axes : l'évaluation de méthodes de criblage *in silico* et leur utilisation dans la recherche de petites molécules à potentiel thérapeutique ciblant le facteur de nécrose tumorale alpha (TNF α).

L'évaluation de méthodes de criblage *in silico* a été fait dans plusieurs buts :

- Mise au point et optimisation de protocoles pour :
 - Connaître les cibles pour lesquelles les méthodes à notre disposition fonctionnent, et quelles sont les caractéristiques de ces cibles
 - Savoir la ou les structures à privilégier, lorsque plusieurs structures d'une même cible sont disponibles, pour obtenir les meilleurs résultats
- Choix des métriques d'évaluations pour connaître celles qui seront les plus adaptées à nos objectifs

Les connaissances ainsi acquises sur les méthodes de criblage *in silico* ont été appliquées pour trouver une molécule potentiellement inhibitrice du TNF α .

Deuxième partie :

Résultats

II.1. Evaluation de méthodes

II.1.1 Evaluation comparative de méthodes 3D de criblage *in silico* : impact de l'alignement moléculaire sur l'enrichissement

Il apparaît que la performance des méthodes de criblage *in silico* dépend de plusieurs facteurs, tels que la sélection de la structure de la référence et sa conformation¹⁷¹, la conformation initiale des composés testés¹⁷² et la qualité de l'alignement produit¹⁷³.

Plusieurs méthodes LBVLS utilisent des algorithmes de recherche conformationnelle partagée avec les méthodes de docking correspondantes, par exemple :

- Surflex-dock¹¹³ / Surflex-sim⁶⁷ : algorithme de fragmentation / reconstruction et fonction de score empirique
- FlexX¹⁷⁴ / FlexS¹⁷⁵ : algorithme de fragmentation / reconstruction et fonction de score empirique
- ICM¹¹⁵ / ICMsim¹⁷⁶ : algorithme de Monte Carlo avec critère de Métropolis et fonction de score empirique
- ROCS / FRED¹⁷⁷ : ligand considéré comme rigide. Algorithme de fragmentation / reconstruction via OMEGA et fonction de score basée sur la complémentarité (stérique pour FRED, stérique et chimique pour ROCS).

Ces méthodes ont été choisies dans le cadre d'une collaboration avec l'équipe du professeur Nilges à Pasteur. De plus, elles permettent d'évaluer différents types d'algorithmes de recherche conformationnelle.

La Directory of Useful Decoys (DUD¹⁶⁰) a été utilisée pour évaluer ces 8 méthodes. Elle a été choisie car elle était, au moment de l'étude, la banque la plus difficile pour les méthodes de docking, du fait de la construction rationnelle de son jeu d'inactifs et au ratio inactifs / actifs élevé¹⁷¹. De plus, les modifications apportées par Good¹⁷⁸ la rendent aussi adaptées à

l'évaluation des méthodes LBVLS. Parmi les 40 systèmes disponibles, 11 ont été choisis : l'adénosine désaminase (ADA), la kinase 2 cycline-dépendante (CDK2), la dihydrofolate réductase (DHFR), la forme antagoniste du récepteur aux oestrogènes (ER), le facteur Xa (FXA), la transcriptase inverse du VIH (HIVRT), la neuraminidase (NA), MAP kinase p38 (P38), la thrombine (THR), la thymidine kinase (TK) et la trypsine (TRP). Ces systèmes ont été choisis car ils étaient représentatifs des différents systèmes présents dans la DUD :

- le nombre d'actifs par système (entre 22 pour TK et 454 pour P38)
- l'ouverture du site de liaison (complètement fermé pour TK et très ouvert pour FXA)
- la diversité des actifs (seulement 6 groupes pour 454 actifs pour P38, 16 groupes pour 43 actifs pour HIVRT), calculée à l'aide d'ICM : calcul de descripteurs puis de la distance de Tanimoto entre les molécules afin de former les groupes.
- la taille et flexibilité des actifs (petits et rigides pour TK, grands et flexibles pour FXA)

De plus, ils apparaissent fréquemment dans les études d'évaluations de méthodes.

Chaque méthode a été évaluée, sur chaque système, selon trois critères : la qualité de l'alignement des actifs testés avec la molécule utilisée comme référence, la discrimination faite entre les actifs et les inactifs, la diversité des actifs ayant les meilleurs scores.

Ne possédant la pose expérimentale que d'un ligand pour évaluer les poses et alignements de tous les actifs, un score a été utilisé. Le calcul du RMSD est la métrique de référence pour évaluer la qualité d'un alignement, mais il n'est pas adapté à l'alignement entre deux molécules différentes. Le score utilisé évalue donc, visuellement, la conservation des interactions (docking) ou des points fonctionnels (similarité 3D) plutôt que la position de tous les atomes. Pour chaque composé, le score est compris entre 0 et 2 :

- 2, lorsque la pose ou l'alignement du composé avec la référence était précis et que les interactions ou les points pharmacophoriques attendus étaient retrouvés
- 1, lorsque l'orientation du composé était similaire à celle de la référence et que 50% des interactions ou points pharmacophoriques étaient retrouvés

– 0, pour les autres composés.

La discrimination entre les actifs et les inactifs a été quantifiée par les facteurs d'enrichissement précoce (1% de la banque classée par une fonction de score) et tardif (10%). Le premier seuil a été choisi car il reflète une réalité expérimentale : seules les composés ayant le meilleur score auront une chance d'être testés expérimentalement. Il est aussi intéressant de savoir si la méthode étudiée retrouve les actifs restants rapidement, ce qui est en partie évalué par l'enrichissement tardif. Dans le cas contraire, une inspection visuelle plus approfondie a été faite, afin de déterminer les causes de ce mauvais classement. Les courbes d'enrichissement ont été tracées, dans le but d'avoir une représentation globale et intuitive de la performance de chaque méthode. Les courbes de ROC ont également été tracées et les AUC correspondantes calculées.

L'étude de la diversité structurale des actifs ayant les meilleurs scores permet de savoir si les méthodes LBVLS et SBVLS peuvent s'émanciper du chénotype de la molécule utilisée comme référence. Pour chaque système, les actifs ont été regroupés en fonction de leur similarité en utilisant des descripteurs d'empreintes chimiques et la distance de similarité de Tanimoto à 0,55. Les courbes d'enrichissement et de ROC ont été tracées en utilisant l'actif le mieux classé de chaque groupe comme représentant de celui-ci.

Comparative Evaluation of 3D Virtual Ligand Screening Methods: Impact of the Molecular Alignment on Enrichment

David Giganti,^{†,‡} H el ene Guillemain,^{†,§} Jean-Louis Spadoni,[§] Michael Nilges,[‡]
Jean-Fran ois Zagury,[§] and Matthieu Montes^{*,§}

Unit e de Bioinformatique Structurale, Institut Pasteur, 26 rue du Dr Roux, 75015 Paris, France, and Chaire de Bioinformatique, Conservatoire National des Arts et M etiers, 292 rue Saint Martin, 75003 Paris, France

Received December 26, 2009

In the early stage of drug discovery programs, when the structure of a complex involving a target and a small molecule is available, structure-based virtual ligand screening methods are generally preferred. However, ligand-based strategies like shape-similarity search methods can also be applied. Shape-similarity search methods consist in exploring a pseudo-binding-site derived from the known small molecule used as a reference. Several of these methods use conformational sampling algorithms which are also shared by corresponding docking methods: for example Surflex-dock/Surflex-sim, FlexX/FlexS, ICM, and OMEGA-FRED/OMEGA-ROCS. Using 11 systems issued from the challenging “own” subsets of the Directory of Useful Decoys (DUD-own), we evaluated and compared the performance of the above-cited programs in terms of molecular alignment accuracy, enrichment in active compounds, and enrichment in different chemotypes (scaffold-hopping). Since molecular alignment is a crucial aspect of performance for the different methods, we have assessed its impact on enrichment. We have also illustrated the paradox of retrieving active compounds with good scores even if they are inaccurately positioned. Finally, we have highlighted possible positive aspects of using shape-based approaches in drug-discovery protocols when the structure of the target in complex with a small molecule is known.

INTRODUCTION

In the early stage of research of drug discovery programs, high-throughput screening (HTS) procedures can be applied for hit identification in large small molecule databases. In the past decade, *in silico* screening has been extensively used to reduce the number of compounds going into HTS, reducing time and costs for hit finding.¹

The classical straightforward concept aiming at identifying analogues by comparing the physicochemical, structural, or pharmacophoric properties of a known active molecule with that of compounds in a collection has been massively applied during the last decades.^{2,3} Initially, these ligand-based virtual ligand screening (LBVLS) methods were based on simple 2D descriptors or fingerprints⁴ derived from the structure of the reference active compound and compared to the corresponding descriptors of database compounds using a similarity metric, such as the Tanimoto coefficient (Tc). These methods were generally efficient, very fast, and provided as a result hits sharing a common chemotype with the active molecule used as the reference.⁵

To increase the structural diversity of the hits provided by LBVLS methods and thus to perform “scaffold-hopping” (i.e., change the chemotype, keep the activity⁶), different methods using more sophisticated 3D descriptors have later been developed, such as pharmacophore screening⁷ or shape similarity searching.^{8,9}

In pharmacophore screening, the knowledge of a set of aligned known active compounds is required, in contrast to shape similarity search methods that only require the structure of a single active compound. Shape similarity search methods thus appear as the LBVLS methods of choice when the structure of only a single active compound is available.

When the structure of the target in complex with a ligand is available, structure-based virtual ligand screening (SBVLS) methods like docking/scoring¹⁰ or structure-based pharmacophore screening^{11,12} are generally preferred. However, the coordinates of a bound ligand can also be used as a reference for shape similarity search methods. Using such methods, virtual screening success stories have recently been reported.^{13–16}

It appears that the performance of virtual ligand screening methods depends on many factors, such as the selection of the query structure and its conformation,¹⁷ the conformational search of the compounds,¹⁸ and the quality of the alignment produced.¹⁹

Recently published shape/volume similarity searches, that is, 3D-ligand-based virtual ligand screening (3D-LBVLS) methods have been developed by the same research groups that had previously developed docking-based virtual ligand screening (docking-BVLS) methods such as Surflex-Sim/Surflex-dock, the ICM package, FlexX-FlexS, and OMEGA-FRED/OMEGA-ROCS. All of these 3D-LBVLS methods share their conformation generator with their respective docking-BVLS method “counterpart”.

We thus decided in the present work to evaluate these four different 3D-LBVLS methods and their docking-BVLS methods counterparts. First, we wished to assess the ability of these four 3D-LBVLS methods and their docking-BVLS

* To whom correspondence should be addressed: Matthieu Montes, PhD, e-mail: matthieu.montes@cnam.fr.

[†] These authors contributed equally to this work.

[‡] Unit e de Bioinformatique Structurale, Institut Pasteur.

[§] Chaire de Bioinformatique, Conservatoire National des Arts et M etiers.

methods counterparts to perform accurate molecular alignments using a large number of active compounds on 11 different targets issued from the DUD data set. Then, we wished to evaluate their performance in virtual screening and scaffold-hopping on the same challenging data sets. Recently, Kirchmair et al.¹⁷ highlighted the importance of the query and its conformation for the quality of enrichment after virtual screening. We decided to proceed further in this direction by exploring the impact of molecular alignment on enrichment. This consisted in evaluating the impact of (1) the quality of the alignment of the query with the structure used as a reference on the enrichment obtained with 3D-LBVLS methods and (2) the docking accuracy on enrichment with docking-BVLS methods. From the results of this study, we expect to outline the most favorable approach to be used when the structure of the target in a complex with a small molecule is available and to illustrate a currently unexplained paradox between the quality of molecular alignment and ranking, that is, enrichment.

MATERIAL AND METHODS

All the 3D-LBVLS methods used in this study have a docking-BVLS counterpart with which they share an important part of the molecular alignment algorithm. For a given docking engine, the algorithm used to position, in a flexible manner, the query molecule into the binding site, according to putative ligand–receptor interaction points, is similar to the algorithm used to superimpose, in a flexible manner, the query molecule on the reference active compound, according to its shape and putative common interaction points. The biggest differences stem from the scoring functions used and from the volume to be explored as a protein binding site is generally wider than the pseudosite defined around a given known ligand. Hence, for each system, we carefully adjusted the binding site definition in all the programs to perform a realistic comparison of the methods in terms of volume to be explored. Keeping this idea of consistence between the different packages, we deliberately chose not to tune the other parameters of the different 3D-VLS programs for the difficult systems, in contrast to other studies.^{20,21}

Computational Methods. *FlexX/FlexS.* FlexX²² and FlexS²³ are a docking/scoring method and a shape-similarity search method, respectively, based on a fragmentation/reconstruction algorithm, to dock compounds flexibly into the binding site (FlexX) or to flexibly superimpose compounds on the reference active compound (FlexS). The query molecule is decomposed into fragments, whose physicochemical properties are represented by a set of spheric (FlexX) or Gaussian (FlexS) functions. The fragment defined as the base fragment is aligned with the reference active compound using the RigFit algorithm²³ (FlexS) or positioned into the binding site via a hashing-technique (FlexX). In FlexX the scoring is made using a modified Bohm empirical scoring function.^{24,25} In FlexS, the scoring is made using a 3D similarity metric.²³ FlexX version 2.0.3 and FlexS 1.20.3 were used for all calculations.

Surflex-dock/Surflex-sim. Surflex-dock²⁶ and Surflex-sim²⁷ use a modified Hammerhead fragmentation/reconstruction algorithm²⁸ to dock compounds flexibly into the binding site (Surflex-dock) and to superimpose flexibly onto the reference

active compound (Surflex-sim), respectively. The query molecule is decomposed into rigid fragments that are superimposed to the Surflex-protomol, that is, molecular fragments covering the entire binding site (Surflex-dock) or to the molecular surface points of the reference active compounds²⁹ (Surflex-sim). In Surflex-dock, the molecules are evaluated by an empirical scoring function. In Surflex-sim, the similarity of each fragment with the reference active compound is evaluated according to the alignment of their respective surface points. In this study, Surflex-dock and Surflex-sim, version 2.2, were used for all calculations.

ICM. ICM is based on Monte Carlo simulations in internal coordinates to optimize the position of molecules using a stochastic global optimization procedure combined with pseudo-Brownian positional/torsional steps and fast local gradient minimization.³⁰ The docking poses were evaluated using ICM-VLS empirical scoring function.³¹ Chemical superposition with ICM-sim were performed using the atomic property fields method (APF³²). ICM, version 3.4, was used for all calculations.

OMEGA/FRED/ROCS. FRED³³ and ROCS use atom-centered Gaussian functions parametrized to provide close approximations to hard sphere volumes. In FRED, version 2.2.3, used in this study, orientations of a single conformer are compared with a bump-grid defined using the active site during the rigid body docking procedure. Poses clashing with the bump-grid are eliminated and the remaining poses are evaluated using ShapeGauss.³⁴ In ROCS,³³ shape-similarity is evaluated by maximizing the volume overlap between the reference active compound and a single conformation of a query molecule using the Tanimoto coefficient. In version 2.3.1, used in this study, a “color force field” represents physicochemical properties in addition to the shape component. The conformational search of the different query compounds (up to 100 conformers per compound) has been carried out prior to all the calculations using OMEGA, version 2.1.³³

A summary of the characteristics of the different programs evaluated in this study is available in Table 1.

DUD Database. To evaluate the ability of the methods to discriminate between active molecules and decoys and thus their enrichment performance, we decided to use the directory of useful decoys database (DUD).²¹ DUD is a benchmarking data set designed for docking method evaluation containing annotated active compounds (from 30 to 120) for 40 targets, including 36 decoys for each active molecule issued from the ZINC database.³⁵ In the DUD, the decoys have been chosen for their similar physicochemical properties (molecular weight, logP, number of hydrogen bond donors/acceptors, etc.) with known active molecules, which renders them harder to discriminate compared to random decoys. To be as challenging as possible, the DUD-own set was selected for each target. Each DUD-own set contains only active and decoy compounds designed for the corresponding target making it more challenging than the DUD-all set.^{36,37} Because of its careful construction and its high rate of decoys (36 decoys for one active compound), it appears today as the most challenging available benchmarking database.¹⁷ Moreover, as stated recently by Hawkins et al.,³⁷ the recent modifications of DUD 2.0 (including new active compounds and structure clustering by Good³⁸) make it more suitable for ligand-based method evaluation.

Table 1. Main Characteristics of the Programs Evaluated in This Study: Flexibility Algorithm, Energy Terms Used in the Scoring Function, and Average Computational Cost in Seconds Per Molecule^a

program	conformational search algorithm	scoring terms	average time per compound (s)
Surflex-dock	fragmentation/incremental construction	steric, hydrogen bonds, hydrophobic, pseudoentropic (rotatable bonds, size)	14.6
Surflex-sim	fragmentation/incremental construction	steric, polar (ionic + hydrogen bonds), hydrophobic	6.7
FRED	via OMEGA (fragmentation/incremental construction)	steric, shape complementarity	1.0
ROCS	via OMEGA (fragmentation/incremental construction)	shape, hydrogen bonds, hydrophobic, aromatic	0.5
FlexX	fragmentation/incremental construction	hydrogen bonds, ionic, hydrophobic, aromatic, pseudoentropic (rotatable bonds)	15.6
FlexS	fragmentation/incremental construction	steric, hydrogen bonds, ionic, hydrophobic	6.9
ICM	metropolis Monte Carlo	van der Waals, electrostatic, hydrophobic, hydrogen bonds desolvation, pseudoentropic (rotatable bond), size correction	17.7
ICMsim	metropolis Monte Carlo	hydrogen bonds, ionic, hydrophobic, aromatic, pseudoentropic (rotatable bonds), atomic fields (electrostatic, size)	2.4

^a All calculations have been performed on Intel Xeon 2.4 GHz processors with 2Go of RAM.

Table 2. Description of the 11 Systems^a

target	PDB	accessible surface (Å ²)	alternate PDB	problematic side chains	no. rot. actives (min-max)	MW actives (min-max) (g/mol)	no. rot. ref	MW ref (g/mol)
ADA	1NDW	477.05	2Z7G	L62, F65	0–8	152.1–316.1	6	256.14
CDK2	1CKP	838.29			1–10	235.1–483.2	2	259.1
DHFR	3DFR	1223.14			1–11	206.1–523.2	10	458.2
ER	3ERT	787.23			5–11	347.2–508.2	8	388.2
FXA	1F0R	872.11	1LPG	Y100, R147	4–12	293.1–574	5	450.1
HIVRT	1RT1	1774.43			1–12	244.1–582	6	302.2
NA	1A4G	1191.96			3–13	194.1–409.2	8	332.2
P38	1KV2	1217.09	1OUK, 1BL6, 1BL7, 2EWA	F160	2–10	239.1–509.2	10	523.3
THR	1BA8	971.39			2–12	250.1–647.2	19	469.2
TK	1KIM	313.44	1QHI, 2KI5	Q125	2–5	184.1–404	2	239.1
TRP	1BJU	756.41			1–12	124.1–620.3	5	283.1

^a The accessible surface of the binding sites for each system has been computed using StrucTools. Minimum and maximum molecular weight (MW), number of rotatable bonds (no. rot) for the set of DUD-own active compounds and DUD-own reference compound are also included.

We downloaded DUD release 2 from the Web site <http://dud.docking.org> and carefully selected 11 out of the 40 targets available (ADA, CDK2, DHFR, ER, FXA, HIVRT, NA, P38, THR, TK, TRP) according to their presence in the literature for benchmarking studies and their diversity in binding site properties and in active compounds according to ICM. Hydrogen atoms were added in the DUD protein structures using ICM.

Performance Metrics. All enrichment graphics were produced with the statistical and graphical tool R (<http://www.r-project.org/>). To complete the information of enrichment graphics, receiver operating characteristics curves were plotted with the ROC package. The area under the ROC curve (AUC) was calculated on the base of the Wilcoxon–Mann–Whitney algorithm.³⁹

Ligand Cluster Definition. We classified consequently the DUD-own active compounds by ligand similarity using chemical descriptor fingerprints and Tanimoto similarity distance (Td) as implemented in ICM. We examined the resulting trees and visually inspected all the DUD-own active compounds to select a harmonized Td cutoff that resulted into at least 2 equilibrated clusters in each DUD-own data set. After multiple tests, Td was defined at 0.55.

RESULTS

Presentation of the 11 Systems. The diversity of the binding site properties of the targets selected for this study is presented in Table 2. Buried or partially buried binding

sites are the most frequent (i.e., ADA, CDK2, DHFR, ER, HIVRT, NA and TK), but 2 proteins display a binding site more accessible to solvent (i.e., FXA and TRP). Some binding sites are essentially hydrophobic (i.e., CDK2, ER, HIVRT and P38), whereas others should display mainly polar interactions with a potential ligand (i.e., ADA, DHFR, NA, TK, and TRP). We also provide the number of active compounds and decoys for each system, which range respectively from 22 for TK to 454 for P38 and from 891 for TK to 9141 for P38.

Positioning: Molecular Alignment and Docking Accuracy. For all 3D-VLS methods, the conformational search is a crucial point for (1) producing a correct alignment of the known active compounds on the reference for 3D-LBVLs methods¹⁹ and (2) producing accurate poses for docking-BVLS methods.⁴⁰

Molecular alignment evaluation is in general performed through cognate ligand docking with one active compound per system, using a database of known protein/small molecule complexes derived from the Protein Data Bank, such as the Fisher set.⁴¹ Since one of the goals of the present work was to evaluate the impact of molecular alignment on enrichment, we decided to use a large number of active compounds issued from the DUD data set. Since very limited structural information was known for all these compounds, we assumed their binding mode to be somewhat similar to the one provided as a reference for each system (considering its scaffold). We thus defined an expert knowledge-based visual

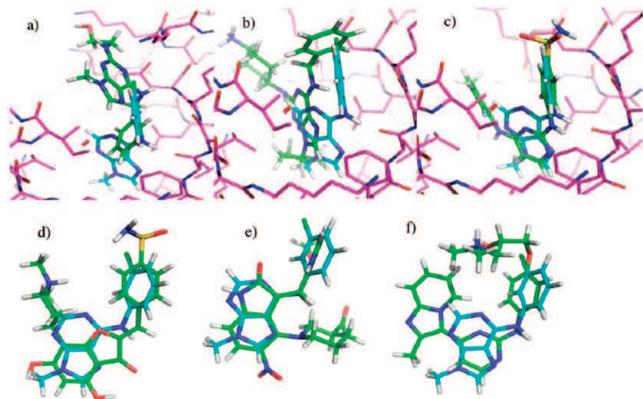


Figure 1. Examples of “0”, “1”, and “2” poses. The quality of molecular alignment with representative CDK2 active compounds is illustrated for docking-BVLS methods in a, b, and c and for 3D-LBVLS methods in d, e, and f. The illustration a represents a “0” pose for the ZINC01641925 compound; b represents a “1” pose for ZINC03591113, and c represents a “2” pose for ZINC03814453. Similarly, d represents a “0” pose for ZINC03814433; e represents a “1” pose for ZINC04617748, and f represents a “2” pose for ZINC03814454.

score for assessing the superimposition accuracy and the docking accuracy of the 3D-VLS methods evaluated in this study. For 3D-LBVLS methods, we used “2” for an accurate superimposition of most of the pharmacophoric interaction points, “1” for an orientation of the compound similar to the reference and at least 50% of well-superimposed pharmacophoric points, and “0” for the others. Similarly, we assessed the docking accuracy using a visual score evaluating the binding mode of the reference active compound provided by the DUD. We used “2” for an accurate pose displaying most of the expected interactions with the residues of the binding site, “1” for an orientation of the compound similar to the reference displaying 50% of the expected interactions with the residues of the binding site, and “0” for the others. We defined acceptable poses as “1 + 2” poses and unacceptable poses as “0” poses. Examples of “0”, “1”, and “2” poses for both 3D-LBVLS and docking-BVLS methods are presented in Figure 1. All compound positioning results are presented in Table 3.

Overall, all the 3D-LBVLS methods are able to superimpose most of the active compounds on the corresponding reference provided by the DUD as the rate of acceptable poses (visual score >0) is above 68% for all the programs. ROCS and FlexS display a very similar performance with 68.46% and 69.25%, respectively, Surflex and ICM-sim being the most efficient with an overall acceptable pose rate of 82.22% and 90.97%, respectively. As expected, the performance of the methods depends on the structure of the active compounds for each target, ADA-compounds displaying a high global rate of unacceptable poses for all the methods (mean rate of 53.21%). Some other systems seem to cause problems to 3D-LBVLS methods such as HIVRT-actives (more than 34% of unacceptable poses for all the methods but Surflex-sim) and FXA-actives (more than 50% of unacceptable poses for all the methods but ICM-sim).

Concerning docking-BVLS methods, similar trends were observed. In general, the methods provided acceptable binding modes, the most efficient being ICM-dock with an overall acceptable pose rate of 82.01%. FRED displays 79.64% of acceptable pose rate (very close to ICM-dock)

and is even the best method on DHFR, ER, and P38. Surflex-dock exhibited the best performance for TK, NA, and TRP despite displaying the least overall docking accuracy with 59.93% of acceptable poses.

As expected, the global performance of all the methods were correct and depended on the structure of the target. Overall, the methods performed accurately on FXA, NA, TK, TRP (except for FRED), and DHFR (except for FlexX). Some targets such as THR (except for ICM-dock), P38, HIVRT, and ER (except for FRED) seemed to cause recurrent problems.

Enrichment in Active Compounds on the DUD-own Databases. Because the 3D-VLS methods evaluated in this study were able to provide accurate molecular alignments, we assessed their performance on enrichment of the DUD-own databases corresponding to the 11 targets we selected for this study using as a reference the compound provided by the DUD. Through the analysis of the corresponding enrichment graphs, we both evaluated their performance in early enrichment (within the top 1% of the database) and in late enrichment (within the top 10% of the database). Results are presented in Figure 2 and Table 4.

Overall, three out of the four 3D-LBVLS methods evaluated in this study (ROCS, FlexS, ICM-sim) display an early enrichment rate of 13%, ROCS displaying the best average performance with 14.12% of active compounds retrieved within the first percent of the corresponding DUD-own database and the best early enrichment rate with 34.69% observed on NA. ICM-sim seems to perform better on late enrichment with 57.25% of actives retrieved in average within the first 10% of the corresponding DUD-own database.

In similarity with the quality of superimposition, we observed a dependence of the enrichment gained by the different methods on the system tested. All the shape similarity search programs evaluated in this study displayed a good performance on ADA, ER, NA, and TK and performed poorly on P38, the best average performance being observed on ER. For THR, FXA, DHFR, and CDK2, only ICM-sim and ROCS for the two latter systems provided acceptable enrichment rates. Concerning the individual performance on each target, ICM-sim appears to be the best method on early and late enrichment on 4 out of the 11 targets (DHFR, FXA, THR, TRP) and displays the best late enrichment rate with 95.92% on TRP. ROCS and FlexS both perform best on 2 systems out of the 11: HIVRT/P38 for FlexS and CDK2/NA for ROCS.

Concerning the docking-BVLS methods, we compared the results with DOCK enrichments published on the same DUD-own databases (the only ones published with these databases to our knowledge). Overall, Surflex-dock displayed the best early and late enrichment rates with 12.36% and 45.36%, respectively, of the active compounds retrieved in the 1% and 10% of the DUD-own database. It appears as the best method for early and late enrichment rates for ADA (7.69%–48.72%), FXA (17.12%–59.59%) and TRP (18.37%–81.6%), followed by ICM-dock, which is the best for DHFR (19.02%–66.83%), ER (17.95%–53.85%), and NA (34.69%–85.71%). FRED and FlexX follow, being the best method respectively for TK (18.18%–22.73%) and THR (6.94%–52.78%). Compared to the methods evaluated herein, DOCK was the best on early enrichment for ADA and DHFR.

Table 3. Molecular Alignment of the DUD-own Active Compounds^a

		ADA (39)	CDK2 (72)	DHFR (410)	ER (39)	FXA (146)	HIVRT (43)	NA (49)	P38 (454)	THR (72)	TK (22)	TRP (49)	mean (1100)
Surflex-sim	2	12.82	12.5	44.39	56.41	4.11	18.6	18.37	9.69	4.17	68.18	40.82	23.15
	1	35.9	51.39	53.66	43.59	33.56	72.09	75.51	69.6	93.06	31.82	59.18	59.07
	0	51.28	36.11	1.95	0	62.33	9.3	6.12	20.7	2.78	0	0	17.78
ROCS	2	12.82	43.06	74.15	41.03	14.38	30.23	79.59	9.47	2.78	86.36	8.16	35.63
	1	20.51	36.11	14.39	56.41	28.77	34.88	14.29	41.19	69.44	9.09	81.63	32.83
	0	66.67	20.83	11.46	2.56	56.85	34.88	6.12	49.34	27.78	4.55	10.2	31.54
FlexS	2	15.38	25	56.1	48.72	35.62	16.28	36.73	14.98	30.56	81.82	18.37	33.48
	1	20.51	19.44	11.71	43.59	13.7	46.51	57.14	74.01	5.56	13.64	2.04	35.77
	0	64.1	55.56	32.2	7.69	50.68	37.21	6.12	11.01	63.89	4.55	79.59	30.75
ICMsim	2	46.15	12.5	86.83	51.28	70.55	18.6	75.51	20.04	88.89	90.91	69.39	54.48
	1	23.08	68.06	11.95	46.15	16.44	46.51	14.29	68.28	9.72	9.09	28.57	36.49
	0	30.77	19.44	1.22	2.56	13.01	34.88	10.2	11.67	1.39	0	2.04	9.03

		ADA (39)	CDK2 (72)	DHFR (410)	ER (39)	FXA (146)	HIVRT (43)	NA (49)	P38 (454)	THR (72)	TK (22)	TRP (49)	mean (1100)
Surflex-dock	2	10.26	23.61	70.98	46.15	25.34	30.23	36.73	6.83	15.28	22.73	61.22	34.05
	1	56.41	45.83	9.02	15.38	57.53	30.23	61.22	18.06	26.39	77.27	36.73	25.88
	0	33.33	30.56	20	38.46	17.12	39.53	2.04	75.11	58.33	0	2.04	40.07
FRED	2	2.56	51.39	79.51	33.33	17.12	41.86	46.94	11.89	15.28	22.73	4.08	36.92
	1	20.51	29.17	8.78	51.28	65.07	32.56	38.78	67.62	40.28	59.09	69.39	42.72
	0	76.92	19.44	11.71	15.38	17.81	25.58	14.29	20.48	44.44	18.18	26.53	20.36
FlexX	2	7.69	18.06	45.61	25.64	16.44	16.28	65.31	9.47	40.28	68.18	36.73	27.31
	1	66.67	81.94	20	41.03	62.33	23.26	30.61	40.31	27.78	18.18	61.22	38.42
	0	25.64	0	34.39	33.33	21.23	60.47	4.08	50.22	31.94	13.64	2.04	34.27
ICM	2	17.95	50	79.51	51.28	28.08	37.21	73.47	12.11	34.72	9.09	24.49	41.29
	1	76.92	41.67	2.93	5.13	57.53	41.86	6.12	65.42	54.17	81.82	71.43	40.72
	0	5.13	8.33	17.56	43.59	14.38	20.93	20.41	22.47	11.11	9.09	4.08	17.99

^a The part of “0”, “1”, and “2” poses are presented in the table as percentage of all the active compounds. Between brackets is the number of active compounds for each system.

Overall, global enrichment was in the favor of 3D-LBVLs methods for 8 out of the 11 systems evaluated in this study. This is particularly striking for ER and TK, where all the 3D-LBVLs methods outperform all the docking-BVLs methods. We observed a mixed performance for both types of methods, with a slightly better outcome of the 3D-LBVLs methods for NA, TRP, HIVRT, and DHFR.

There were exceptions to this statement: for CDK2 and FXA, docking-BVLs methods globally outperformed the 3D-LBVLs methods and for P38, all the methods exhibited a poor enrichment.

Chemotype Enrichment Analysis by 3D-VLS Methods. A critical assessment of the robustness of the methods could be performed by evaluating their ability to retrieve structures with different chemotypes and thus to perform “scaffold-hopping”.^{42,43} Similarly to the work published by Good,³⁸ we thus assessed the structural diversity of the different DUD-own active compounds by assigning them to their respective clusters according to a defined Tanimoto similarity distance descriptor threshold ($T_d < 0.55$).

We defined cluster enrichment via a simple procedure: a cluster is counted as present if at least one of its members is retrieved within the subset sampled. Results are presented on Figure 3 using cluster enrichment graphs. Overall, the enrichment in active compounds and in clusters follows a similar trend for 3D-LBVLs methods. Slight differences can be noticed in the individual performance of the methods for some systems. Surflex-sim performed efficiently in cluster enrichment for CDK2, P38, NA, and TK and was one of the best methods for these systems. Similarly, FlexS displayed a good performance on cluster enrichment for TRP, HIVRT, P38, NA, and FXA. Regarding docking-BVLs methods, their global performance was good for cluster

enrichment. This was particularly striking on ER, FXA, and CDK2, where all docking-BVLs methods displayed a very good performance. By comparing the relative individual performance of docking-BVLs methods, it seems that the global trend was conserved for most of the systems used in this study. For 8 out of 11 systems, a method that displayed the best performance in active compounds enrichment on a defined target also displayed the best performance in cluster enrichment. Striking differences in cluster enrichment could be observed with FlexX that displayed much better performance for DHFR and with ICM that displayed much better performance for THR and TRP. Overall, both 3D-LBVLs and docking-BVLs methods retrieved more than 50% of the clusters in the top 10% for 6 out of the 11 systems.

To highlight global trends between 3D-LBVLs methods and docking-BVLs methods in enrichment, we created mean enrichment graphs using the means of the ranks obtained after screening the DUD-own databases. As the performance of each individual method varies depending on the target, an advantage of such representation is to enhance the differences in the global performance for 3D-LBVLs methods and docking-BVLs methods. Mean enrichment graphs for cluster enrichment ($T_d > 0.55$) and active compounds enrichment are presented on Figure 4.

As expected, global trends in enrichment of combined docking-BVLs methods and combined 3D-LBVLs methods were similar to the trends of individual methods. The global performance of docking-BVLs methods was relatively poor and similar to 3D-LBVLs methods in cluster enrichment for HIVRT and THR. Overall docking-BVLs methods displayed the best performance on 5 out of the 11 targets, it was thus equivalent on this test to 3D-LBVLs methods.

Impact of Molecular Alignment on Enrichment. It is obvious that, at least for 3D-LBVLs methods, there should

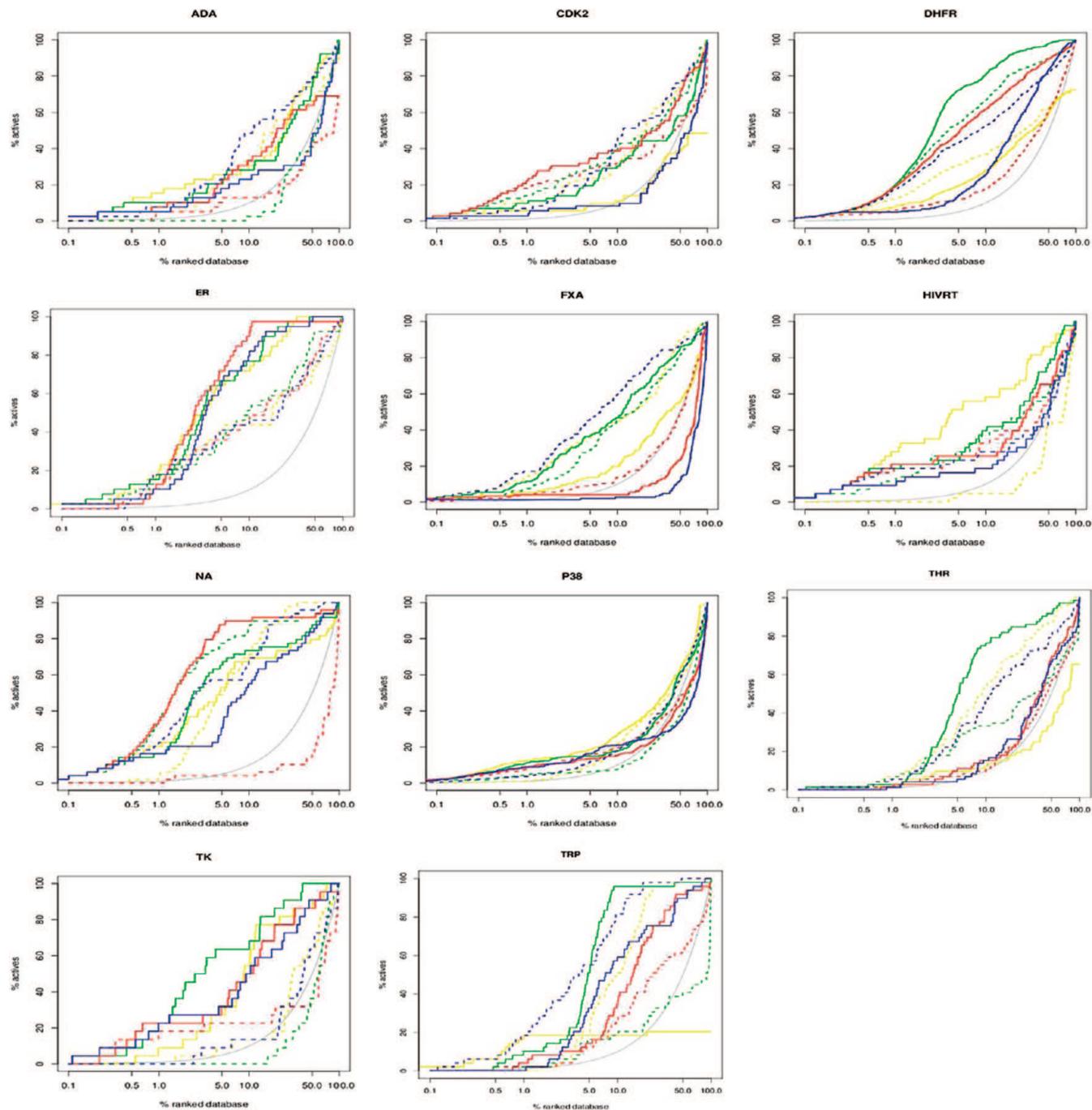


Figure 2. Enrichment graphs with docking-BVLS methods (dotted lines) and 3D-LBVLS methods (plain lines). The gray line represents random enrichment. Each color is for a package: blue for Surflex-dock/Surflex-sim, green for ICM/ICM-sim, yellow for FlexX/FlexS, and red for FRED/ROCS.

be a direct correlation between the quality of the alignment of the active compounds on the reference compound and their rank in the database. To highlight this, we plotted mean enrichment graphs for the acceptable and unacceptable poses of the active compounds of the corresponding DUD-own database provided by the methods evaluated in this study. These mean enrichment graphs are presented in Figure 5.

It clearly appears from the analysis of these graphs that, for 3D-LBVLS methods, compounds that are accurately superimposed to the reference active compound provided by the DUD are present in the early ranks. For 8 out of the 11 systems (all but HIVRT, P38 and NA), more than 95% of the active compounds available in early enrichment (1% threshold) are accurately superimposed (visual score = 2).

At a 10% threshold, most of the active compounds retrieved are superimposed in an acceptable manner (visual score = 1 and 2) but up to 50% of these actives are superimposed with a visual score of 1. This highlights that the less accurate the superposition provided by the methods the lower the score. It also shows that the scoring functions used for the shape-similarity search methods evaluated in this study are efficient and reliable.

Concerning the docking-BVLS methods, at a 1% threshold, more than 95% of the active compounds were accurately positioned in the binding site only for DHFR. For all systems but TK, CDK2, and TRP, a non-negligible part of the ligands that ranked in the top 10% of the corresponding DUD-own database displayed an inaccurate positioning, the most

Table 4. Early (1%) and Late (10%) Enrichments for the DUD-own Active Compounds^a

enrichment all actives	Surflex-sim		ROCS		FlexS		ICMsim		DOCK	
	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%
ADA	5.13	23.08	7.69	33.33	15.38	30.77	10.26	28.21	15	40
CDK2	2.78	8.33	22.22	38.89	5.56	11.11	9.72	30.56	10	25
DHFR	4.88	25.61	19.02	61.95	8.05	27.8	20.24	80.73	25	60
ER	10.26	82.05	10.26	89.74	15.38	71.79	17.95	76.92	10	20
FXA	1.37	2.74	4.11	4.11	5.48	19.86	11.64	46.58	5	35
HIVRT	9.3	18.6	20.93	25.58	27.91	58.14	18.6	39.53	5	25
NA	16.33	55.1	34.69	89.8	20.41	69.39	16.33	73.47	10	60
P38	9.03	21.15	8.81	15.42	12.56	26.21	11.23	17.84	2	25
thrombin	1.39	15.28	0.75	15.28	2.78	12.5	2.78	76.39	5	35
TK	22.73	50	22.73	50	9.09	54.55	22.73	63.64	0	25
trypsin	0	59.18	4.08	34.69	18.37	18.37	10.2	95.92	0	30
mean	7.56	32.83	14.12	41.71	12.82	36.41	13.79	57.25	7.91	34.55
st. dev.	6.99	24.91	10.48	28.79	7.57	22.72	5.85	25.77	7.35	13.87
median	5.13	23.08	10.26	34.69	12.56	27.8	11.64	63.64	5	30

enrichment all actives	Surflex-dock		FRED		FlexX		ICM		DOCK	
	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%	1.00%	10.00%
ADA	7.69	48.72	5.13	12.82	5.13	33.33	0	2.56	15	40
CDK2	6.94	44.44	18.06	31.94	11.11	29.17	13.89	38.89	10	25
DHFR	17.56	52.93	5.85	17.56	11.95	36.83	19.02	66.83	25	60
ER	17.95	43.59	12.82	48.72	15.38	43.59	17.95	53.85	10	20
FXA	17.12	59.59	4.11	13.7	14.38	44.52	7.53	45.21	5	35
HIVRT	18.6	27.91	18.6	32.56	0	4.65	11.63	37.21	5	25
NA	22.45	67.35	0	4.08	4.08	69.39	34.69	85.71	10	60
P38	3.74	18.06	7.93	19.6	1.1	21.59	4.63	7.49	2	25
thrombin	5.56	47.22	1.39	11.11	6.94	52.78	5.56	33.33	5	35
TK	0	13.64	18.18	22.73	0	13.64	0	0	0	25
trypsin	18.37	81.6	2.04	26.53	0	51.02	2.04	18.37	0	30
mean	12.36	45.91	8.56	21.94	6.37	36.41	10.63	35.4	7.91	34.55
st. dev.	7.63	20.32	7.13	12.5	5.96	18.7	10.42	27.16	7.35	13.87
median	17.12	47.22	5.85	19.6	5.13	36.83	7.53	37.21	5	30

^a Corresponding results with DOCK from the DUD original paper (Huang et al. *J. Med. Chem.* 2006) are also presented for information.

striking examples being for THR and ER where 10.45% and 6.45%, respectively, of the active compounds retrieved in the top 10% were docked inaccurately.

Overall, using docking-BVLS, 0.57% of the active compounds retrieved in the top 1% of the ranked database were inaccurately positioned. This rate reached 3.38% when considering the top 10% of the ranked database. Concerning 3D-LBVLS methods, the results were respectively 0.13% for 1% of the ranked database and 1.11% for 10% of the ranked database.

We could also observe that out of the active compounds ranked late in enrichment (from 10% to 100%), 16.15% and 10.82% were positioned accurately by docking-BVLS and 3D-LBVLS, respectively, methods. This highlights that even though 3D-VLS methods provide accurate molecular alignments, a non-negligible part of the active compounds can still be missed because of scoring.

DISCUSSION

Molecular Alignment and Docking Accuracy. From the observation of the results, we can conclude that both 3D-LBVLS and docking-BVLS methods are able to produce conformations similar to the bioactive conformation. Docking-BVLS methods are more challenging since the conformational space of a binding pocket is more complex than the pseudoreceptor defined by the reference ligand in 3D-LBVLS methods. But they display a good performance in terms of docking accuracy even on difficult targets where

artifacts because of the construction of the DUD may sometimes affect the results. In particular, when the protein is considered as rigid (as in all the methods evaluated in this study), it may be very difficult for some ligands to be docked into the binding site of the structure provided by the DUD. For example, in the P38 structure, there are 82 active compounds (i.e., 18% of the P38 active compounds) that cannot be docked into the binding site because of major clashes occurring in the current protein conformation. We retrieved from the PDB database structures corresponding to some of the 82 active compounds that caused problems with the P38 structure, and it seems that there is a major shift of F169 that caused this docking issue (see figure 6). Similar problems can be observed with the structures of FXA, ADA, and TK, where some of the active compounds cannot bind properly to the currently provided structure because of the orientation of some key side chains.

Concerning the relative performance of the methods, it seems that for both docking-BVLS and 3D-LBVLS, ICM displays the best performance in terms of molecular alignment and docking accuracy. This may be caused by a more effective treatment of the flexibility of the compounds in ICM via its biased probability Monte Carlo procedure.⁴⁴ FRED displays an overall very good performance despite being the fastest docking method. It is even the most effective docking method on DHFR, P38, and ER, which pose problems to the other docking methods. These targets display the largest binding sites, making the conformational search within these

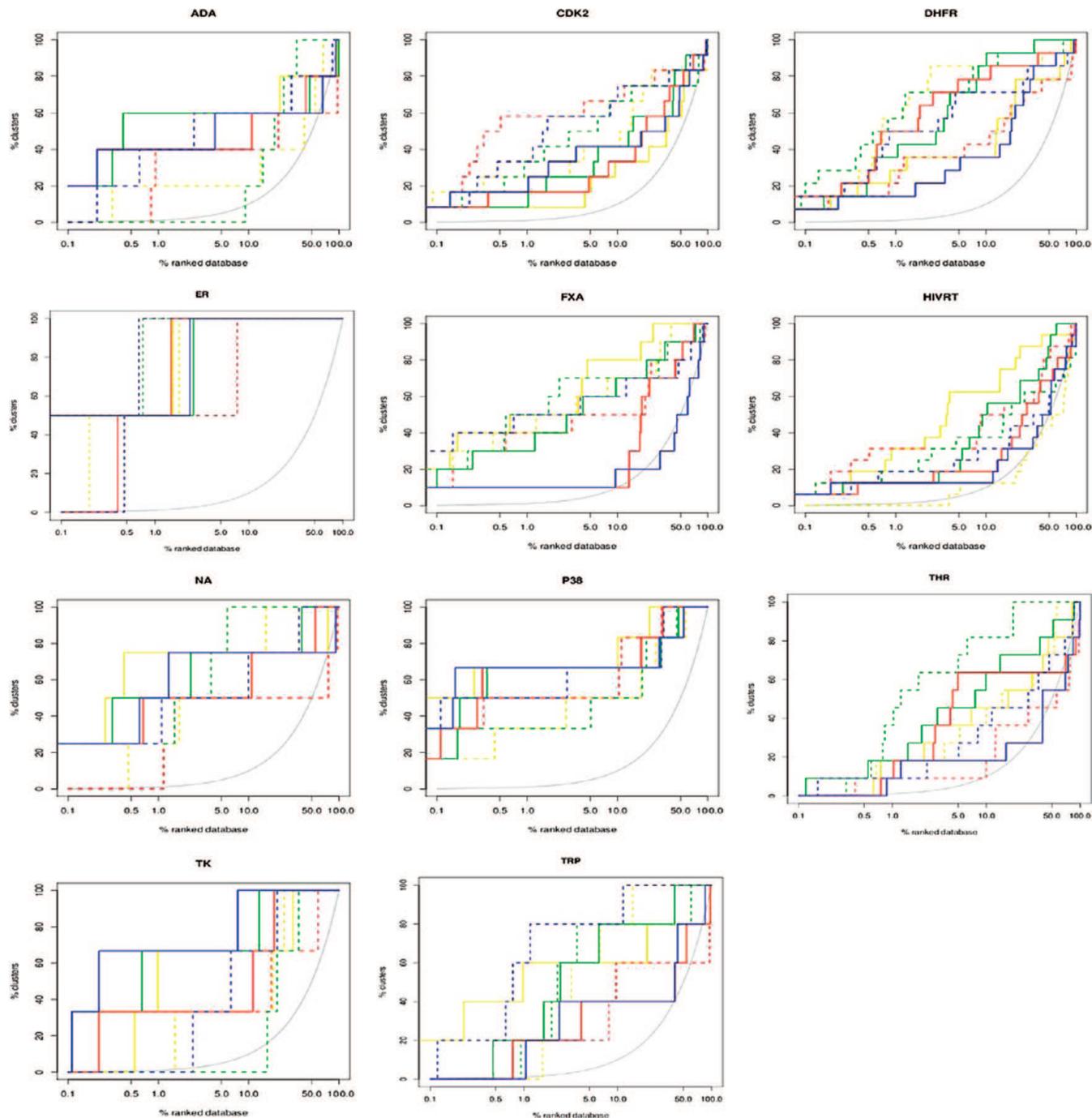


Figure 3. Cluster enrichment graphs with docking-BVLS methods (dotted lines) and 3D-LBVLs methods (plain lines). Each color is for a package: blue for Surflex-dock/Surflex-sim, green for ICM/ICM-sim, yellow for FlexX/FlexS, and red for FRED/ROCS.

sites more difficult. One of the possible reasons for the good performance of FRED on these 3 systems could be that FRED focuses on the optimization of only translational and rotational degrees of freedom during the calculation, which seems more effective for exploring such large binding sites than using “fully flexible” ligands. For FRED calculation, the torsional degrees of freedom are estimated using multiconformer libraries generated by OMEGA. Also, FRED calculations were performed using shape-gauss to evaluate the ligand poses, which is known to be relatively clash-tolerant when using default settings.³⁴ This could also be a possible reason for its success compared to the other methods for P38, taking into consideration that a non-negligible part

of the DUD-active compounds cannot bind to the provided structure because of the conformation of F169.

Surflex-dock appears to be the best on 3 systems (TK, NA, TRP) despite displaying the least overall docking accuracy (59.93% of acceptable poses), principally, because of its very poor performance on P38 and THR.

Overall the different programs evaluated in this study are able to perform acceptable molecular alignments that lead to an accurate superimposition of the active compounds on the reference for 3D-LBVLs methods and acceptable docking poses for docking-BVLS methods.

Comparative Performance in Enrichment. The quality of virtual ligand screening methods is evaluated by their

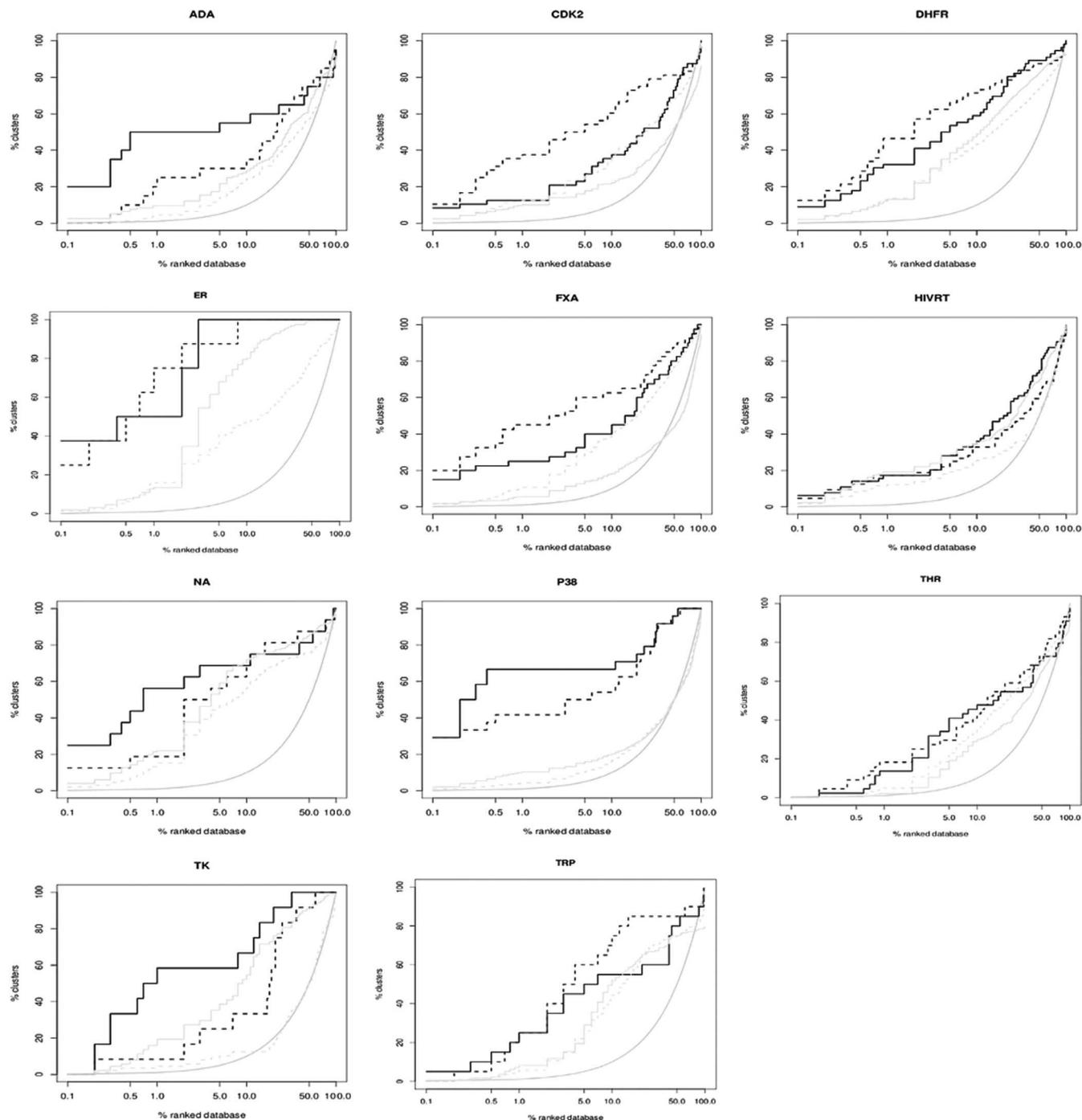


Figure 4. Mean enrichment graphs (gray) and mean cluster enrichment graphs (black) obtained by docking-BVLS methods (dotted lines) and 3D-LBVLS methods (plain lines).

ability to retrieve active compounds within the early ranked database, the so-called enrichment.

Overall, mean early and late enrichments are clearly in favor of 3D-LBVLS methods (9.48%/34.91% for docking-BVLS versus 12.07%/42.05% for 3D-LBVLS). This is particularly striking for ER where systematically all the 3D-LBVLS methods outperform docking-BVLS methods. Von Korff⁴⁵ suggested that this should not be surprising because of the construction of the DUD database as, decoys are mostly topologically dissimilar to the active compounds.²¹ However, as shown by Cleves,³⁶ the impact of this bias is important principally for 2D similarity search methods. An opposite opinion is given by Kirchmair¹⁷ when he suggests that docking-BVLS methods are in fact favored in the DUD

considering that docking has, more likely, data present on the dimensions of the active site. We believe that shape-similarity search methods are less sensitive to the possible bias highlighted by Von Korff and Cleves as several features directly related to the topology such as the bond and atom types are not primarily considered in 3D-LBVLS methods and, as Kirchmair stated,¹⁷ that ligand-based methods have to represent the binding site properties using a single active compound, which is already very challenging on its own.

Even considering possible bias, docking-BVLS methods are still very effective on 6 out of the 11 systems evaluated in this study (NA, TRP, HIVRT, DHFR, CDK2 and FXA), the most striking performance being with CDK2 and FXA.

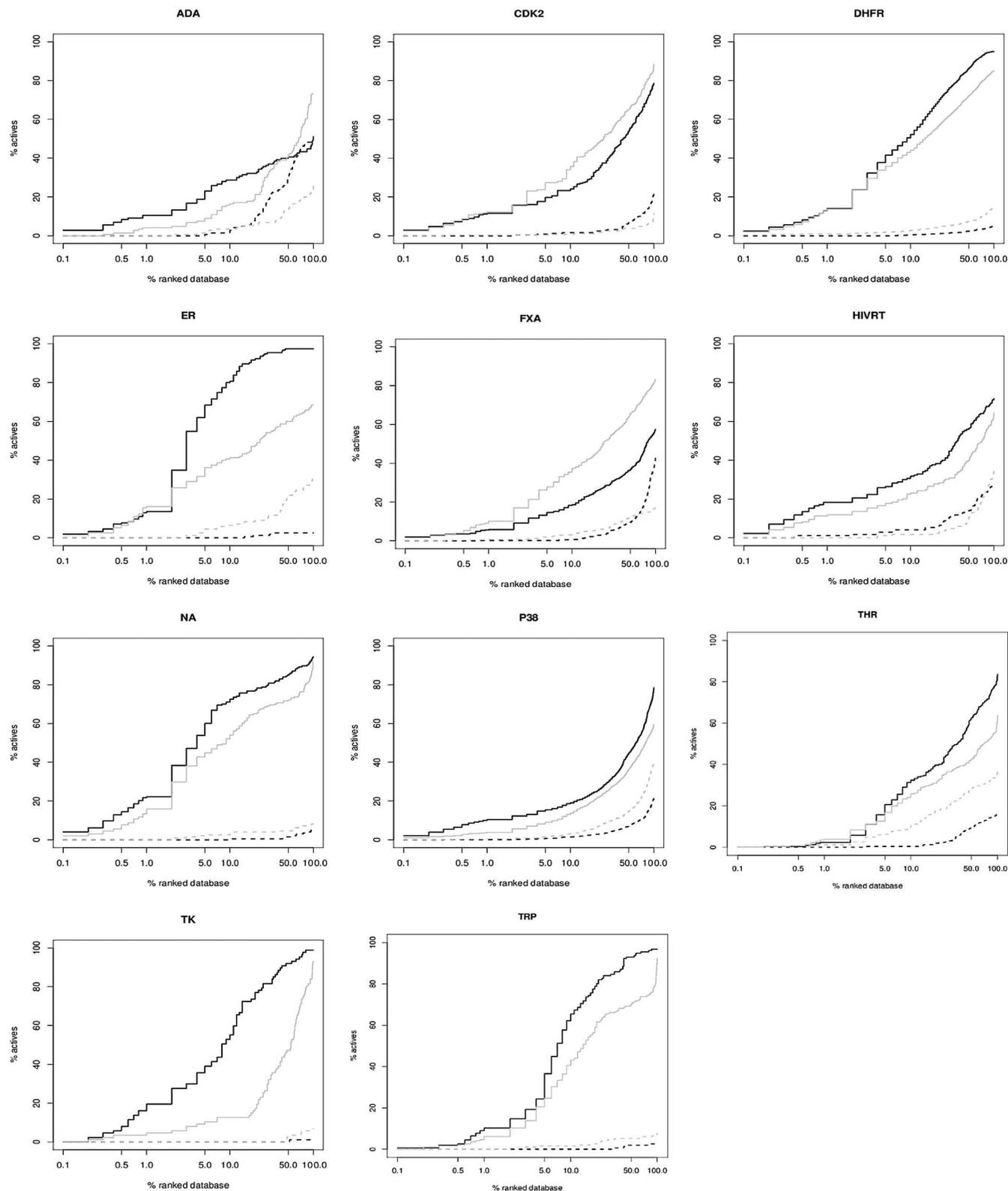


Figure 5. Molecular alignment of the active compounds using 3D-LBVLS methods (plain lines) and docking-BVLS methods (dotted lines). Light gray line shows the random enrichment. Black lines are for acceptable poses and gray lines for unacceptable poses.

Surflex-dock appears to be the most effective docking-BVLS method on enrichment. It also outperforms most of the 3D-LBVLS methods in several systems. By comparing our results with Cross et al,²⁰ the Surflex-dock ringflex parameter seems clearly valuable at least for serine proteases where we could see a significant impact (data not shown). It improves the enrichment for 5 out of the 11 systems evaluated in this study (ADA, CDK2, ER, FXA, and THR).

FRED appears to show its limits in enrichment except on the systems where the more complex docking methods fail, such as HIVRT, P38, and TK. This shows that in particularly difficult systems, multiconformer rigid-body docking methods can perform at least, as well as flexible docking. As for docking accuracy, one possible explanation can be the clash tolerance of the shape-gauss scoring function implemented in FRED for these targets where several clashes can occur

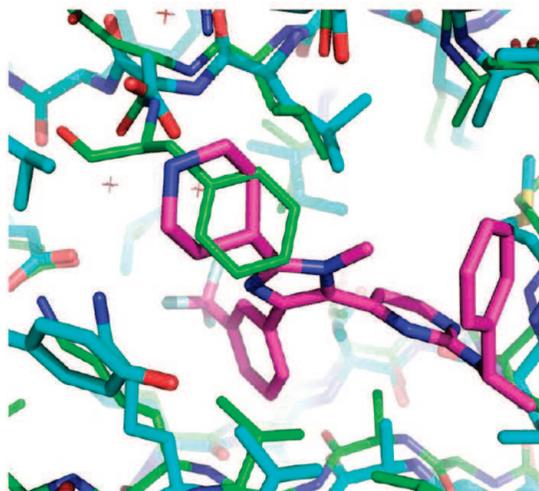


Figure 6. Illustration of a clash between F169 of the P38 structure provided by the DUD (green) and one of the compounds from the P38 DUD-own active set, ZINC03815672, positioned using ICM (pink). The blue structure represents an alternate PDB that could have accepted this compound.

with a non-negligible part of the active compounds of the corresponding DUD-own data set, for example, in the currently provided structure for P38 (see the docking-accuracy section), and in the small and buried binding site of TK and HIVRT.

The variance observed between the results is important within the DUD targets and, as observed by von Korff,⁴⁵ highlights the importance of multiple systems for benchmarking studies.

Comparison with Other Studies. On 2 out of the 11 targets evaluated in this study (ADA and DHFR), DOCK results from the original DUD paper²¹ display a better performance in early enrichment compared to the other docking methods. For the other targets, its performance is comparable to FlexX.

As expected, our results with ROCS are almost identical to those obtained by Kirchmair.¹⁷ The slight differences we observe can be explained by the different settings used for the multiconformer generation of the DUD-own sets with OMEGA but seem to have a very low impact on the final performance of the program. It is reassuring to have similar results using the same data sets and encouraging for using standardized data sets, such as the DUD for benchmarking studies. As proposed by the creators of the DUD,²¹ a broadened use of such data sets will be beneficiary for the whole community to compare their benchmarking results.

By comparing the ROC AUC with the early and late enrichment rates, we can draw several observations. ROC AUC is a good measure of the global discriminating performance of the different methods since we use the same conditions for testing. In general, similar conclusions can be drawn from the analysis of ROC curves/ROC AUC and enrichment graphs around 10% or 20% of the ranked database but in early enrichment it is not necessarily the case.¹⁷ For “real life” project decisions, the important point about the methods is the definition of the selection threshold to use, that is, the number of compounds to test experimentally after a virtual screening. This measure, in our opinion, cannot be necessarily provided using a metric, such as ROC AUC or even the analysis of ROC curves but rather by

classical enrichment values comparisons. For instance, in the case of ADA, by observing the ROC curves and AUC, the best method seems to be ROCS but in fact in early and late enrichment, it appears that Surflex-dock is the best method to use even if its ROC AUC is 10% lower. This highlights the point stated by Kirchmair¹⁷ that both ROC AUC and enrichment rates need to be considered in method evaluation and thus in the definition of new drug-discovery protocols. Once more, the use of standardized metrics in benchmarking studies will benefit the whole community.

Comparative Performance in Scaffold Enrichment. As expected, methods that display a good performance in enrichment also display a good performance in scaffold-enrichment. It is widely accepted that docking methods should be appropriate to retrieve compounds with different molecular scaffolds, as a protein’s binding site is more permissive than a pseudoreceptor derived from ligand-based hypothesis. Indeed, docking-BVLS methods perform well in scaffold enrichment, retrieving in average more than 50% of the different clusters in the top 10% for 7 out of the 11 systems studied here. However, for TK, NA, ADA, and P38, 3D-LBVLS methods outperform significantly docking-BVLS methods, highlighting their good performance in scaffold-hopping. By analyzing the trends from the mean enrichment graphs, we can conclude that overall, 3D-VLS methods have a similar acceptable performance in scaffold enrichment, depending on the system. It is thus of importance to underscore that, in the systems tested in the present study, docking-BVLS methods do not perform systematically better in scaffold-hopping than 3D-LBVLS methods. This makes 3D-LBVLS methods also reliable for retrieving diverse molecular scaffolds.

As a confirmation, we also tested the ability to retrieve compounds within the same cluster as the reference active compound provided by the DUD (thus more similar compounds i.e., the inner ability of LBVLS methods). As expected, in this trivial test for ligand-based methods, 3D-LBVLS methods and ROCS in particular display a much better overall performance, outperforming systematically docking-BVLS methods (11 out of the 11 systems studied, data not shown).

Molecular Alignment and Enrichment. It should be obvious that, using 3D-VLS methods, compounds that are retrieved in the top scores should be positioned accurately in the binding site/superimposed accurately on the reference compound (i.e., accurate molecular alignment). We thus analyzed the molecular alignment of the different active compounds after 3D-VLS.

As expected, for the 3D-LBVLS programs evaluated in this study, more than 80% of the active compounds retrieved in the top 10% are positioned in an acceptable manner, while a very low number display unacceptable poses: 19/338 for Surflex-sim, 8/500 for ROCS, 10/712 for ICM-sim, and 19/399 for FlexS. Only few exceptions occur for FXA and CDK2 with Surflex-sim and for ADA with ROCS. This is not surprising since the scoring in 3D-LBVLS methods is directly based on the molecular alignment and not on a (pseudo) binding energy evaluation as in docking-BVLS methods.

In the case of the docking-BVLS methods, all the programs but FRED retrieve more than 80% of the active compounds in the top 10% with an acceptable positioning but we can

observe that a non-negligible part of the active compounds retrieved in the top 10% display unacceptable binding modes: 66/576 for Surflex-dock, 39/268 for FRED, 36/518 for ICM, 37/467 for FlexX, and particularly in the case of THR and P38. One possible reason might be that the active compounds of these systems, where the positioning is difficult, display high flexibility and are to be positioned in relatively large binding sites which makes the conformational search more difficult for docking-BVLS methods.

This illustrates one of the surprising current paradoxes in virtual screening that is retrieving active compounds in the top scores while displaying an inaccurate molecular alignment. This paradox is much more frequent with docking-BVLS methods compared to 3D-LBVLS methods as the conformational space to explore is wider in a "real" binding pocket relatively to a pseudobinding site derived from a ligand-based hypothesis and thus more susceptible to errors. It is difficult to pinpoint a specific cause for this paradox but there might be several factors including (1) the weight of the different parameters in scoring functions which generally do not take correctly into account solvation/desolvation and entropy and (2) the rigid-body treatment of the atoms of the binding pocket or the limited handling of flexibility of the compounds during the simulation (especially for highly flexible compounds).

CONCLUSION

We have investigated the performance of 8 different 3D-VLS programs on 11 DUD systems using their corresponding DUD-own database. The use of the DUD-own database is more challenging compared to the general DUD data set and is more appropriate for benchmarking studies. The programs evaluated in this study are all able to perform acceptable molecular alignments with standard parameters. In this exercise, ICM appeared to be the most efficient underscoring the good performance of the probability biased Monte Carlo procedure used for simulating the compounds' flexibility. We identified several problems with some of the structures provided by the DUD. In particular, for P38, a non-negligible part of the ligands provided as active compounds appeared unable to bind the provided structure because of major clashes biasing negatively the results for docking-BVLS methods evaluation.

Concerning the performance of retrieving active compounds among decoys, that is, the enrichment, it was overall acceptable but variable depending on the target and the program used. Surflex-dock and ICM had the best global performance in enrichment. In terms of scaffold enrichment, 3D-LBVLS methods showed a comparable performance to docking-BVLS methods, which is quite surprising as docking-methods are generally considered to be the methods of reference for retrieving diverse molecular scaffolds. This is very interesting for the definition of new drug-discovery projects since 3D-LBVLS methods use very limited computational resources compared to the more sophisticated docking-BVLS methods (excluding FRED) thus opening new questions about which method to use when the structure of a small-molecule/protein target complex is known. The variance observed between the results is important within the DUD targets and highlights the importance of multiple systems for benchmarking studies.

Finally this study illustrates the molecular alignment paradox in enrichment that consists in finding a non-negligible number of active compounds displaying good scores, and thus good ranks but with an inaccurate positioning into the binding pocket. Considering that docking methods are in most of the cases able to produce accurate binding modes, this highlights the current limits of the scoring functions used in docking-BVLS methods that still need to be optimized to avoid such problems.

Abbreviations. 3D-LBVLS, 3D ligand-based virtual ligand screening; docking-BVLS, docking-based virtual screening; 3D-VLS, 3D virtual ligand screening; Tc, Tanimoto coefficient; ADA, adenosine deaminase; CDK2, cyclin dependent kinase 2; DHFR, dihydrofolate reductase; ER, antagonists for estrogen receptor; FXA, coagulation factor Xa; HIVRT, HIV reverse transcriptase; NA, neuraminidase; P38, mitogen-activated protein; THR, thrombin; TK, thymidine kinase; TRP, trypsin.

ACKNOWLEDGMENT

We are very grateful to Openeye Scientific Software for generously providing ROCS, FRED and OMEGA. We thank Prof Jain for generously providing the Surflex package. We would like to also thank Molsoft LLC and BiosolveIT for providing academic licenses for the ICM suite and the FlexX/FlexS package. HG is recipient of a CIFRE fellowship from ANRT.

REFERENCES AND NOTES

- (1) Alvarez, J. C. High-throughput docking as a source of novel drug leads. *Curr. Opin. Chem. Biol.* **2004**, *8* (4), 365–370.
- (2) Lajiness, M. S.; Johnson, M. A.; Maggiora, G. M. Implementing drug screening programs using molecular similarity methods. *Prog. Clin. Biol. Res.* **1989**, *291*, 173–176.
- (3) Johnson, M.; Lajiness, M.; Maggiora, G. Molecular similarity: A basis for designing drug screening programs. *Prog. Clin. Biol. Res.* **1989**, *291*, 167–171.
- (4) Guner, O. F.; Hughes, D. W.; Dumont, L. M. An integrated approach to three-dimensional information management with MACCS-3D. *J. Chem. Inf. Comput. Sci.* **1991**, *31* (3), 408–414.
- (5) Good, A. C.; Hermsmeider, M. A.; Hindle, S. A. Measuring CAMD technique performance: A virtual screening case study in the design of validation experiments. *J. Comput.-Aided Mol. Des.* **2004**, *18* (7–9), 529–536.
- (6) Schneider, G.; Neidhart, W.; Giller, T.; Schmid, G. Scaffold-hopping by topological pharmacophore search: A contribution to virtual screening. *Angew. Chem., Int. Ed. Engl.* **1999**, *38* (19), 2894–2896.
- (7) Kirchmair, J.; Ristic, S.; Eder, K.; Markt, P.; Wolber, G.; Laggner, C.; Langer, T. Fast and efficient in silico 3D screening: Toward maximum computational efficiency of pharmacophore-based and shape-based approaches. *J. Chem. Inf. Model.* **2007**, *47* (6), 2182–2196.
- (8) Putta, S.; Eksterowicz, J.; Lemmen, C.; Stanton, R. A novel subshape molecular descriptor. *J. Chem. Inf. Comput. Sci.* **2003**, *43* (5), 1623–1635.
- (9) Haigh, J. A.; Pickup, B. T.; Grant, J. A.; Nicholls, A. Small molecule shape-fingerprints. *J. Chem. Inf. Model.* **2005**, *45* (3), 673–684.
- (10) Kroemer, R. T. Structure-based drug design: Docking and scoring. *Curr. Protein Pept. Sci.* **2007**, *8* (4), 312–328.
- (11) Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. J. Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. *Curr. Med. Chem.* **2004**, *11* (1), 71–90.
- (12) Bowman, A. L.; Lerner, M. G.; Carlson, H. A. Protein flexibility and species specificity in structure-based drug discovery: Dihydrofolate reductase as a test system. *J. Am. Chem. Soc.* **2007**, *129* (12), 3634–40.
- (13) Rush, T. S., 3rd; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48* (5), 1489–95.
- (14) Bologna, C. G.; Revankar, C. M.; Young, S. M.; Edwards, B. S.; Arterburn, J. B.; Kiselyov, A. S.; Parker, M. A.; Tkachenko, S. E.;

- Savchuck, N. P.; Sklar, L. A.; Oprea, T. I.; Prossnitz, E. R. Virtual and biomolecular screening converge on a selective agonist for GPR30. *Nat. Chem. Biol.* **2006**, *2* (4), 207–212.
- (15) Freitas, R. F.; Oprea, T. I.; Montanari, C. A. 2D QSAR and similarity studies on cruzain inhibitors aimed at improving selectivity over cathepsin L. *Bioorg. Med. Chem.* **2008**, *16* (2), 838–853.
- (16) Bostrom, J.; Berggren, K.; Elebring, T.; Greasley, P. J.; Wilstermann, M. Scaffold hopping, synthesis, and structure–activity relationships of 5,6-diaryl-pyrazine-2-amide derivatives: a novel series of CB1 receptor antagonists. *Bioorg. Med. Chem.* **2007**, *15* (12), 4077–4084.
- (17) Kirchmair, J.; Distinto, S.; Markt, P.; Schuster, D.; Spitzer, G. M.; Liedl, K. R.; Wolber, G. How to optimize shape-based virtual screening: choosing the right query and including chemical information. *J. Chem. Inf. Model.* **2009**, *49* (3), 678–692.
- (18) Moffat, K.; Gillet, V. J.; Whittle, M.; Bravi, G.; Leach, A. R. A comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS. *J. Chem. Inf. Model.* **2008**, *48* (4), 719–729.
- (19) Lemmen, C.; Lengauer, T. Computational methods for the structural alignment of molecules. *J. Comput.-Aided Mol. Des.* **2000**, *14* (3), 215–232.
- (20) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. Comparison of several molecular docking programs: Pose prediction and virtual screening accuracy. *J. Chem. Inf. Model.* **2009**, *49* (6), 1455–1474.
- (21) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49* (23), 6789–6801.
- (22) Rarey, M.; Kramer, B.; Lengauer, T. Time-efficient docking of flexible ligands into active sites of proteins. *Proc.-Int. Conf. Intell. Syst. Mol. Biol.* **1995**, *3*, 300–308.
- (23) Lemmen, C.; Lengauer, T.; Klebe, G. FLEXS: A method for fast flexible ligand superposition. *J. Med. Chem.* **1998**, *41* (23), 4502–4520.
- (24) Bohm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* **1994**, *8* (3), 243–256.
- (25) Klebe, G. The use of composite crystal-field environments in molecular recognition and the de novo design of protein ligands. *J. Mol. Biol.* **1994**, *237* (2), 212–235.
- (26) Jain, A. N. Surflex: Fully automatic flexible molecular docking using a molecular similarity-based search engine. *J. Med. Chem.* **2003**, *46* (4), 499–511.
- (27) Jain, A. N. Ligand-based structural hypotheses for virtual screening. *J. Med. Chem.* **2004**, *47* (4), 947–961.
- (28) Welch, W.; Ruppert, J.; Jain, A. N. Hammerhead: Fast, fully automated docking of flexible ligands to protein binding sites. *Chem. Biol.* **1996**, *3* (6), 449–462.
- (29) Jain, A. N. Morphological similarity: A 3D molecular similarity method correlated with protein–ligand recognition. *J. Comput.-Aided Mol. Des.* **2000**, *14* (2), 199–213.
- (30) Abagyan, R.; Totrov, M.; Kuznetsov, D. ICM (internal coordinate mechanics). *J. Comput. Chem.* **1994**, *15*, 488.
- (31) Schapira, M.; Totrov, M.; Abagyan, R. Prediction of the binding energy for small molecules, peptides, and proteins. *J. Mol. Recognit.* **1999**, *12* (3), 177–190.
- (32) Totrov, M. Atomic property fields: Generalized 3D pharmacophore potential for automated ligand superposition, pharmacophore elucidation, and 3D QSAR. *Chem. Biol. Drug Des.* **2008**, *71* (1), 15–27.
- (33) OpenEye Scientific Software. <http://www.eyesopen.com>.
- (34) McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. Gaussian docking functions. *Biopolymers* **2003**, *68* (1), 76–90.
- (35) Irwin, J. J.; Shoichet, B. K. ZINC—A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182.
- (36) Cleves, A. E.; Jain, A. N. Effects of inductive bias on computational evaluations of ligand-based modeling and on drug discovery. *J. Comput.-Aided Mol. Des.* **2008**, *22* (3–4), 147–159.
- (37) Hawkins, P. C.; Warren, G. L.; Skillman, A. G.; Nicholls, A. How to do an evaluation: pitfalls and traps. *J. Comput.-Aided Mol. Des.* **2008**, *22* (3–4), 179–190.
- (38) Good, A. C.; Oprea, T. I. Optimization of CAMD techniques 3. Virtual screening enrichment studies: A help or hindrance in tool selection. *J. Comput.-Aided Mol. Des.* **2008**, *22* (3–4), 169–178.
- (39) Sing, T.; Sander, O.; Beerenwinkel, N.; Lengauer, T. ROC-R: Visualizing classifier performance in R. *Bioinformatics* **2005**, *21* (20), 3940–3941.
- (40) Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335–373.
- (41) Fischer, B.; Fukuzawa, K.; Wenzel, W. Receptor-specific scoring functions derived from quantum chemical models improve affinity estimates for in-silico drug discovery. *Proteins* **2008**, *70* (4), 1264–1273.
- (42) Perez-Nueno, V. I.; Ritchie, D. W.; Rabal, O.; Pascual, R.; Borrell, J. I.; Teixido, J. Comparison of ligand-based and receptor-based virtual screening of HIV entry inhibitors for the CXCR4 and CCR5 receptors using 3D ligand shape matching and ligand-receptor docking. *J. Chem. Inf. Model.* **2008**, *48* (3), 509–533.
- (43) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* **2006**, *49* (20), 5912–5931.
- (44) Abagyan, R.; Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **1994**, *235* (3), 983–1002.
- (45) von Korff, M.; Freyss, J.; Sander, T. Comparison of ligand- and structure-based virtual screening on the DUD data set. *J. Chem. Inf. Model.* **2009**, *49* (2), 209–231.

CI900507G

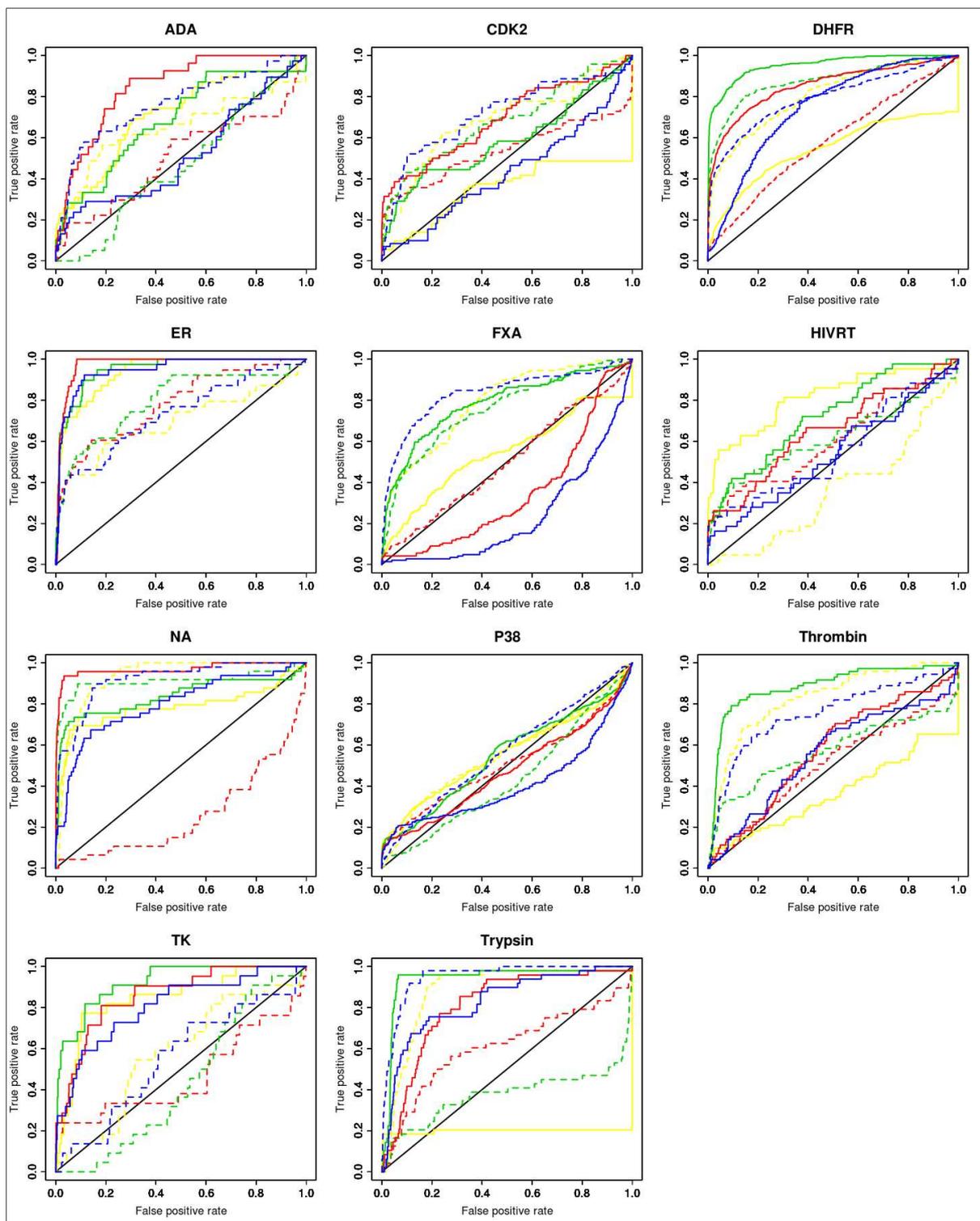


Figure 10: Courbes de ROC globales. Les méthodes SBVLS sont représentées par un trait plein et les méthodes LBVLS par des pointillés. Les méthodes Surflex-dock/Surflex-sim, FRED/ROCS, FlexX/FlexS et ICM/ICMsim sont représentées en bleu, rouge, jaune et vert, respectivement.

Les courbes de ROC permettent d'observer les mêmes tendances qu'avec les courbes d'enrichissement. Elles permettent de confirmer les résultats observés sur les courbes d'enrichissement en s'émancipant du ratio actifs / inactifs. Cette fois, l'interprétation ne se fait plus en fonction du pourcentage parcouru de la banque, mais du taux de faux positifs.

DISCUSSION

Les deux types de méthodes peuvent produire une conformation similaire à la conformation bioactive de la molécule de référence. La difficulté est plus importante pour les méthodes SBVLS, l'espace conformationnel du site de liaison étant plus complexe que celui d'un pseudo-récepteur défini à partir d'une molécule de référence. D'autant plus que la construction de la DUD peut entraîner la présence d'artefacts. Par exemple, la conformation du site de liaison présente dans la DUD n'est pas toujours adaptée à tous les actifs, notamment en ce qui concerne P38. Plusieurs conformations peuvent être observées pour le résidu F169 du site de liaison. La conformation retenue dans la DUD empêche le bon positionnement de 18% des composés. Les enrichissements par les méthodes SBVLS sont d'ailleurs relativement faible, n'excédant pas une valeur de 10 à 1% et une seule méthode réussissant à faire un enrichissement de plus de 20 à 10%. Ainsi, leurs méthodes LBVLS respectives les surpassent quasi systématiquement (seule exception : FRED surpasse ROCS à 10%).

La même chose, dans une moindre mesure, peut être observée pour ADA, FXA et TK. Globalement, ICM et ICMsim sont les méthodes qui ont les meilleurs positionnement et alignements, respectivement, peut-être en raison d'une meilleure exploration de l'espace conformationnel par l'algorithme de Monte Carlo. Cependant, toutes les méthodes sont capables de faire des positionnements ou alignements précis et de se distinguer sur une cible particulière.

Les enrichissements moyens à 1 et 10 % sont clairement en faveur des méthodes LBVLS, ce qui ne serait pas surprenant puisque les inactifs sont, pour la plupart, topologiquement différents des actifs¹⁷⁹, et les méthodes LBVLS se basant en grande partie sur la similarité de forme avec la molécule de référence. Ainsi, les molécules ayant une topologie différente de la molécule de référence seront évaluées défavorablement. Cependant, cette différence jouerait surtout un rôle important pour les méthodes de similarité 2D¹⁸⁰. Kirchmair¹⁷¹ suggère au contraire que le docking est favorisé car il possède les informations du site de liaison.

L'hypothèse de Kirchmair semble être la plus acceptable, les méthodes de similarité 3D ne considérant pas directement les liaisons et types d'atomes (liés à la topologie), mais devant représenter un pseudo-site de liaison, ce qui est complexe en soi. En effet, les méthodes SBVLS font un meilleur enrichissement que les méthodes LBVLS pour 6 systèmes sur 11. Surflex-dock est la méthode SBVLS la plus efficace et fait un meilleur enrichissement que les méthodes LBVLS sur plusieurs systèmes. La variance observée entre les résultats est importante parmi les cibles de la DUD, ce qui souligne l'importance d'utiliser plusieurs systèmes pour l'évaluation de méthodes. Les résultats sont similaires à ceux obtenus par Kirchmair¹⁷¹.

La diversité chimique de l'enrichissement a aussi été étudiée. Comme attendu, les méthodes faisant un bon enrichissement en actifs font aussi un bon enrichissement en chemotypes. Bien que plusieurs clusters aient été établis pour les actifs de chaque cible, ceux-ci ne montrent généralement pas une diversité très élevée. Les méthodes de docking sont supposées être meilleures, le site de liaison étant plus permissif qu'un pseudo-site de liaison dérivé à partir d'une molécule de référence, ce qui est effectivement le cas pour la majorité des cibles. Cependant, les méthodes de LBVLS fonctionnent mieux pour 4 des 11 systèmes. Même pour des systèmes pour lesquels les actifs sont très variés (par exemple HIVRT avec 16 groupes de chemotypes pour 43 actifs), les méthodes LBVLS fonctionnent bien : 3 méthodes de LBVLS fonctionnent mieux que leur méthode SBVLS respective à 1%, et une égalité est observée à 10%. De manière surprenante, les méthodes LBVLS ne font pas un enrichissement spectaculaire des actifs de P38 bien qu'ils soient peu variés (6 groupes pour 454 actifs), mais ils surpassent les méthodes SBVLS quasi systématiquement, exception faite de FRED enrichissant mieux que ROCS à 10%.

La capacité des méthodes à retrouver les molécules les plus similaires au ligand de référence a également été observée. Comme attendu, les méthodes LBVLS fonctionnent systématiquement mieux que les méthodes SBVLS. Cette performance a sans doute un impact sur l'enrichissement global observé, les différents groupes de chemotypes n'étant pas de taille similaires au sein d'un même système et le ligand de référence n'appartenant pas toujours au plus grand groupe.

Il paraît évident que les actifs ayant les meilleurs scores devraient être positionnés dans le site de liaison, ou superposés sur le composé de référence, de manière précise. Pour les méthodes

LBVLS, plus de 80% des actifs retrouvés dans le top 10% de la banque sont alignés de manière acceptable. Ceci est en accord avec le fait que les métriques de similarité sont basées directement sur l'alignement moléculaire. Toutes les méthodes SBVLS, à part FRED, retrouvent plus de 80% des actifs du top 10% avec une pose acceptable. Cependant, le nombre de poses inacceptables n'est pas négligeable, notamment pour THR et P38. Les sites de liaisons de ces protéines sont grands et l'exploration de l'espace conformationnel est donc plus difficile pour les méthodes SBVLS. Cependant, un seul complexe par cible a été utilisé pour évaluer la qualité du positionnement, ne prenant pas (ou peu) en compte les modes de liaisons qui diffèrent selon les actifs.

Ces observations soulèvent un des paradoxes des méthodes de criblage *in silico* : certains actifs ayant obtenus les meilleurs scores sont mal positionnés ou alignés. Ce paradoxe est plus fréquemment observé avec les méthodes SBVLS. Trois explications possibles peuvent être avancées :

- Les fonctions de score ne prennent généralement pas en compte la solvation et l'entropie correctement
- Le site de liaison est considéré de manière rigide (ne prenant ainsi pas en compte tous les modes de liaisons potentiels) et la gestion de la flexibilité des ligands n'est pas toujours suffisante, notamment pour les ligands très flexibles.
- L'espace conformationnel à considérer est plus grand pour les méthodes SBVLS que pour les méthodes LBVLS, plus de positionnements étant possibles pour chaque composé.

Dans cette étude, une seule structure a été utilisée pour chaque système. L'utilisation de plusieurs structures par système pourrait permettre de gérer la flexibilité du site de liaison.

II.1.2 Structures multiples pour le criblage *in silico* de petites molécules : définition des critères de choix de la bonne structure cible à partir des propriétés des sites de liaison

Des changements conformationnels se produisent lors de la liaison du ligand dans le site de liaison du récepteur (induced fit). Mais, de manière classique, la structure du site de liaison reste rigide pendant le docking. Plusieurs approches ont été développées pour prendre en compte la flexibilité du site de liaison de manière dynamique pendant le docking. La flexibilité du site de liaison peut aussi être représentée en amont, en utilisant plusieurs structures ayant des conformations différentes. Cette méthode est appelée « ensemble docking ». Les structures utilisées peuvent avoir été obtenues expérimentalement^{181; 182} ou dérivées à partir de simulations^{183; 184}. Ces approches sont plus coûteuses en temps de calcul que les approches de docking classiques et le taux de faux positifs observé augmente parfois. Cette augmentation peut être expliquée par des erreurs de score mais aussi par le choix des structures constituant les ensembles^{181; 185; 186}. Cependant, la plupart des méthodes de docking prenant en compte la flexibilité du récepteur le font de manière implicite^{100; 187}.

Dans la littérature récente, plusieurs problèmes avec les structures sélectionnées dans la DUD ont été mis en évidence^{188; 189}. Par exemple, le résidu F169 de la protéine P38 MAP kinase est orienté de telle sorte qu'un nombre non négligeable d'actifs ne peuvent pas être positionnés convenablement dans le site de liaison, aboutissant à des enrichissements artificiellement faibles. Puisque plusieurs structures de cette protéine sont disponibles dans la PDB, ce biais aurait pu être évité en choisissant une structure plus adaptée, ou en faisant de l'ensemble docking. Bien que plusieurs études aient été faites sur ce problème, il n'y a toujours pas de consensus pour une approche ou une autre¹⁹⁰⁻¹⁹². Pour les ensemble docking, les petits ensembles de protéines (moins de 5) semblent être à privilégier^{181; 192; 193}. Selon Bottegoni et al¹⁹⁴, l'ensemble docking fonctionne, en moyenne, mieux que le docking sur une structure seule. Des études préliminaires devraient donc être faites pour déterminer les meilleures structures à utiliser^{192; 195; 196}.

Pour un docking, le choix de la conformation de départ est primordial^{159; 188}. Nous avons donc décidé de rechercher des tendances dans les propriétés des sites de liaison conduisant aux

performances optimales des études de docking. Les systèmes de la DUD pour lesquels plusieurs structures expérimentales étaient présentes dans la PDB ont été sélectionnés (tous sauf PDGFRb). Jusqu'à 4 structures expérimentales ont été choisies, pour lesquelles les topologies (volume, ouverture) des sites de liaison sont significativement différentes. Pour chaque structure, les performances en terme d'enrichissement pour 2 méthodes de docking (Surflex-dock et ICM) ont été évaluées à l'aide des actifs et inactifs spécifiques de chaque système. Cette étude a été faite dans le but de :

- Confirmer l'impact du choix d'une structure sur les résultats d'un criblage *in silico*
- Identifier les tendances potentielles de propriétés structurales partagées par les structures donnant les meilleurs ou les moins bons résultats
- Réduire le biais possible des tendances qui pourrait être observé en utilisant une seule structure pour un docking rigide

Puisque la puissance de calcul disponible est maintenant suffisante, les performances de Surflex-dock et ICM ont été évaluées sur tous les ensembles de structures possibles sélectionnées pour cette étude. Cette évaluation a été faite pour identifier les structures optimales composant les ensembles et comparer leurs propriétés structurales.

MATERIELS ET METHODES

DUD. La Directory of Useful Decoys¹⁶⁰ est une banque d'évaluation contenant 40 systèmes. Chaque système est composé d'une structure cristallographique de la protéine cible et du ligand cocrystallisé, d'actifs connus et, pour chaque actif, de 36 inactifs sélectionnés pour leurs propriétés physicochimiques similaires aux actifs. Ils sont en revanche structurellement différents. Tous les systèmes, pour lesquels au moins une structure expérimentale de la cible était disponible, ont été retenus.

Surflex-dock. L'algorithme d'exploration de l'espace conformationnel de Surflex-dock est dérivé de l'algorithme de fragmentation / reconstruction de Hammerhead¹¹³. Le site de liaison est représenté par un protomol, c'est à dire des fragments moléculaires (NH, C=O, CH₄) recouvrant tout le site de liaison. Les poses sont évaluées par une fonction de score empirique.

Pour chaque structure, le site de liaison a été défini à 4 Å autour du ligand co-cristallisé. La version 2.5 de Surflex-dock a été utilisée pour cette étude.

ICM. ICM¹¹⁵ est basé sur l'algorithme de Métropolis Monte Carlo pour optimiser le positionnement des molécules dans le site de liaison. Les poses sont évaluées avec la fonction de score empirique ICM-VLS. La version 3.6 d'ICM a été utilisée pour cette étude.

Ensemble docking. Pour chaque cible, tous les ensembles possibles (jusqu'à onze ensemble pour un système donné) ont été construits à partir des structures expérimentales de la PDB. Les expériences de docking sur les ensembles ont été faits avec Surflex-dock et ICM. Pour un composé donné sur une cible donnée, le meilleur score obtenu avec les différentes structures de l'ensemble est retenu. Ainsi, pour chaque ensemble, les résultats du docking consistent en une liste avec les scores des meilleures poses.

Métriques des performances. Tous les graphes ont été faits avec l'outil statistique et graphique R. Le module ROCR a été utilisé pour tracer les courbes de ROC et l'algorithme Wilcoxon-Mann-Whitney pour calculer l'aire sous la courbe de ROC¹⁹⁷.

Pour déterminer la catégorie (polaire / non polaire, grande / petite) à laquelle appartient chaque système, les moyennes du taux de surface polaire accessible au solvant et du volume des sites de liaison constituant chaque système. Un système est dans la catégorie des petits sites de liaison lorsque le volume moyen est inférieur ou égal à 350 Å³. Lorsque la moyenne des taux de surface apolaire accessible au solvant est inférieur à 70%, le système est dans la catégorie des sites de liaison polaires. Ce taux est calculé avec GetArea¹⁹⁸ en faisant le ratio de l'aire de la surface non polaire accessible au solvant (ASASA) et de l'aire de la surface accessible au solvant (SASA).

RESULTATS ET DISCUSSION

Pour les systèmes de la DUD pour lesquels au moins une structure expérimentale est disponible (tous sauf PDGFRB), toutes les structures présentes dans la PDB ont été étudiées, c'est à dire entre 2 et 300 structures par système. Jusqu'à 4 structures par système ont été conservées dans le but de couvrir toute la flexibilité expérimentale disponible pour le volume et l'ouverture de la poche. Les performances en terme d'enrichissement de Surflex-dock et

ICM ont été observées afin de déterminer les tendances parmi les propriétés structurales des conformations donnant les meilleurs ou pires enrichissements.

Pour 19 systèmes sur 39, la conformation pour laquelle la meilleure performance est observée pour chaque système est commune à Surfex-dock et ICM. Pour 21 systèmes sur 39, la moins bonne performance est aussi associée à la même conformation pour Surfex-dock et ICM. La conformation de la structure de la cible semble donc avoir un impact sur la qualité de l'enrichissement, ce qui est en accord avec d'autres études^{159; 193; 199}. De plus, que ce soit avec Surfex-dock (pour 17 systèmes sur 39) ou ICM (18 sur 39), la meilleure performance est observée pour les conformations dont le site de liaison a le plus petit volume. A l'opposé, la moins bonne performance est souvent observée pour les plus grands sites de liaison (20 sur 39 pour Surfex-dock et 21 sur 39 pour ICM). Pour Surfex-dock, parmi les systèmes ayant un site de liaison avec un petit volume moyen, la conformation la plus fermée est le plus souvent associée à la meilleure performance (10 sur 16) tandis que la conformation la plus ouverte correspondrait à la moins bonne performance (10 sur 16). Pour les systèmes dont le site de liaison a un grand volume moyen, la conformation la plus fermée entraîne les pires enrichissements pour 11 systèmes sur 19, que ce soit avec Surfex-dock ou ICM.

Pour chaque système, tous les ensembles de 2 à 4 structures ont été construits. Pour chaque composé, le meilleur score, obtenu par docking classique sur chaque structure de l'ensemble, est retenu. L'analyse a été faite sur tous les systèmes où plus d'un ensemble pouvaient être construit, soit 21 systèmes. Que ce soit pour Surfex-dock ou ICM, les plus petits ensembles aboutissent quasi systématiquement aux meilleurs enrichissements (20 et 19 sur 21 systèmes, respectivement), ce qui est en accord avec les études précédentes^{181; 192; 195}. La meilleure structure seule est généralement présente dans le meilleur ensemble (systématiquement pour Surfex-dock et pour 17 systèmes sur 21 pour ICM) et le plus petit site de liaison est présent dans cet ensemble pour 17 et 16 systèmes sur 21, respectivement. De la même manière, la structure seule associée à la plus mauvaise performance est toujours ou presque dans le plus mauvais ensemble (21 et 20 systèmes sur 21, respectivement). Bien que pour Surfex-dock, 17 systèmes sur 21 ont la conformation la plus grande dans le plus mauvais ensemble, aucune tendance ne peut être déterminée pour ICM.

L'utilisation d'ensembles de plusieurs structures permet de prendre en compte la flexibilité du site de liaison, entraînant une meilleure adaptation de ce site aux molécules criblées, qu'elles

soient actives ou inactives. Il était donc nécessaire de déterminer les conformations maximisant l'adaptation aux actifs s'en trop augmenter l'adaptation aux inactifs. Ainsi, une prise en compte discrète de la flexibilité (à l'aide d'un petit nombre de conformations) peut paraître plus adaptée qu'une prise en compte continue (en faisant de la dynamique moléculaire par exemple).

Pour savoir s'il est nécessaire d'utiliser plusieurs structures, les performances des meilleurs ensembles et des meilleures structures seules ont été comparées. Aucune tendance ne se dégage, pour Surflex-dock et ICM, la meilleure structure seule aboutissant à un meilleur enrichissement pour 18 systèmes sur 39, et inversement pour 19 systèmes sur 39. Aucune différence notable n'a été observée pour 2 systèmes.

CONCLUSION

A partir de deux méthodes de docking, plusieurs tendances ont été observées mettant en jeu les propriétés des sites de liaison :

- Les structures entraînant les meilleurs et les moins bons enrichissements sont généralement communes aux deux méthodes
- Les meilleurs enrichissements sont observés pour la conformation de plus faible volume (à l'opposé, les moins bonnes performances sont associées aux plus grandes conformations)
- Les ensembles les plus petits fonctionnent le mieux et contiennent presque toujours la meilleure structure seule.
- Les ensembles avec les pires enrichissements contiennent généralement la pire structure seule.

En revanche, il ne semble pas y avoir de tendance permettant de choisir plutôt le meilleur ensemble que la meilleure structure seule, le nombre de systèmes en faveur de chacun d'eux étant similaire.

Small and Polar												
	Mean volume (Å ³)	Mean % hydrophobicity	# DUD actives	# Available PDB structures	PDB ID	Resolution (Å)	RMSD (Å)	Volume (Å ³)	Relative opening	SASA (Å ²)	ASASA (Å ²)	% Hydrophobicity
NA	243	45.0	49	5	1A4G	2.2		227	+	810.94	356.16	43.9
					1A4Q	1.9	0.471	258	+++	854.53	394.47	46.2
HSP90	342	60.4	37	62	2CDD	1.9		228	++	595.76	326.53	54.8
					3K99	2.1	3.279	233	+	565.48	353	62.4
					2XJX	1.66	2.785	357	++++	586.08	349.07	59.6
					1UYF	2	2.299	549	+++	775.09	503.33	64.9
PNP	159	60.7	50	7	1B8O	1.5		98	+	398.64	211.62	53.1
					2QPL	2.1	0.579	148	+++	824.75	513.52	62.3
					1V48	2.2	0.734	230	++	886.51	591.93	66.8
TRP	138	63.6	49	300	1BTY	1.5		60	+	441.9	272.75	61.7
					3AAU	1.8	0.196	171	+++	773.09	484.91	62.7
					1V2O	1.62	0.951	183	++	810.34	538.76	66.5
VEGFR2	322	63.8	88	3	2XIR	1.5		221	+	561.89	387.11	68,9
					1VR2	2.4	2.864	422	+++	658.96	386.68	58,7
EGFR	279	64.5	475	10	2J5E	3.1		258	++	590	381.75	64.7
					2ITX	2.98	0.794	278	++++	598.66	390.75	65.3
					2RGP	2	2.333	282	+	466.08	314.1	67.4
					3POZ	1.5	2.404	299	+++	462.56	281.29	60.8
SAHH	259	66.3	33	2	1A7A	2.8		177	+	522.19	329.13	63.0
					1LI4	2.01	0.349	341	+++	363.05	252.56	69.6
TK	146	67.4	22	4	1E2P	2.5		114	+	275.68	200.36	72.7
					1E2I	1.9	7.991	128	+++	467.39	306.21	65.5
					1E2N	2.2	2.568	195	+	510.22	326.11	63.9

Table 1 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen inférieur à 350Å³ (petits) et un pourcentage moyen d'hydrophobicité inférieur à 70% (polaires).

Small and Hydrophobic												
	Mean volume (Å ³)	Mean % hydrophobicity	# DUD actives	# Available PDB Structures	PDB ID	Resolution (Å)	RMSD (Å)	Volume (Å ³)	Relative opening	SASA (Å ²)	ASASA (Å ²)	% Hydrophobicity
COX2	264	71.0	426	6	3PGH	2.5		248	++	757.73	523.27	69.1
					1CX2	3	0.832	264	+++	802.35	562.03	70.1
					1DDX	3	1.221	280	+	893.57	660.19	73.9
GART	320	71.0	40	6	1C2T	2.1		269	+	797.69	561.75	70.4
					1GAR	1.96	2.561	293	+++	1610.82	1163.58	72.2
					1JKX	1.6	2.524	399	++	1095.99	771.79	70.4
PR	302	72.7	27	12	1A28	1.8		207	+	395.43	328.24	83.0
					2OVH	2	5.861	396	+++	793.22	495.17	62.4
COMT	171	73.5	11	12	2CL5	1.6		84	++	1046.81	755.33	72.2
					1H1D	2	3.456	157	+	1028.2	749.55	72.9
					3A7D	2.4	0.905	272	+++	1132.68	854.69	75.5
AR	280	74.5	79	48	1T7T	1.7		240		376.4	284.97	75.7
GR	315	75.3	78	7	3B66	1.65	1.006	320		404.38	296.45	73.3
					1M2Z	2.5		228	++	387.61	299.76	77.3
					3CLD	2.84	1.325	262	+	413.39	323.81	78.3
					1NHZ	2.3	1.524	317	++++	703.95	494.52	70.3
MR	173	76.5	15	3	3K22	2.1	1.059	452	+++	493.64	371.79	75.3
					2A3I	1.95		172		622.61	472.22	75.9
					2AA2	1.95	0.798	174		713.29	549.85	77.1
COX1	243	77.8	25	22	1PTH	3.4		200	++	347.91	269.87	77.6
					3KK6	2.75	5.736	256	+	431.26	334.67	77.6
					1PGG	4.5	0.527	272	+++	400.29	312.88	78.2
HIVRT	179	78.9	43	52	2JLE	2.9		133	++	547.7	410.53	74.9
					3FFI	2.6	1.717	169	++++	718.36	541.72	75.4
					1DTQ	2.8	1.937	177	+	549	478.49	87.2
					3LAM	2.76	1.751	236	+++	774.32	604.6	78.1
RXR	253	80.1	20	17	1FBY	2.25		228		649.47	509.52	78.5
					3DZY	3.1	1.207	278		649.62	531.41	81.8
ER ago	259	84.9	67	10	3ERD	2.03		245		291.99	248.15	85.0
					2P15	1.94	0.972	272		458.98	389.22	84.8
INHA	331	86.4	86	25	2X22	2.1		515	+	751.32	653.26	86.9
					1P44	2.6	3.275	553	+++	895.1	769.33	85.9

Table 2 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen inférieur à 350Å³ (petits) et un pourcentage moyen d'hydrophobicité supérieur à 70% (hydrophobes).

Large and Polar												
	Mean volume (Å ³)	Mean % hydrophobicity	# DUD actives	# Available PDB Structures	PDB ID	Resolution (Å)	RMSD (Å)	Volume (Å ³)	Relative opening	SASA (Å ²)	ASASA (Å ²)	% Hydrophobicity
AMPC	670	57.3	21	44	1FSW	1.9		592	++	748.23	397.21	53,1
					2R9W	1.8	0.494	688	+	673.43	359.79	53,4
					2HDR	2.2	0.527	672	++++	866.63	527.39	60,9
					1XGJ	1.97	2.722	728	+++	929.97	573.72	61,7
FXA	631	59.2	146	100	2RA0	2.3		415	++	797.61	460.25	57,7
					2D1J	2.2	0.829	422	+	827.86	492.57	59,5
					1LQD	2.7	0.591	478	++++	826.69	470.66	56,9
					1IQF	3.2	1.054	577	+++	879.35	551.15	62,7
HIVPR	513	59.9	62	87	1NH0	1.03		451	++	1003.31	612.23	61,0
					2PWC	1.78	0.761	465	+	880.66	523.98	59,5
					1XL2	1.5	0.322	642	+++	932.73	552.85	59,3
ACE	595	60.4	49	4	3BKL	2.18		448	+	657.23	407.95	62,1
					2OC2	2.25	0.323	664	+++	697.22	415.67	59,6
					1O86	2	0.275	672	++	684.48	407.95	59,6
SRC	515	60.6	159	4	2H8H	2.2		278	+	842.3	523.48	62,1
					1Y57	1.91	2.017	698	+++	864.17	510.5	59,1
FGFR1	611	60.7	120	4	1FGI	2.4		355	+	809.6	462.71	57,2
					1AGW	2.5	0.414	569	+++	767.8	482.73	62,9
					2FGI	2.5	0.511	909	++	1104.25	685.99	62,1
GPB	366	61.3	52	52	6GPB	2.86		174	++	269.17	181.18	67,3
					1A8I	1.78	1.281	293	+	335.56	198.76	59,2
					1K06	1.8	1.772	630	+++	414	237.28	57,3
HMGR	390	66.5	35	8	1DQA	2		314	+	842.41	576.33	68,4
					1HWL	2.1	0.45	367	+++	1164.86	764.05	65,6
					1HWJ	2.26	0.507	490	++	1157.2	757.43	65,5
ACHE	496	68.6	107	75	2C5F	2.6		378	++	569.06	392.2	68,9
					1GQS	3	0.561	434	+	621.14	423.45	68,2
					2CEK	2.2	0.685	558	++++	717.6	493.22	68,7
					1OCE	2.7	0.711	612	+++	627.45	430.46	68,6
THR	608	68.7	72	4	1DOJ	1.7		568	+	696.97	470.88	67,6
					1JWT	2.5	13.717	648	+++	782.92	547.52	69,9
PARP	508	69.7	35	6	1EFY	2.2		506	+++	640.37	450.63	70,4
					4PAX	2.8	0.637	509	+	664.49	459.34	69,1

Table 3 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen supérieur à 350Å³ (grands) et un pourcentage moyen d'hydrophobicité inférieur à 70% (polaires).

Large and Hydrophobic												
	Mean volume (Å ³)	Mean % hydrophobicity	# DUD actives	# Available PDB structures	PDB ID	Resolution (Å)	RMSD (Å)	Volume (Å ³)	Relative opening	SASA (Å ²)	ASASA (Å ²)	% Hydrophobicity
CDK2	456	70.0	72	143	3PY1	2.05		304	+	675.73	497.63	73,6
					2I40	2.8	2.6	504	+++	999.26	635.47	63,6
					2B55	1.85	1.792	560	++	972.25	708.62	72,9
P38	367	71.1	454	96	3BX5	2.4		185	+	542.64	373.81	68,9
					3FSF	2.1	1.221	442	+++	848.18	581.67	68,6
					1KV2	2.8	2.559	473	++	974.64	737.97	75,7
ADA	507	72.3	39	12	1NDW	2		360	+	544.4	382.81	70,3
					1NDV	2.3	1.357	575	+++	770.89	570.77	74,0
					2E1W	2.5	1.424	585	++	796.15	578.72	72,7
PDE5	524	74.1	88	3	1TBF	1.3		519	++	1039.01	789.05	75,9
					1T9S	2	1.073	524	+++	1079.81	774.38	71,7
					1XOZ	1.37	0.999	540	+	1045.72	780.25	74,6
PPAR	531	74.1	85	83	3OSW	2.55		332	+++	873.94	639.04	73,1
					1WM0	2.9	1.272	380	+	716.24	515.46	72,0
					2OM9	2.8	1.349	700	++++	953.04	698.15	73,3
					1FM9	2.1	1.698	710	++	899.83	702.78	78,1
ER antago	426	75.4	39	18	3ERT	1.9		358	+	794.89	599.04	75,4
					2IOK	2.4	1.826	436	+++	1074.83	783.25	72,9
					2IOG	1.6	2.025	483	++	785.79	611.62	77,8
DHFR	706	76.1	410	7	3DFR	1.7		542	++	696.07	544.38	78,2
					1DIU	NMR	0.651	672	+	686.11	542.39	79,1
					1BZF	NMR	1.185	904	+++	761.21	541.65	71,2
ALR2	356	76.2	26	5	1AH3	2.3		272	+++	624.68	493.17	78,9
					1EK0	1.48	0.618	440	+	616.48	452.66	73,4

Table 4 : Propriétés physicochimiques des systèmes dont les sites de liaison ont un volume moyen supérieur à 350Å³ (grands) et un pourcentage moyen d'hydrophobicité supérieur à 70% (hydrophobes).

Small and polar targets

		SF					ICM				
		Relative opening	EF1%	EF10%	AUC	Average time per compound (s)	EF1%	EF10%	AUC	Average time per compound (s)	
NA	1A4G	+	24.79	6.75	0.869	11.60	26.85	8.58	0.923	17.80	
	1A4O	+++	30.98	6.34	0.836	13.79	39.24	8.18	0.938	22.73	
HSP90	2CDD	++	0.00	1.63	0.665	14.53	19.22	3.53	0.627	14.77	
	3K99	+	2.75	1.90	0.593	17.66	13.73	2.99	0.544	16.10	
	2XJX	++++	5.49	1.63	0.700	15.00	13.73	3.26	0.594	15.36	
	1UYF	+++	2.75	1.09	0.498	13.76	0.00	1.90	0.631	14.69	
PNP	1B8O	+	15.20	5.43	0.755	7.60	8.69	4.02	0.635	9.90	
	2OPL	+++	15.20	3.02	0.638	5.64	10.86	5.23	0.788	7.19	
TRP	1V48	++	0.00	1.81	0.561	7.07	0.00	0.40	0.590	7.16	
	1BTY	+	16.45	7.16	0.933	27.92	2.06	3.68	0.747	47.07	
	3AAU	+++	2.06	3.27	0.755	25.81	0.00	2.04	0.675	45.25	
VEGFR2	1V2O	++	26.73	9.00	0.973	26.76	4.11	7.16	0.882	50.71	
	2XIR	+	11.73	3.41	0.576	18.08	14.08	3.53	0.609	15.76	
	1VR2	+++	4.69	1.48	0.560	10.00	3.52	2.84	0.625	17.40	
EGFR	2J5E	++	8.03	2.23	0.536	9.70	26.01	5.22	0.750	11.60	
	2ITX	++++	9.30	2.44	0.493	10.02	7.40	4.00	0.713	12.25	
	2RGP	+	8.46	2.48	0.558	11.88	5.50	3.14	0.641	10.93	
	3POZ	+++	8.88	2.08	0.494	11.88	4.02	2.82	0.609	13.12	
SAHH	1A7A	+	0.00	3.05	0.811	8.90	0.00	3.97	0.751	11.30	
	1LI4	+++	5.87	1.59	0.497	13.00	9.64	3.05	0.643	10.00	
TK	1E2P	+	9.22	3.19	0.593	8.87	0.00	0.46	0.534	6.52	
	1E2I	+++	13.83	2.74	0.643	8.28	4.61	2.74	0.607	6.80	
	1E2N	+	0.00	1.37	0.555	7.49	0.00	0.91	0.478	6.42	

Table 5 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surflex-dock et ICM pour chaque structure des systèmes petits et polaires.

Small and hydrophobic targets

		SF					ICM				
		Relative opening	EF1%	EF10%	AUC	Average time per compound (s)	EF1%	EF10%	AUC	Average time per compound (s)	
COX2	3PGH	++	3.77	1.15	0.573	11.56	0.00	0.16	0.530	12.58	
	1CX2	+++	3.76	3.50	0.785	12.20	0.24	1.29	0.560	11.16	
	1DDX	+	5.87	1.71	0.574	11.80	0.71	1.41	0.500	12.21	
GART	1C2T	+	2.55	5.55	0.880	25.60	10.21	5.30	0.780	40.15	
	1GAR	+++	2.55	6.56	0.916	21.87	17.87	6.06	0.861	43.12	
	1JKX	++	0.00	4.29	0.861	23.31	2.55	2.02	0.692	46.91	
PR	1A28	+	7.91	2.99	0.667	8.80	31.64	5.22	0.738	8.80	
	2OVH	+++	7.91	1.12	0.419	8.31	27.69	4.10	0.609	9.73	
COMT	2CL5	++	0.00	1.85	0.716	4.26	0.00	1.85	0.712	6.33	
	1H1D	+	21.77	3.71	0.733	4.50	0.00	2.78	0.700	5.99	
	3A7D	+++	32.66	2.78	0.491	6.39	0.00	0.93	0.501	5.84	
AR	1T7T		26.88	4.18	0.726	1.43	21.76	4.94	0.731	7.98	
	3B66		3.84	0.63	0.452	1.92	12.80	3.42	0.572	7.93	
GR	1M2Z	++	9.05	3.08	0.564	12.60	10.34	2.57	0.450	11.70	
	3CLD	+	2.59	1.67	0.511	13.51	11.63	1.67	0.591	12.96	
	1NHZ	++++	1.29	0.77	0.478	11.30	2.59	0.51	0.340	11.72	
	3K22	+++	0.00	0.39	0.209	14.10	0.00	0.13	0.230	11.49	
MR	2A3I		28.93	5.34	0.757	12.40	36.17	6.68	0.788	12.90	
	2AA2		36.17	6.00	0.796	13.90	43.40	5.34	0.767	13.10	
COX1	1PTH	++	0.00	1.21	0.507	8.72	0.00	3.22	0.734	8.08	
	3KK6	+	4.16	1.21	0.503	8.10	0.00	1.61	0.644	9.43	
	1PGG	+++	4.16	2.01	0.521	8.53	4.16	3.22	0.670	8.03	
HIVRT	2JLE	++	9.69	2.56	0.593	11.22	9.69	1.63	0.515	12.45	
	3FFI	++++	7.27	2.33	0.547	10.72	2.42	1.16	0.516	12.20	
	1DTQ	+	7.27	1.86	0.628	10.95	0.00	1.86	0.559	12.20	
	3LAM	+++	2.42	2.33	0.578	10.45	7.27	3.03	0.657	11.62	
RXR	1FBY		5.50	8.50	0.941	26.49	5.50	7.00	0.922	27.00	
	3DZY		0.00	0.00	0.444	16.29	0.00	5.00	0.750	27.28	
ER ago	3ERD		7.57	2.69	0.646	8.10	22.71	5.69	0.746	8.51	
	2P15		9.08	2.54	0.613	7.00	12.11	5.24	0.730	7.36	
INHA	2X22	+	4.72	1.86	0.550	13.84	8.27	1.51	0.575	11.77	
	1P44	+++	4.07	2.57	0.550	10.90	1.16	2.33	0.525	12.80	

Table 6 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surflex-dock et ICM pour chaque structure des systèmes petits et hydrophobes.

Large and polar targets										
		SF					ICM			
		Relative opening	EF1%	EF10%	AUC	Average time per compound (s)	EF1%	EF10%	AUC	Average time per compound (s)
AMPC	1FSW	++	0.00	1.92	0.656	6.60	4.80	1.44	0.490	6.74
	2HDR	++++	0.00	0.96	0.593	5.72	0.00	0.48	0.530	7.22
	2R9W	+	0.00	0.96	0.643	5.28	0.00	1.44	0.517	7.66
	1XGJ	+++	0.00	1.44	0.688	5.70	14.41	3.84	0.530	6.62
FXA	2RA0	++	18.09	6.99	0.877	21.87	14.61	6.65	0.865	42.04
	2D1J	+	8.35	4.25	0.778	22.04	5.57	2.47	0.714	37.80
	1LQD	++++	25.74	6.16	0.845	22.30	25.74	7.60	0.913	37.90
HIVPR	1IQF	+++	9.74	4.18	0.743	22.14	22.96	5.41	0.855	37.79
	1NH0	++	9.68	4.35	0.728	35.14	3.23	2.26	0.602	55.71
	2PWC	+	10.16	2.92	0.713	34.09	1.61	1.45	0.550	55.21
ACE	1XL2	+++	3.23	1.61	0.649	35.51	0.00	0.00	0.685	58.95
	3BKL	+	4.19	1.84	0.434	15.00	2.09	1.64	0.580	18.98
SRC	2OC2	+++	2.09	1.43	0.402	13.39	4.19	1.64	0.502	18.84
	1O86	++	2.09	2.66	0.476	12.90	16.74	4.09	0.660	17.23
FGFR1	2H8H	+++	8.91	3.15	0.542	21.21	14.64	3.97	0.668	19.14
	1Y57	+	8.91	2.33	0.555	15.94	12.10	4.72	0.742	19.47
GPB	1FGI	+	5.08	1.58	0.444	19.45	2.54	1.33	0.400	23.08
	1AGW	+++	5.92	2.42	0.568	17.90	1.69	0.75	0.646	25.04
	2FGI	++	6.77	2.67	0.598	17.96	3.38	3.58	0.614	26.16
HMGR	6GPB	++	2.00	3.27	0.747	11.00	4.01	4.43	0.752	17.56
	1A8I	+	4.01	2.12	0.674	9.60	8.03	7.51	0.835	13.64
	1K06	+++	0.00	1.15	0.492	9.60	2.01	1.35	0.513	14.44
ACHE	1DQA	+	0.00	1.15	0.470	25.11	2.89	0.86	0.668	36.76
	1HWL	+++	37.51	6.02	0.827	21.43	32.65	4.42	0.681	34.38
	1HWJ	++	37.51	6.31	0.858	24.16	43.29	4.87	0.792	36.67
THR	2C5F	++	0.00	0.37	0.424	15.66	0.00	0.94	0.582	20.38
	1GQS	+	0.96	0.94	0.436	17.18	0.00	0.56	0.683	21.55
	2CEK	++++	3.83	2.15	0.549	16.68	0.96	1.22	0.505	21.76
	1OCE	+++	0.96	0.47	0.436	14.96	3.83	2.15	0.581	22.85
PARP	1DOJ	+	15.45	6.27	0.793	31.28	1.40	4.18	0.797	41.65
	1JWT	+++	1.40	1.95	0.611	28.81	7.02	3.62	0.631	47.44
PARP	1EFY	+++	18.28	5.74	0.738	4.70	3.05	5.74	0.756	5.10
	4PAX	+	9.14	3.73	0.615	4.80	12.18	4.59	0.615	5.82

Table 7 : Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surfex-dock et ICM pour chaque structure des systèmes grands et polaires.

Large and hydrophobic targets

		SF					ICM				
		Relative opening	EF1%	EF10%	AUC	Average time per compound (s)	EF1%	EF10%	AUC	Average time per compound (s)	
CDK2	3PY1	+	2.84	2.65	0.608	15.38	1.42	1.25	0.583	14.54	
	2I40	+++	0.00	1.11	0.515	11.66	5.68	3.34	0.669	14.48	
	2B55	++	5.68	2.37	0.626	11.52	14.19	4.18	0.705	14.52	
P38	3BX5	+	0.67	0.60	0.488	13.75	0.00	0.04	0.318	11.38	
	3FSF	+++	1.78	1.90	0.585	12.71	0.00	0.13	0.414	11.00	
	1KV2	++	5.12	2.20	0.650	11.00	4.45	0.95	0.367	10.80	
ADA	1NDW	+	16.51	3.10	0.700	8.80	0.00	0.00	0.320	10.77	
	1NDV	+++	5.50	1.81	0.580	8.07	0.00	1.29	0.549	11.24	
	2E1W	++	13.76	3.35	0.624	7.70	0.00	0.77	0.595	11.05	
PDE5	1TBF	++	4.70	1.82	0.565	20.56	17.61	4.56	0.791	28.87	
	1T9S	+++	12.91	3.65	0.662	20.94	7.04	3.42	0.704	27.16	
	1XOZ	+	5.87	3.53	0.707	23.06	4.70	2.62	0.705	26.07	
PPAR	3OSW	+++	1.18	3.06	0.737	37.98	0.00	0.59	0.516	54.50	
	1WM0	+	0.00	0.71	0.579	41.73	1.18	0.59	0.556	57.12	
	2OM9	++++	1.18	2.83	0.745	36.24	1.18	2.00	0.560	54.10	
	1FM9	++	27.16	7.65	0.901	37.90	16.53	5.18	0.748	51.10	
ER antago	3ERT	+	16.34	5.15	0.757	25.50	13.62	3.09	0.630	32.00	
	2IOK	+++	8.17	4.38	0.748	25.60	19.06	6.18	0.787	33.85	
	2IOG	++	5.45	3.86	0.785	25.62	5.45	2.83	0.591	34.53	
DHFR	3DFR	++	18.45	4.32	0.716	12.50	4.19	2.94	0.620	15.36	
	1DIU	+	5.41	1.22	0.482	11.67	6.15	2.03	0.560	15.39	
	1BZF	+++	11.32	3.59	0.731	11.01	0.49	0.71	0.552	15.82	
ALR2	1AH3	+++	27.49	3.85	0.537	5.20	3.93	2.69	0.650	6.00	
	1EK0	+	19.63	3.85	0.545	4.88	0.00	0.77	0.525	6.47	

Table 8 :Facteurs d'enrichissement à 1% et 10% et ROC AUC obtenus par Surflex-dock et ICM pour chaque structure des systèmes grands et hydrophobes.

II.1.3 Les courbes de prédictibilité comme métrique d'évaluation des méthodes de criblage *in silico*

Il existe beaucoup de méthodes de criblage *in silico*, ce qui a conduit à un travail considérable pour déterminer les meilleures manières de les évaluer^{200; 201}, notamment concernant l'évaluation de leur capacité à augmenter la concentration d'actifs dans les premiers pourcentages de la banque, classée à l'aide d'une fonction de score¹⁶⁴.

La sélection d'un seuil optimal pour les tests expérimentaux est primordial et est estimé en analysant les résultats d'évaluations rétrospectives des méthodes de criblage *in silico*. Ces évaluations consistent à tester la capacité d'une méthode à discriminer, pour une cible particulière, les composés actifs des composés inactifs¹⁶². La métrique d'évaluation doit donc être choisie avec soin, que ce soit pour l'évaluation d'un protocole de criblage ou pour comparer plusieurs méthodes entre elles¹⁶⁴. Différentes métriques sont apparues : le facteur d'enrichissement, la courbe d'enrichissement, la courbe de ROC et son aire sous la courbe (ROC AUC), parmi les plus utilisées. D'autres métriques ont été proposées, basées sur les précédentes, mais essayant de les améliorer : l'aire partielle sous la courbe de ROC (pROC AUC)²⁰², BEDROC ou encore RIE. Toutes ces métriques ont été longuement étudiées et leurs forces et faiblesses débattues^{203; 204}, mais aucune métrique standard n'a encore été définie. Le facteur et la courbe d'enrichissement sont très utilisés car ils représentent le but même du criblage *in silico*, mais sont dépendants du taux d'actifs dans la banque²⁰¹. De plus, ils ne prennent en compte qu'un seul type d'erreur de classification, la sensibilité (taux de vrais positifs)¹⁶².

D'un autre côté, les courbes de ROC représentent à la fois la spécificité (taux de vrais négatifs) et la sensibilité. De plus, elles sont indépendantes du taux d'actifs présents dans la banque d'évaluation. Cependant, toutes les valeurs de sensibilité ont le même poids, qu'elles correspondent à une spécificité faible ou élevée.

Dans cette étude, la courbe de prédictibilité est utilisée pour tenter de compléter la ROC AUC dans la comparaison des performances de différentes méthodes. Il serait en effet intéressant d'évaluer la capacité d'une méthode à prédire la probabilité d'une molécule d'être active plutôt qu'inactive. Cette courbe représente la distribution d'une population de risques prédits par un marqueur ou un modèle de prédiction de risque, ici le score attribué à chaque composé. La

métrique est décrite par Huang et Pepe²⁰⁵ de la manière suivante. Supposons que A représente l'activité d'un composé, avec $A=1$ pour un actif et $A=0$ pour un inactif. Y représente la fonction de score étudiée et le risque calculé à partir du score devient $Risk(Y)=P(A=1|Y)$. En supposant que le risque $Risk(Y)$ augmente de manière monotone avec la fonction Y , la courbe de prédictibilité pour Y est $R(v)$ avec v compris entre 0 et 1. Ainsi, $R(v)$ est le risque correspondant au $v^{\text{ème}}$ centile de la fonction. La fonction inverse $R^{-1}(p)=P\{Risk(Y)\leq p\}$ est la proportion de composés ayant un risque d'être actifs inférieur ou égal à p . Une propriété intéressante de cette représentation est qu'elle fournit une échelle commune aux différentes fonctions, permettant de faire des comparaisons entre des fonctions qui ne seraient pas forcément comparables avec leurs échelles respectives.

Ces comparaisons peuvent être faites à partir de $R(v)$, une fonction faisant une meilleure prédiction de risque ayant tendance à avoir une plus grande variabilité de valeur. Ainsi, pour $v=0,9$, la fonction de score pour laquelle $R(v)$ est le plus élevé a une meilleure prédiction pour les risques élevés que l'autre. Deux seuils peuvent aussi être choisis : p_L (tous les composés sous ce seuil seront considérés comme inactifs) et p_H (tous les composés au-dessus de ce seuil seront considérés actifs). La fonction de score ayant fait la meilleure prédiction est celle pour laquelle le moins de composés ont un risque compris entre les deux seuils.

Le gain total a été proposé par Bura et Gastwirth en 2001²⁰⁶ comme mesure globale de la courbe de prédictibilité :

$$TG = \int_0^1 |R(v) - \rho| dv$$

c'est à dire l'aire présente entre la courbe de prédictibilité étudiée et la courbe correspondant à la prévalence ρ en actifs de la banque. Cette courbe représente donc une fonction de score non informative, toutes les molécules ayant la même probabilité d'être actives.

Deux méthodes, Surflex-dock et ICM, ont été évaluées sur les 39 cibles dont une structures expérimentale était disponibles dans la DUD, et les résultats sont représentés avec les courbes de ROC et de prédictibilité. Pour chaque cible et chaque méthode, la distribution des scores obtenus a été représentée par une courbe de densité.

MATERIELS ET METHODES

DUD. La Directory of Useful Decoys¹⁶⁰ est une banque d'évaluation contenant 40 systèmes. Chaque système est composé d'une structure cristallographique de la protéine cible et du ligand cocristallisé, d'actifs connus et, pour chaque actif, de 36 inactifs sélectionnés pour leurs propriétés physicochimiques similaires aux actifs. Ils sont en revanche structurellement différents. Tous les systèmes, pour lesquels une structure expérimentale de la cible était disponible, ont été retenus.

Surflex-dock. L'algorithme d'exploration de l'espace conformationnel de Surflex-dock est dérivé de l'algorithme de fragmentation / reconstruction de Hammerhead¹¹³. Le site de liaison est représenté par un protomol, c'est à dire des fragments moléculaires (NH, C=O, CH₄) recouvrant tout le site de liaison. Les poses sont évaluées par une fonction de score empirique. Pour chaque structure, le site de liaison a été défini à 4 Å autour du ligand co-cristallisé. La version 2.5 de Surflex-dock a été utilisée pour cette étude.

ICM. ICM¹¹⁵ est basé sur l'algorithme de Métropolis Monte Carlo pour optimiser le positionnement des molécules dans le site de liaison. Les poses sont évaluées avec la fonction de score empirique ICM-VLS. La version 3.6 d'ICM a été utilisée pour cette étude.

Métriques des performances. Tous les graphes ont été faits avec l'outil statistique et graphique R. Le module ROCR a été utilisé pour tracer les courbes de ROC. Les courbes de prédictibilité ont été tracées avec le module PredictABEL. Contrairement aux courbes de ROC en d'enrichissement qui représentent la banque classée par une fonction de score de 0 à 1 (ou de 0 à 100), la courbe de prédictibilité représente la banque de 1 à 0. Une droite au niveau de la prévalence des actifs a été tracée. Elle correspond à une distribution aléatoire des actifs dans la banque. Une échelle logarithmique pour la probabilité prédite a été utilisée car la prévalence en actifs des banques est trop faible pour pouvoir bien observer l'allure des courbes.

RESULTATS

Les fonctions de score des méthodes de criblages in silico sont censées donner des scores plus favorables aux actifs qu'aux inactifs. Pour Surflex-dock, les actifs sont censés avoir un score plus élevé tandis que la fonction de score de ICM doit leur donner un score plus négatif. Les

courbes de densité des distributions de scores permettent d'illustrer cette hypothèse et de montrer que ce n'est pas systématique. En effet, les deux distributions sont décalées en faveur des actifs pour un grand nombre de systèmes, que ce soit avec ICM (GART, GPB, HMGR, NA, PARP, SRC, Trypsin) ou Surflex-dock (AMPC, AR, FXA, GART, HIVPR, HMGR, PPAR gamma). En revanche, celles-ci se chevauchent (avec ICM et Surflex-dock) ou sont décalées (seulement avec ICM) en faveur des inactifs pour de nombreux autres systèmes (ADA, AMPC, GR, HSP90, INHA, P38 pour ICM ; ACE, ACHE, GR, HIVRT, INHA, PR pour Surflex-dock). Plusieurs pics sont aussi visibles sur certaines distributions de score (par exemple ADA, AMPC, TK, Trypsin) qui pourraient être le résultat de la présence de plusieurs chemotypes parmi les actifs. Cependant, ces pics ne sont pas toujours présents avec les deux méthodes de docking. Ces distributions de score donnent donc un premier aperçu de la qualité de l'enrichissement d'une banque en actifs. Mais elles ne permettent pas de faire des comparaisons de performances entre méthodes, les échelles des scores n'étant pas les mêmes.

Les courbes de ROC et d'enrichissement ont été tracées en tant que référence. Ces courbes permettent non seulement de quantifier l'enrichissement en actifs d'une banque par une fonction de score, mais également de comparer les performances de deux fonctions de score. Il apparaît pour certains systèmes que les deux métriques donnent des résultats contradictoires ou qu'il soit difficile de déterminer quelle méthode fonctionne le mieux au début de la banque. Par exemple, pour VEGFR2, les courbes de ROC et d'enrichissement montrent que ICM fonctionne mieux sur l'ensemble de la banque. Mais il est difficile de déterminer quelle méthode fonctionne le mieux dans le premier pourcentage, alors que l'enrichissement dans le premier pourcentage de la banque est ce qui sera testé expérimentalement par la suite. La courbe de prédictibilité permet de déterminer que ICM a effectivement un plus grand pouvoir prédictif au tout début de la banque. Un autre exemple est CDK2 : les courbes d'enrichissement montrent de meilleurs résultats pour Surflex-dock au tout début de la banque, tandis que les courbes de ROC mettent en évidence une meilleure discrimination de la part de ICM. Les courbes de prédictibilité appuient les courbes de ROC, ICM ayant une capacité de prédiction plus grande, même au début de la banque.

L'utilisation des courbes de prédictibilité pourrait apporter un angle nouveau lorsque les courbes de ROC et d'enrichissement ne donnent pas de résultats clairs ou allant dans le même

sens. En effet, contrairement aux courbes de ROC et d'enrichissement qui représentent la sensibilité et/ou la spécificité d'une méthode, la courbe de prédictibilité représente la capacité d'une méthode à prédire le caractère actif d'une molécule. De plus, cette fonction de probabilité permet aussi d'avoir une échelle commune pour deux fonctions de score, permettant une comparaison directe basée sur le score attribué à une molécule, et non plus seulement sur son rang. Enfin, le gain total pourrait être utilisé de la même manière que la ROC AUC, en tant que mesure globale de la courbe de prédictibilité.

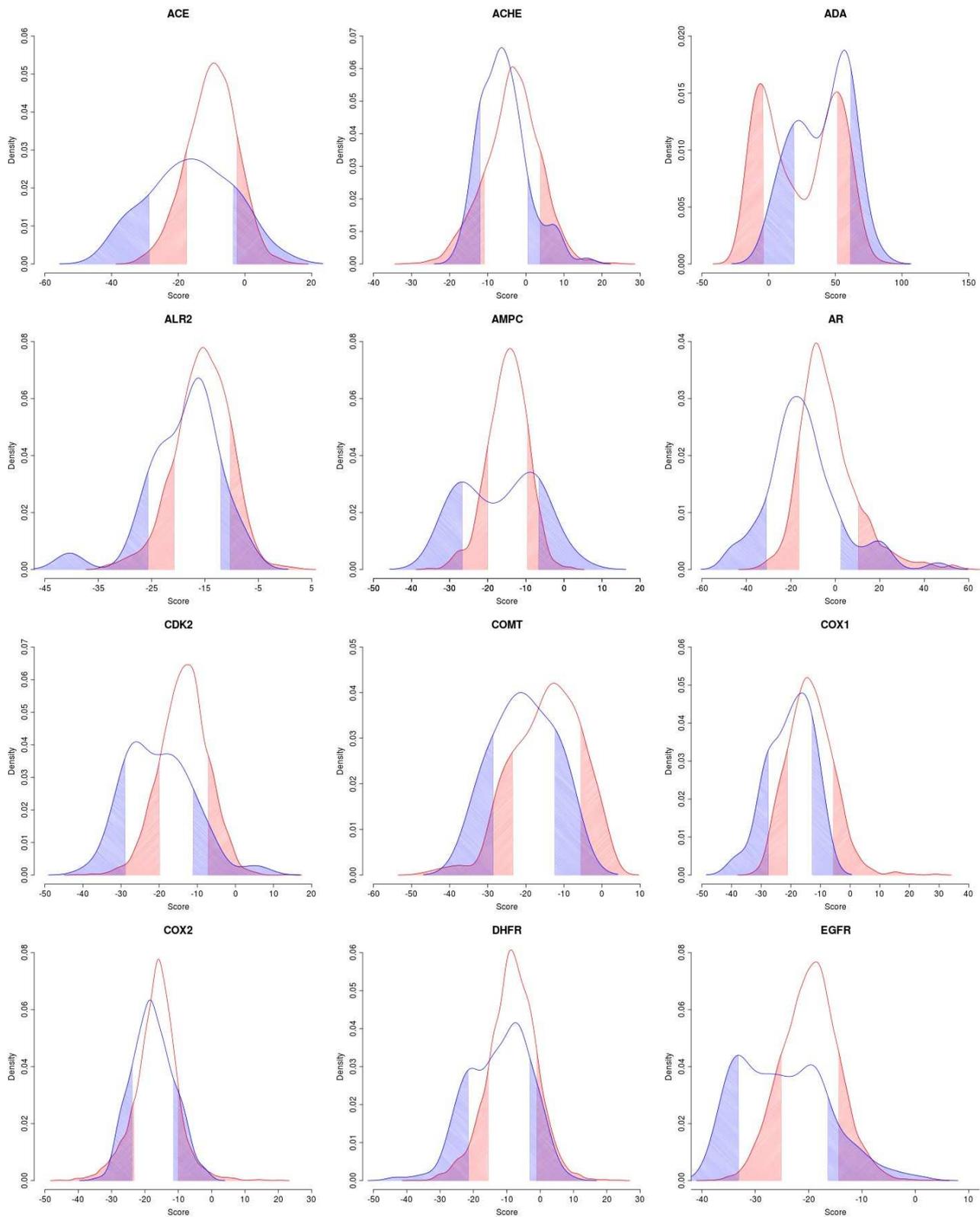


Figure 11 : Courbes de densité des distribution de scores attribués aux actifs (bleus) et aux inactifs (rouges) par ICM pour les 39 systèmes de la DUD. Pour chaque distribution, la partie hachurée représente les scores se trouvant au-delà de l'écart-type à la moyenne.

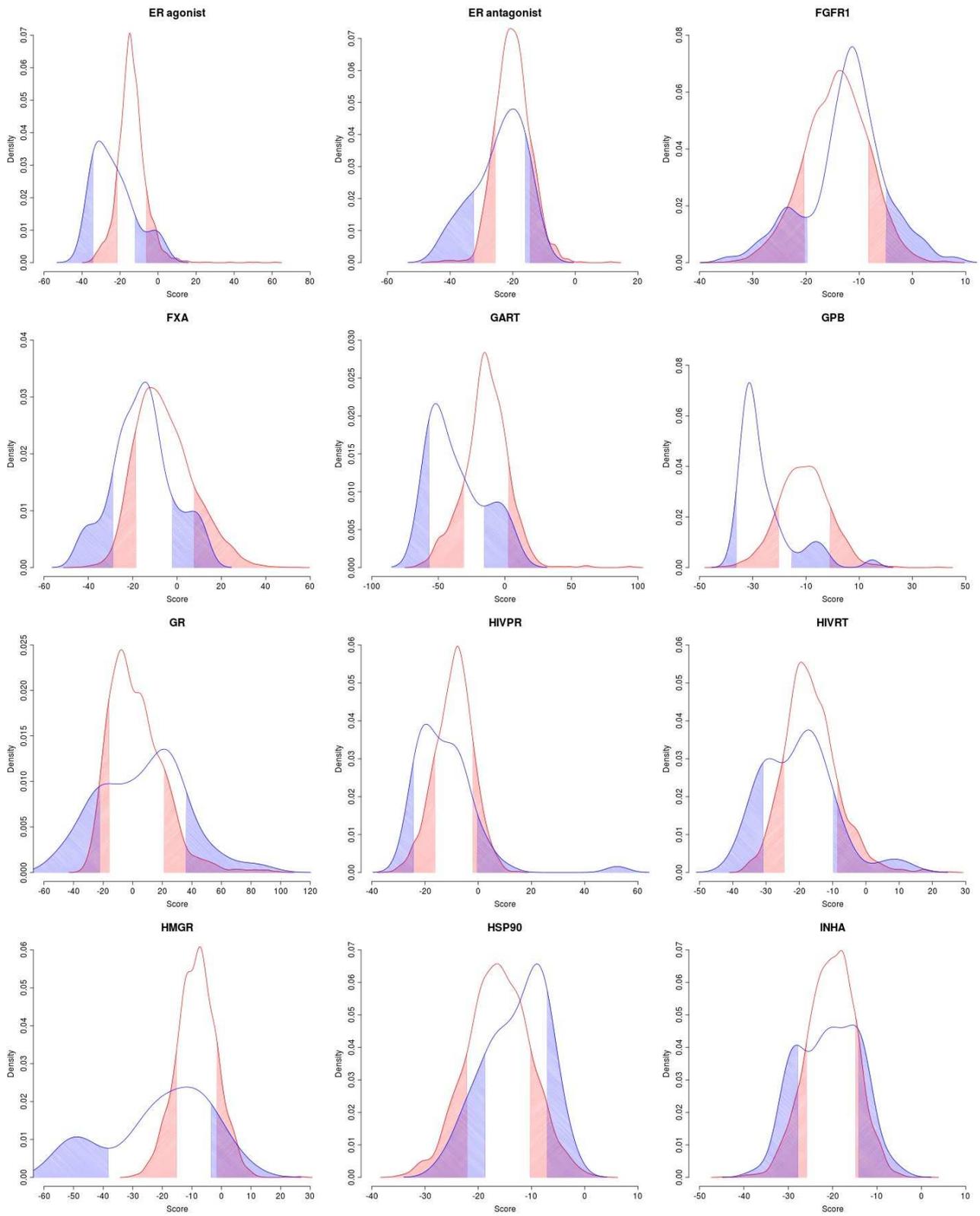


Figure 11 (suite)

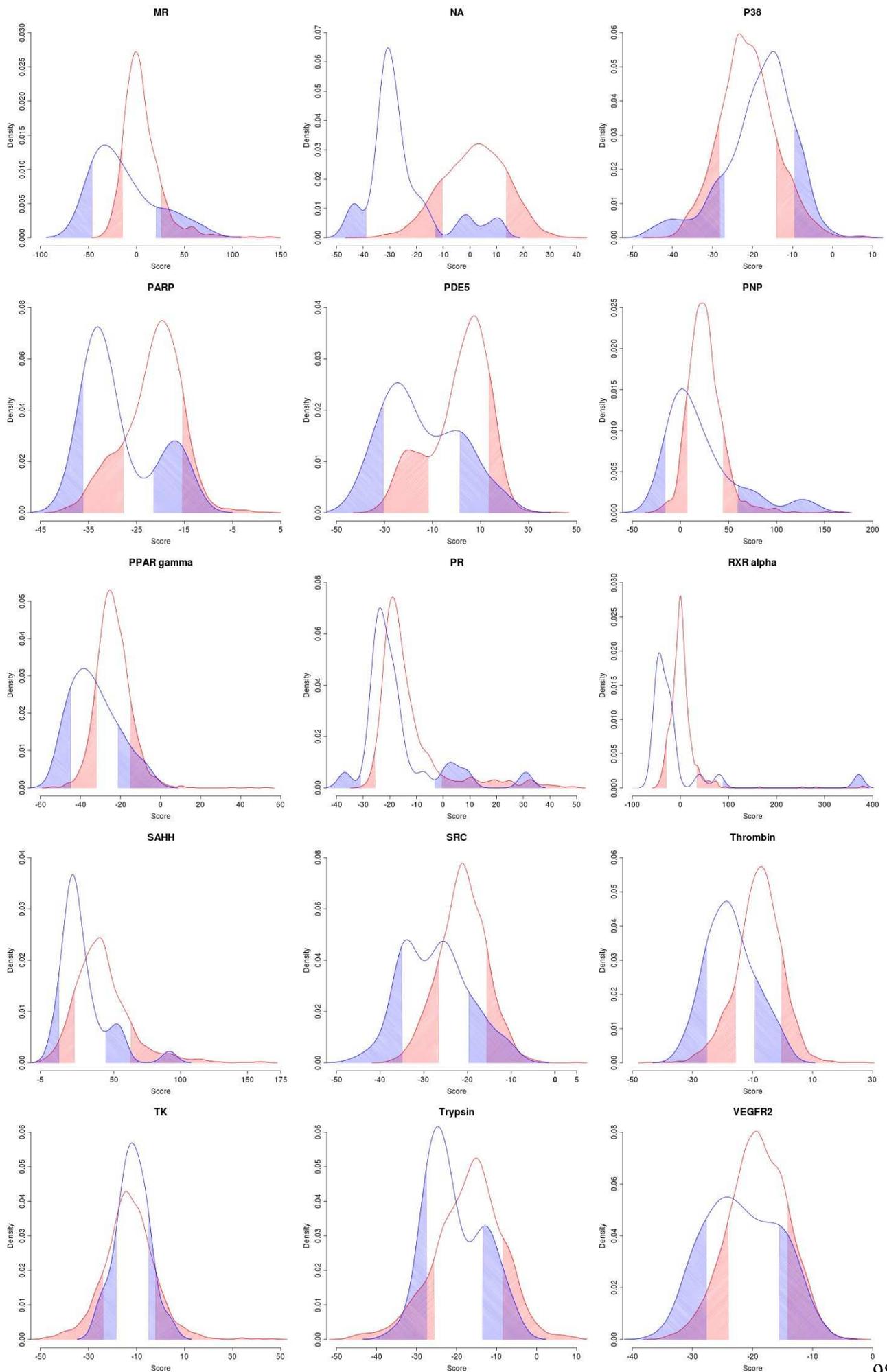


Figure 11 (fin)

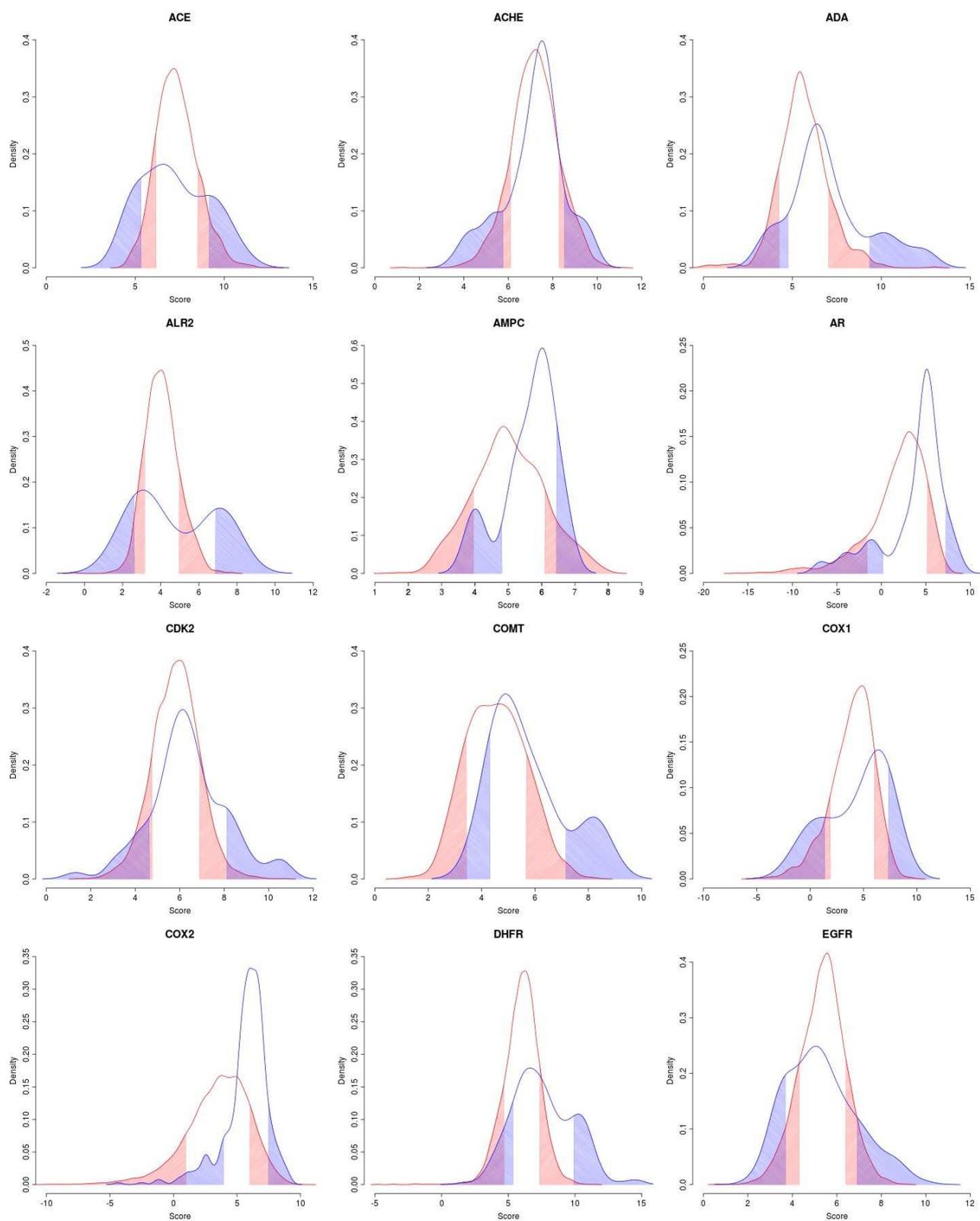


Figure 12 : Courbes de densité des distribution de scores attribués aux actifs (bleus) et aux inactifs (rouges) par Surflex-dock pour les 39 systèmes de la DUD. Pour chaque distribution, la partie hachurée représente les scores se trouvant au-delà de l'écart-type à la moyenne.

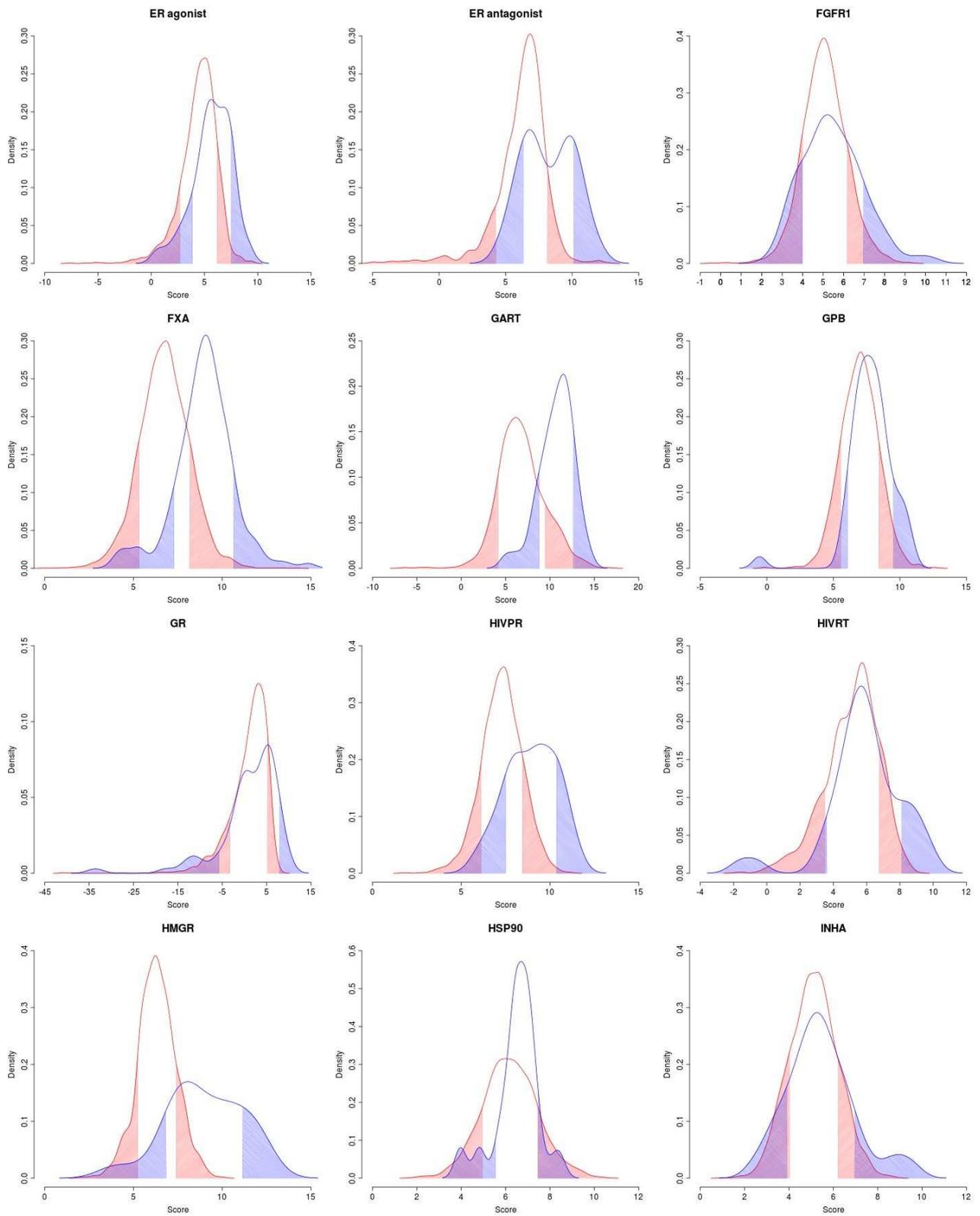


Figure 12 (suite)

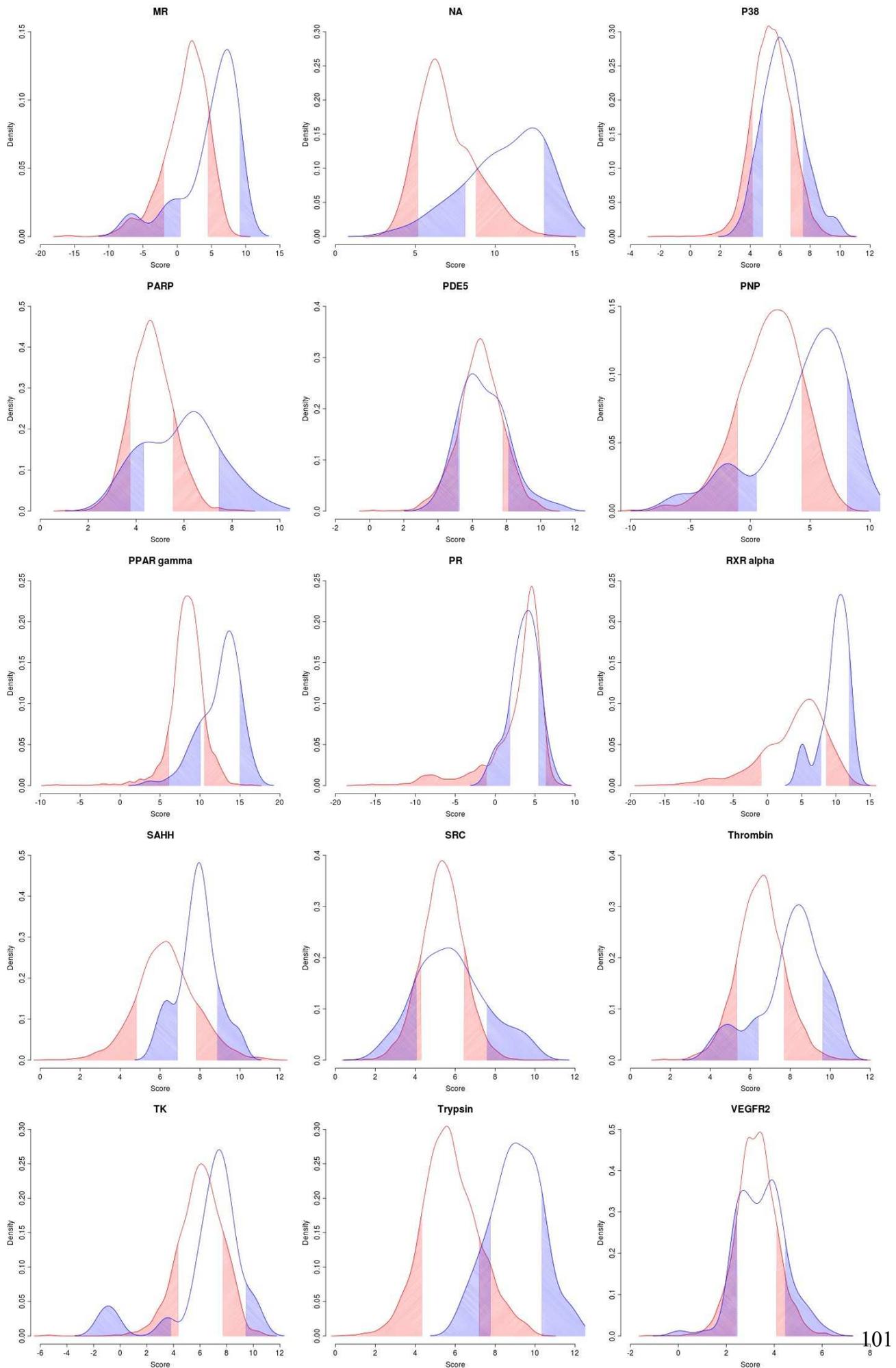


Figure 12 (fn)

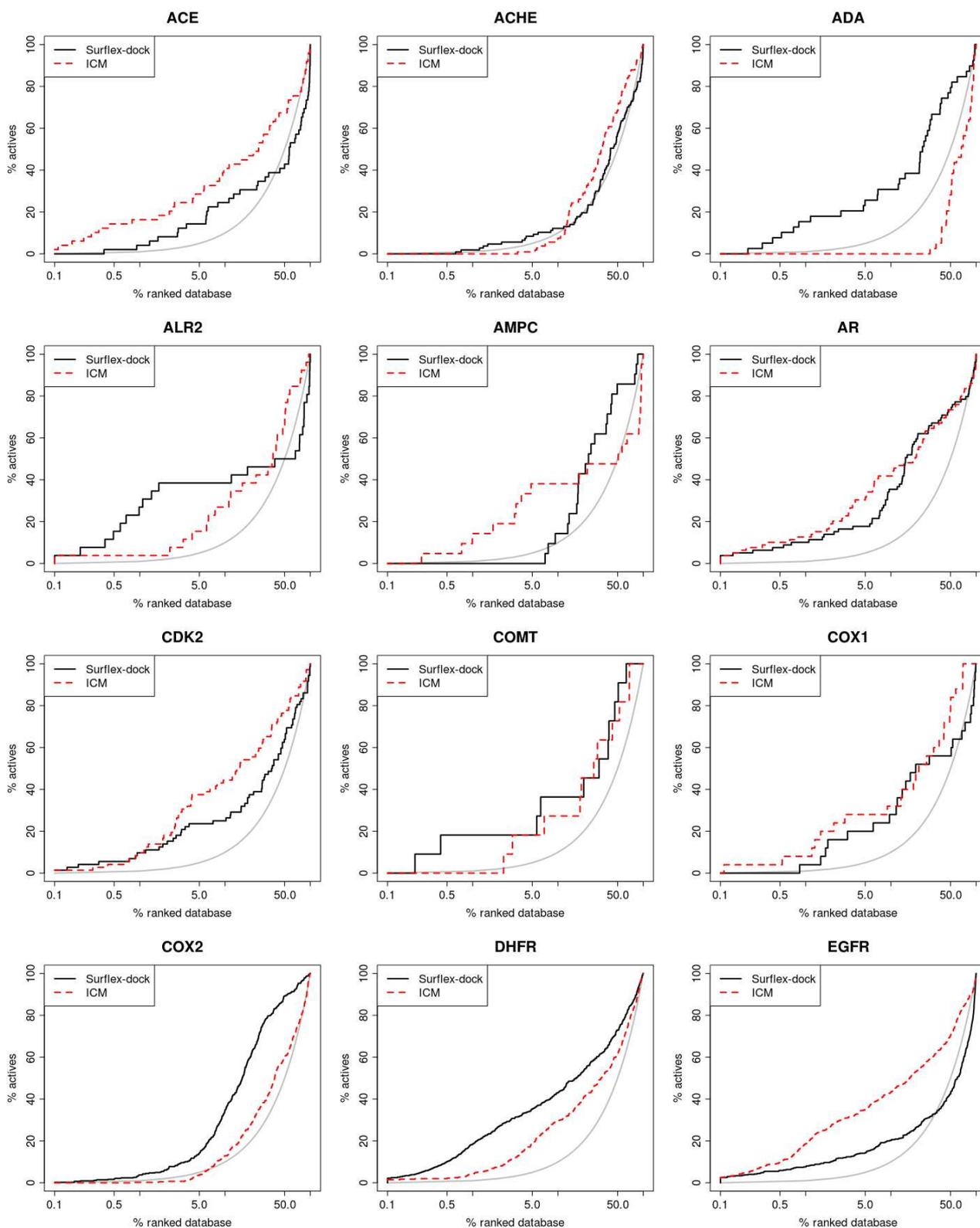


Figure 13 : Courbes d'enrichissement en semi-log pour Surflex-dock (trait plein noir) et ICM (pointillés rouges) sur les 39 systèmes de la DUD. La distribution homogène des actifs est représentée en gris.

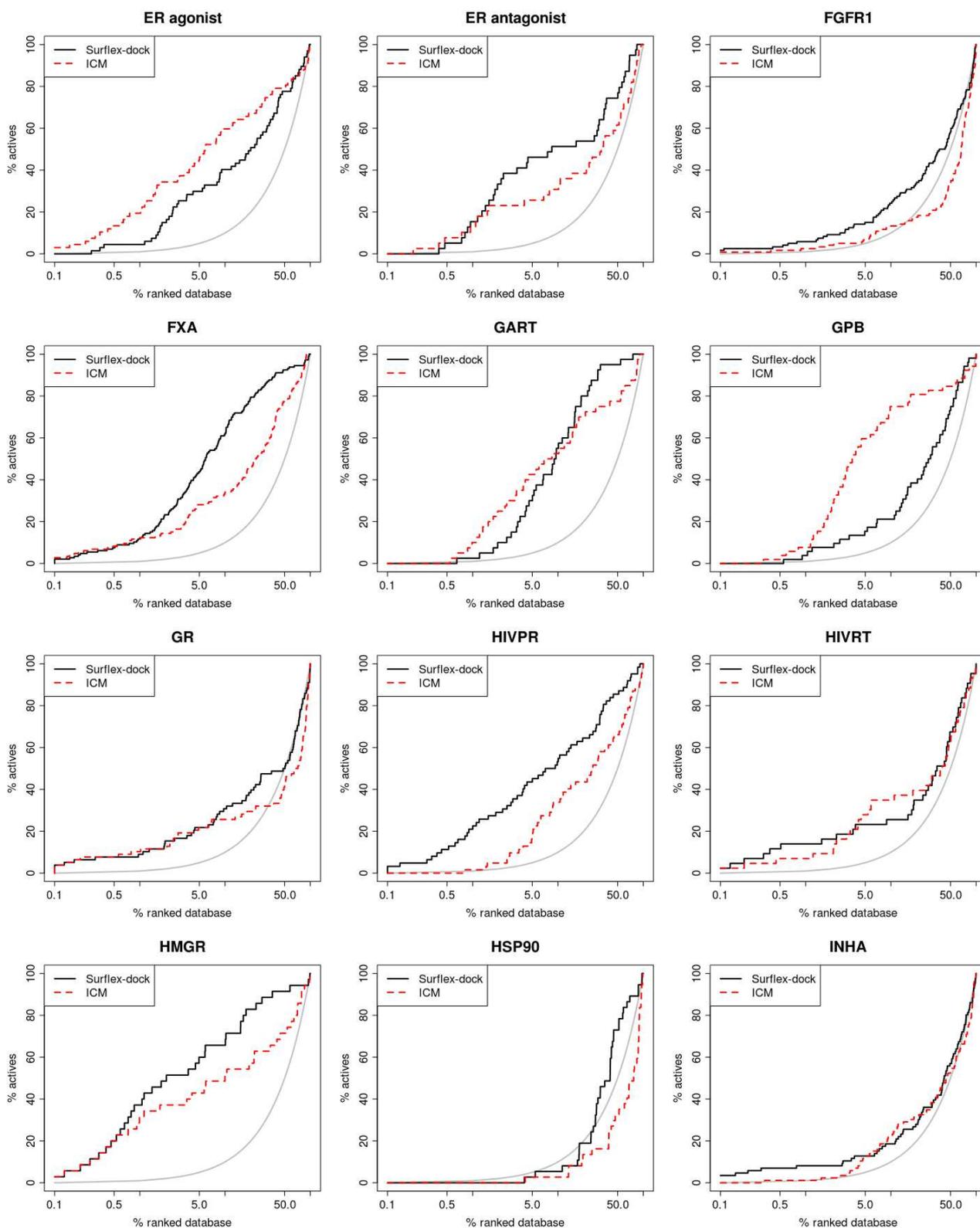


Figure 13 (suite)

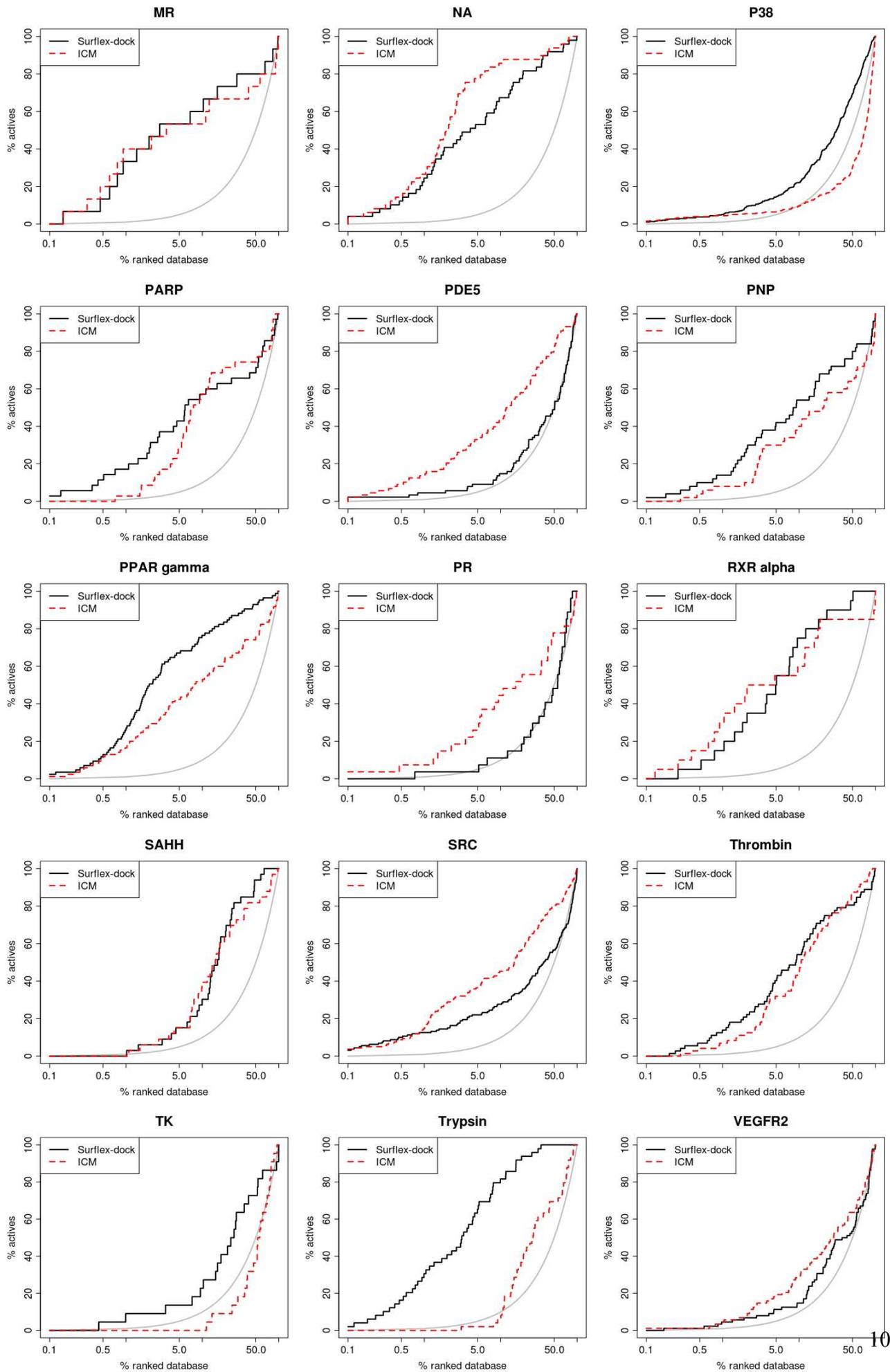


Figure 13 (fin)

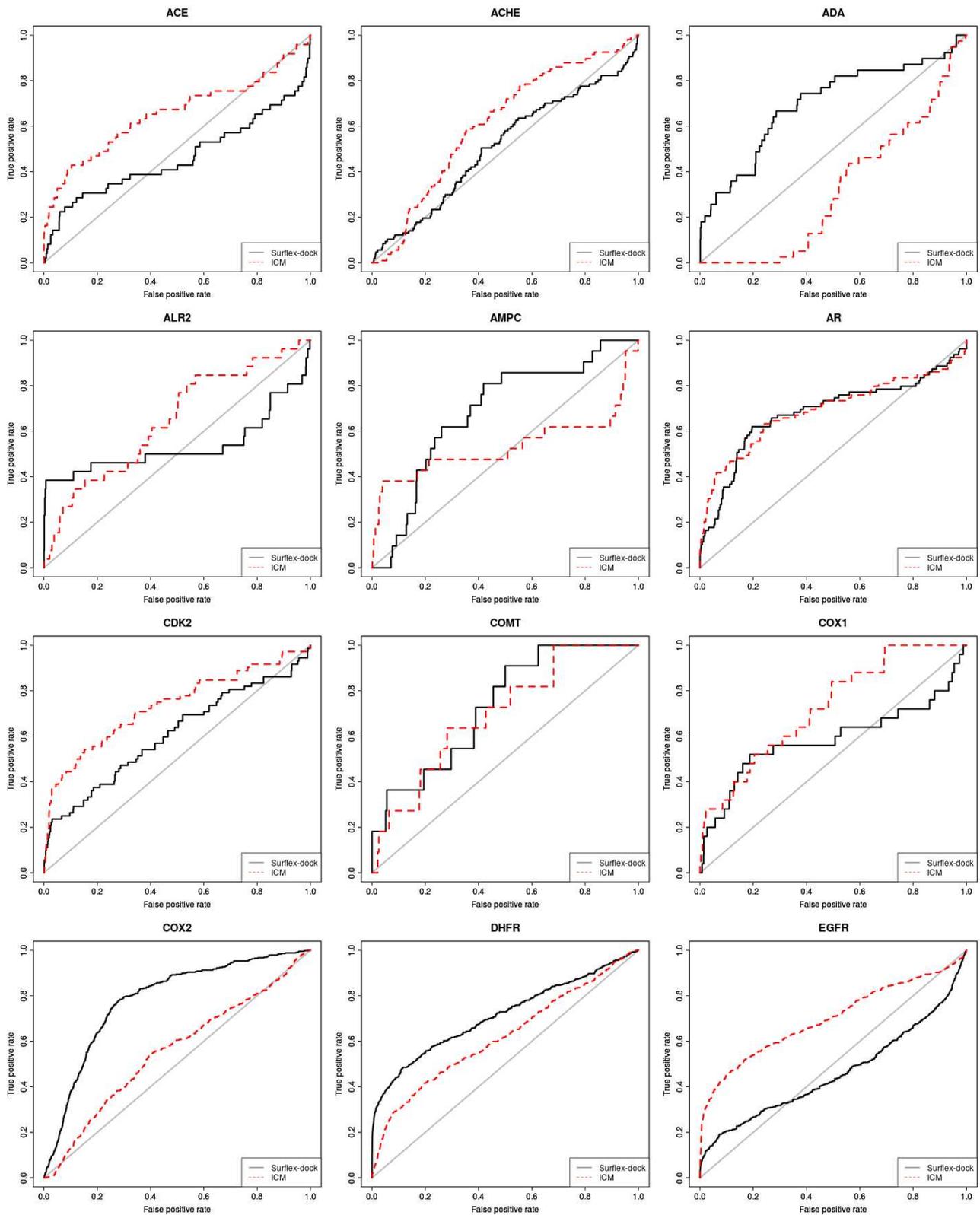


Figure 14 : Courbes de ROC pour Surflex-dock (trait plein noir) et ICM (pointillés rouges) sur les 39 systèmes de la DUD. La distribution homogène des actifs est représentée en gris.

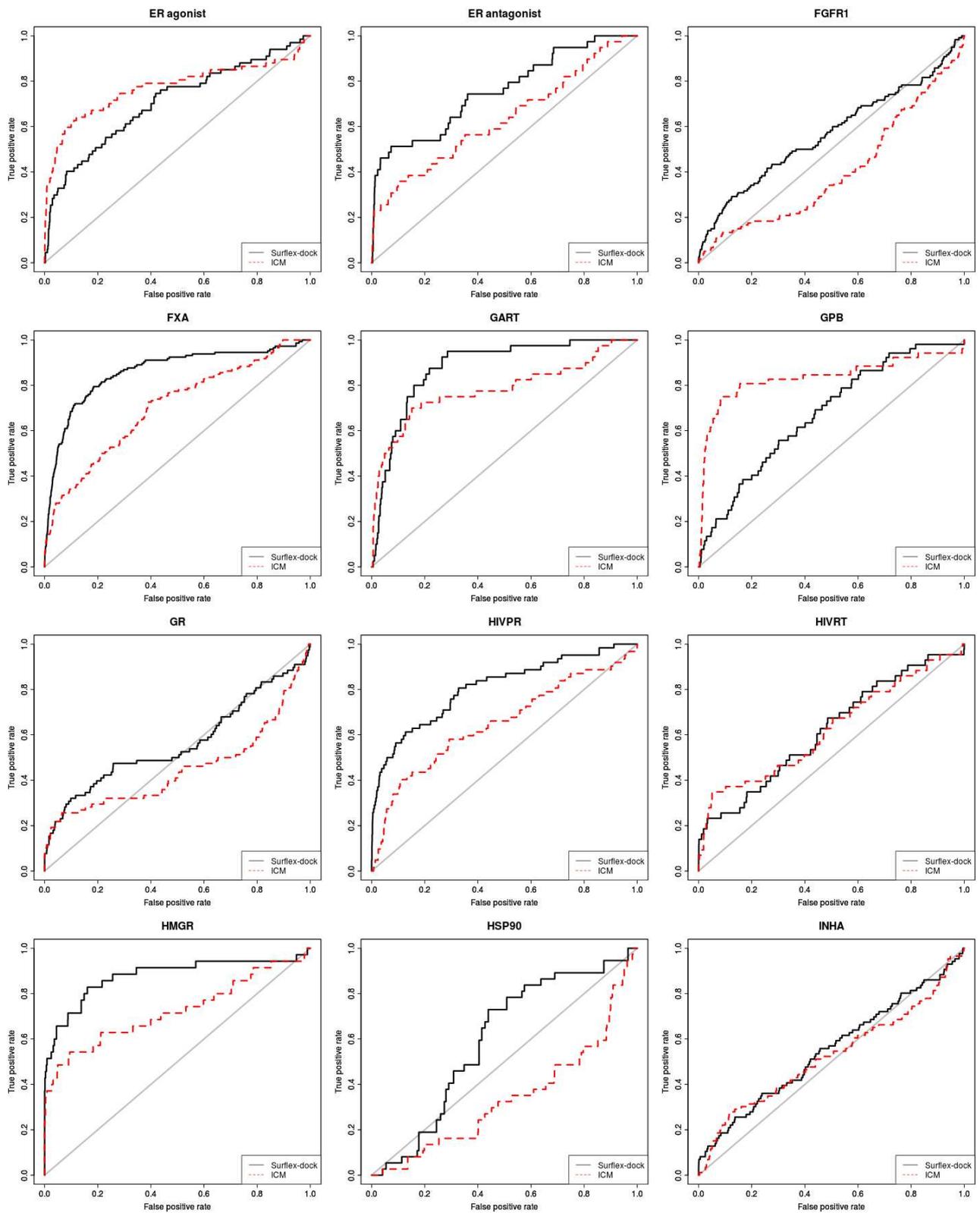


Figure 14 (suite)

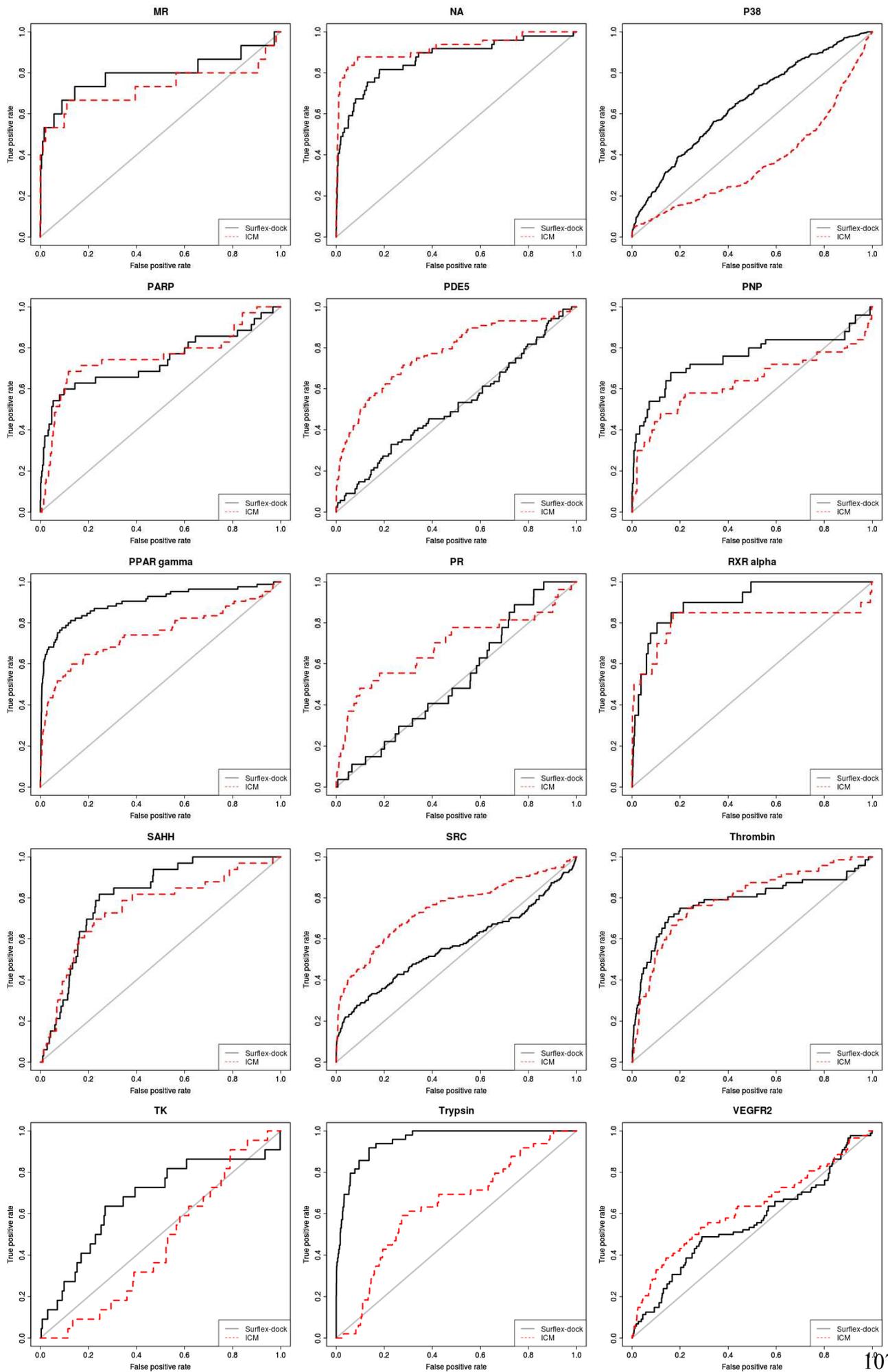


Figure 14 (fin)

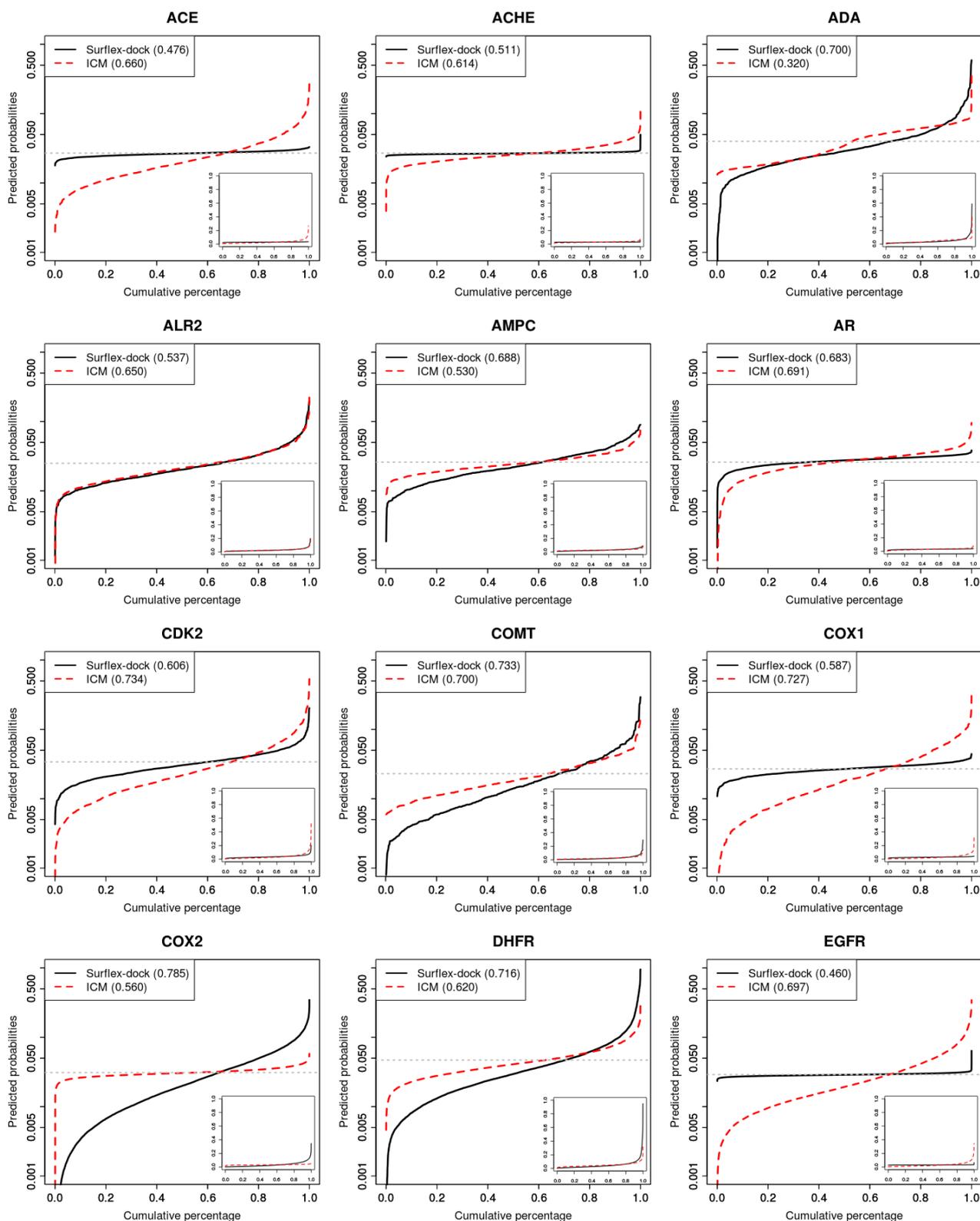


Figure 15 : Courbes de prédictibilité pour Surfex-dock (trait plein noir) et ICM (pointillés rouges) en semi-log, sur les 39 systèmes de la DUD. Les mêmes graphes avec une échelle linéaire sont représentés en encadré. Les pointillés gris au niveau de la prévalence représentent une distribution homogène des actifs dans la banque.

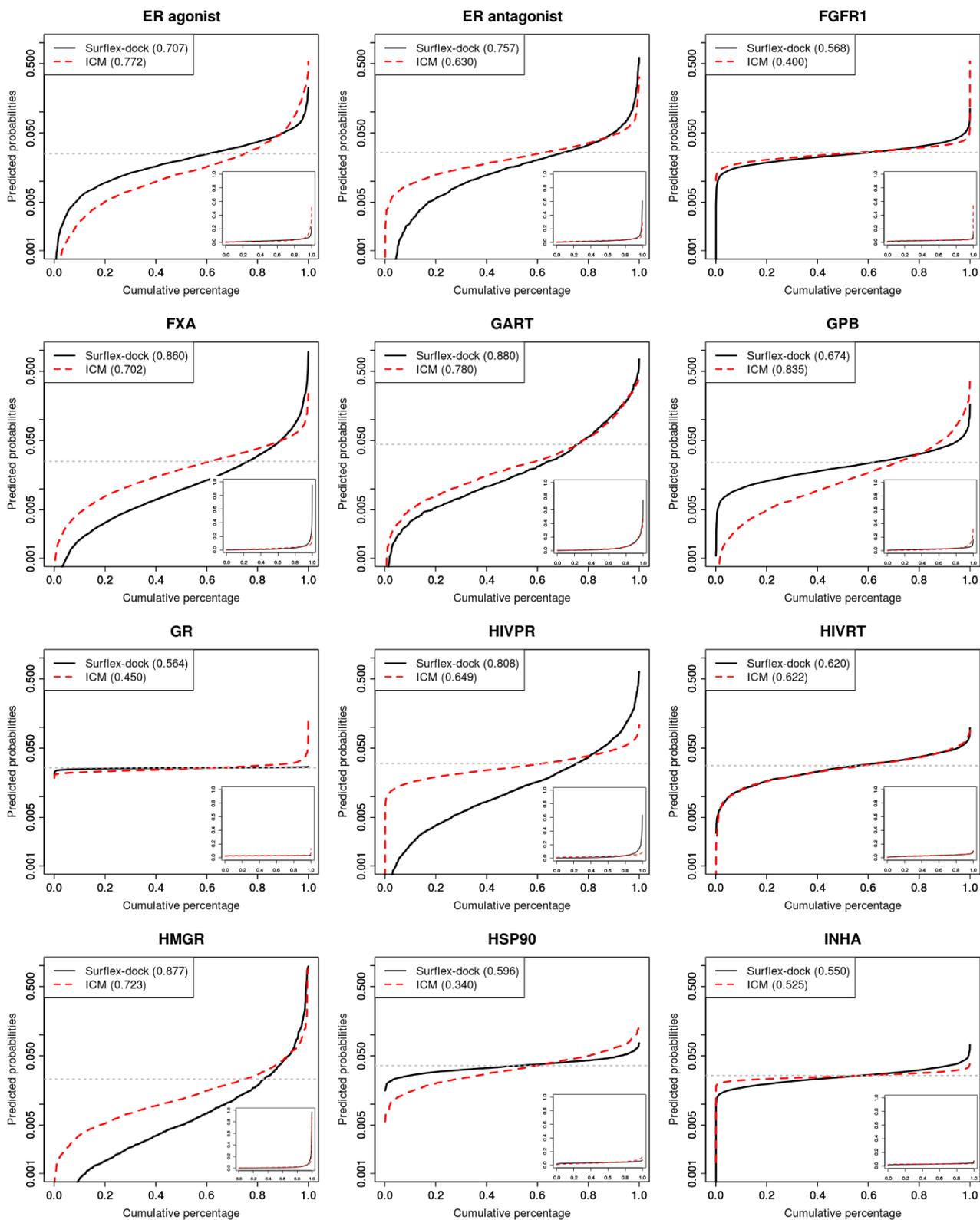


Figure 15 (suite)

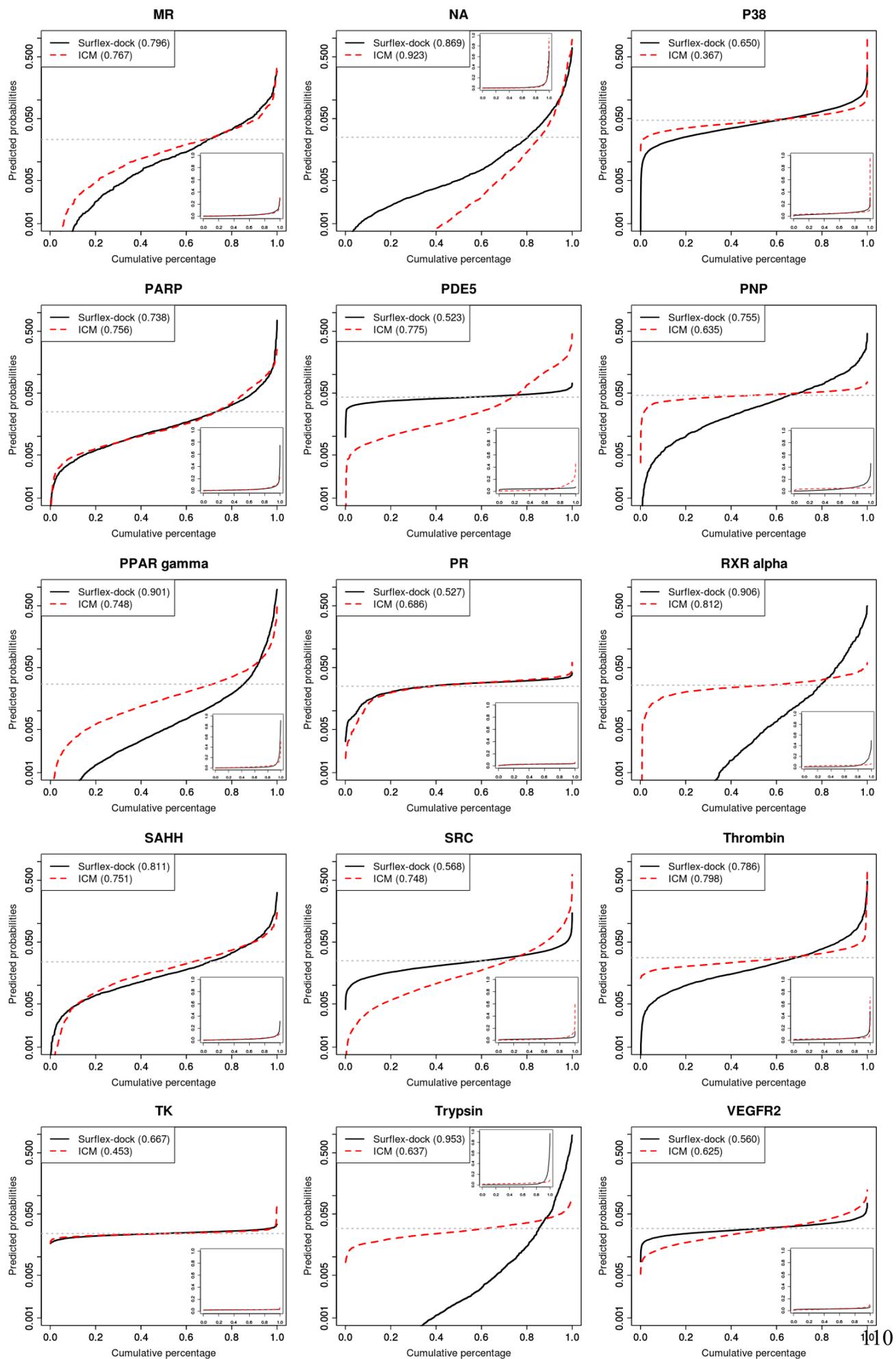


Figure 15 (fin)

II.2 Application : Petites molécules inhibitrices du TNF α oralement actives *in vivo*

Le TNF α a été mis en évidence en 1962 par O'Malley et al²⁰⁷ en montrant que le sérum de souris infectées avec *Serratia marcescens* était capable d'induire la mort de cellules de sarcomes. Ce sérum contenait un facteur conduisant à la régression tumorale : le facteur nécrosant les tumeurs, ensuite rebaptisé le facteur de nécrose tumorale en 1975²⁰⁸. Cette équipe a également montré que les macrophages en étaient la source principale. Depuis, il a été montré qu'elle est produite par des types cellulaires variés regroupant les macrophages, les lymphocytes T, mais également les fibroblastes, les synoviocytes ou encore les keratinocytes. C'est une cytokine connue pour ses propriétés apoptotiques sur les cellules tumorales ainsi que pour son rôle dans l'inflammation, la défense de l'hôte contre les infections et la régulation du système immunitaire^{209; 210}. Sa surproduction a été associée avec des maladies inflammatoires chroniques telles que la polyarthrite rhumatoïde, la maladie de Crohn ou le psoriasis²¹⁰. En raison de son rôle clé dans les maladies inflammatoires, elle est ciblée avec succès par les biothérapies, notamment avec des anticorps monoclonaux (infliximab, adalimumab) et un récepteur soluble du TNF α (etanercept).

L'Infliximab est le premier anticorps ciblant le TNF α à avoir été commercialisé en France, d'abord en tant que traitement contre la polyarthrite rhumatoïde, puis contre la maladie de Crohn. C'est un anticorps chimérique, d'origine humaine à 75% et murine à 25%, ce qui entraîne une perte d'efficacité chez l'homme au fil du temps, avec l'apparition d'anticorps anti-infliximab. L'adalimumab, quant à lui, est un anticorps entièrement humanisé et aussi utilisé dans un premier temps pour traiter la polyarthrite rhumatoïde. L'etanercept est une protéine fusion constituée de la portion extracellulaire du récepteur RII du TNF α et du fragment Fc d'une IgG1. Cette protéine est un récepteur soluble, avec une action plus faible que celles des anticorps car elle ne ciblerait que le TNF α soluble

Cependant, ces biothérapies présentent quelques inconvénients qu'une petite molécule n'aurait pas : effets secondaires²¹¹, résistance au traitement, coût prohibitif (jusqu'à 20000 dollars par

patient et par an). Cependant, après plusieurs tentatives²¹²⁻²¹⁷ aucune petite molécule n'est arrivée sur le marché pour l'instant. Alzani et al²¹² ont montré que la suramine avait une action directe inhibitrice sur le TNF α , en entraînant la dissociation du trimère. Mancini et al²¹³ ont ensuite mis au point des dérivés de la suramine. Cinq analogues de produits naturels, structuralement divers, ont également été identifiés par Choi et al²¹⁵ et deux autres par Chan et al²¹⁶. He et al²¹⁴, dont l'étude a servi de base pour notre composé, ont identifié le produit de synthèse SPD304, capable de dissocier le TNF α trimérique en TNF α dimérique, biologiquement inactif. Enfin, Leung et al²¹⁷ ont testé des molécules déjà présentes sur le marché et ont mis en évidence l'activité inhibitrice de deux molécules sur le TNF α : la darifénacine et l'ézétimibe.

Dans cette étude, une nouvelle famille de petites molécules inhibant le TNF α *in vitro* a été identifiée. Elles sont aussi actives chez des modèles animaux, administrées par voies péritonéale et orale. Ces molécules pourraient aboutir à de nouveaux composés pouvant traiter les maladies inflammatoires chroniques.

Un filtrage ADME-tox, à l'aide de FAF-drugs²¹⁸, a été fait sur la ChemBridge, permettant de passer de 900000 à 700000 composés pour le criblage. Un criblage *in silico* a donc été fait avec Surflex-dock, dans un premier temps, suivi d'un criblage *in vitro* sur les composés sélectionnés, ce qui a permis d'identifier des dérivés benzène sulfonamide. Pour le criblage *in silico*, la structure dimérique du TNF α (code PDB : 2AZ5), issue de l'étude de He et al²¹⁴, a été utilisée comme cible, le site de liaison étant déterminé à 4 Å autour du ligand cocrystallisé. Les hydrogènes ont été ajoutés à l'aide de Chimera²¹⁹.

Des analogues des composés donnant les meilleurs résultats *in vitro* ont ensuite été recherchés à l'aide de méthodes de recherche de similarité 2D/3D, disponibles sur²²⁰. Le seuil de similarité a été fixé à 60%.

A chaque étape de criblage *in silico*, les composés ont été sélectionnés pour les essais expérimentaux après une inspection visuelle des molécules ayant obtenues les meilleurs scores.

In Vivo Orally Active Small-Molecule Inhibitors of TNF α **.

Matthieu Montes^{1*}, Hadley Mouhsine^{1,2*}, H  l  ne Guillemain^{1,2}, Bruno Baron³, Lucille Desallais¹, Gabriel Moreau¹, Aur  lien Latouche⁴, Nesrine Ben Nasr¹, Rojo Ratsimandresy¹, Jean-Louis Spadoni¹, Herv   Do¹, Patrick England³, Jean-Fran  ois Zagury^{1#}

¹Chaire de Bioinformatique, Laboratoire G  nomique, Bioinformatique et Applications, EA 4627, Conservatoire National des Arts et M  tiers, 292 rue Saint Martin, 75003 Paris, FRANCE

²Vaxconsulting SARL, 66 avenue des Champs-Elys  es, 75008, Paris, France

³Plate-forme de Biophysique des Macromol  cules et de leurs Interactions , Proteopole Institut Pasteur, 25 rue du Dr Roux, 75015 Paris

⁴Equipe MSDMA, Centre d'  tudes et de recherche en informatique et communications, EA 4629, Conservatoire National des Arts et M  tiers, 292 rue Saint Martin, 75003 Paris, FRANCE

*These authors contributed equally to this work

#To whom correspondence should be addressed : Prof Jean-Fran  ois Zagury, MD, PhD, e-mail : zagury@cnam.fr

** We thank Prof Jain for providing Surflex-dock. HM and HG are recipients of CIFRE fellowships from ANRT. NBN is a recipient of a fellowship from CNAM. LD is supported by a fellowship of the DGA.

ABSTRACT

TNF α is a key cytokine associated with chronic inflammatory diseases, whose direct targeting by protein biotherapies has been an undeniable success of biotech companies. In a quest for equally effective small molecule drugs, we identified by *in silico* screening a new family of compounds that bind TNF α and inhibit its activity *in vitro* and *in vivo*, orally and intraperitoneally. This breakthrough could open avenues for innovative TNF α -targeted therapeutics.

Tumor Necrosis Factor α (TNF α) is a key cytokine of the immune system mainly involved in inflammation, host defense against infections, regulation of the immune system and tumor regression ^[1, 2]. The overproduction of TNF α has been notably associated with chronic inflammatory diseases notably rheumatoid arthritis, Crohn's disease or Psoriasis ^[2]. Protein biotherapies targeting TNF α have been a revolution for the treatment of chronic inflammatory diseases with anti-TNF α monoclonal antibodies (infliximab, adalimumab) or soluble receptors of TNF α (etanercept). However, these biotherapies present several drawbacks including side-effects^[3], treatment resistance, and prohibitive costs (up to \$20,000 per patient per year)^[4], that could be addressed by small molecule-based therapies. However, despite many efforts ^[5-10], no small molecule drug targeting TNF has been released so far. We report here a new family of small molecules that inhibit TNF α *in vitro* and are active intraperitoneally and *per os* in TNF α -linked animal models. These molecules could represent lead compounds to treat chronic inflammatory diseases.

We have used high-throughput *in silico* screening approaches to identify small molecules targeting a binding pocket of the TNF α identified by He et al ^[9] in 2005 (Supplementary Figure 1). We carried out a hierarchical *in silico* and *in vitro* screening of a collection of 700,000 commercially available compounds using the TNF α structure as a template (Supplementary Figure 2): this led to the identification of benzene sulfonamide derivatives represented by compound 1 (Figure 1).

First, we demonstrated that compound 1 could enhance in a dose-dependent manner the intrinsic tryptophan fluorescence (ITF) of TNF α , confirming that it could bind directly to the TNF α trimer (Figure). Interestingly, the previously described compound SPD304, active only *in vitro*, induces an opposite ITF quenching ^[9], suggesting that the TNF α binding mode of both compounds are different. Second, compound 1 inhibited the induction of apoptosis by TNF α in the L929 cell line with an IC₅₀ of 12 μ M without exhibiting any cellular toxicity at 100 μ M (Figure 3). The targeted pocket is identical in both human and murine cytokines and accordingly, we found that compound 1 inhibited both human and murine TNF α with similar IC₅₀s. This aspect is important to explain why our compound could be directly tested in mouse models.

The *in vivo* activity of compound 1 was assessed using two different TNF α -linked murine disease models: the LipopolySaccharide-DGalactosamine (LPS-DGal) septic shock assay ^[11] and the Dextran Sulfate Sodium (DSS)-induced colitis assay ^[12]. As shown in figure 4, in the

LPS/D-Gal septic shock mouse model, compound 1 exhibited a fully protective effect with an intraperitoneal injection of 1 mg per mouse ($p < 7 \cdot 10^{-4}$). In comparison, etanercept, one of the reference anti-TNF protein biotherapies, showed a similar protection at 600 μ g per mouse. In addition, we observed a fully protective effect with a *per os* administration of 5 mg of compound 1 per mouse 8 hours before the induction of the shock ($p < 4 \cdot 10^{-3}$, Figure 5). These results demonstrate that compound 1 has a highly protective activity *in vivo* and suggest that it has an acceptable level of oral bioavailability. In the DSS-induced colitis assay (Figure 6), compound 1 also exhibited a protective effect, as repeated intraperitoneal injections of compound 1 resulted in the maintenance of mice colon lengths compared to the controls ($p < 5 \cdot 10^{-3}$).

Using an *in silico/in vitro* screening approach, we have thus identified a novel compound that inhibits TNF α directly and is active both *in vitro* and *in vivo*. Other small molecules targeting TNF α have been described previously^[9, 10] but, to our knowledge, this is the very first protein-protein interaction inhibitor that provides significant protection in animal models. In the septic shock assay, the effect of compound 1 was comparable with that of the reference anti-TNF α soluble receptor, etanercept. In the colitis model, our compound was efficient (no shortening of the colon) but less significantly than in the septic shock model. This can be explained by the two following reasons 1. the compound was administered only once every three days in the colitis test, and 2. unlike in the septic shock model, the TNF α pathway is not mandatory in the colitis model.

The compound described in this report constitutes a considerable step forward for the development of alternative therapeutics to the biologics currently used to treat chronic inflammatory diseases, with its demonstrated oral availability that would make it a first-in-class. As shown by the undeniable success of anti-TNF α protein biotherapies, TNF α is an attractive drug target for the treatment of chronic inflammatory diseases. Small molecule drugs targeting directly TNF α would potentially provide several advantages over biologics: first of all, easier compliance and lower production costs, but also the possibility to stop treatment immediately in case of adverse side-effects and an additional class of therapeutics to fight resistance arising in patients. Our work shows it is feasible to develop bioavailable small molecule anti-cytokine inhibitors that could be efficient alternatives to the powerful biologics used since 15 years.

EXPERIMENTAL SECTION

Materials, cell line and mice

Compounds were obtained from Chembridge (San Diego, CA, USA). Dimethyl Sulfoxide (DMSO), Lipopolysaccharide (LPS), Thiazolyl Blue Tetrazolium Bromide (MTT) and Cremophor® EL were obtained from Sigma-Aldrich (Saint Quentin Fallavier, France). Human TNF α was obtained from R&D Systems (Lille, France). Dulbecco's Modified Eagle Medium (DMEM), Fetal Bovine Serum (FBS), Penicillin-Streptomycin, L-Glutamine and Phosphate Buffered Saline were obtained from Pan Biotech (Brumath, France). Absolute ethanol was obtained from Prolabo (Strasbourg, France). Actinomycin D and D-Galactosamine were obtained from Fisher (Illkirch, France). Dextran Sodium Sulfate (36,000 – 50,000) was obtained from MP Biochemicals (Illkirch, France) L929 cell line has been grown in the Laboratory for years. 7 weeks-old female Balb/C mice and 18-22 grs female C57BL/6 mice were obtained from Charles River Laboratories (L'Arbresle, France). Mice used in all experiments were handled according to the guidelines and approved protocols of the Direction Départementale de Protection des Populations, Paris, France.

In silico screening

Structure preparation. The binding site has been defined at 4 Å around the co-crystallized SPD304 ligand in the structure of the TNF α dimer (PDB id: 2AZ5, Supplementary Figure 1). Hydrogen atoms were added using Chimera ^[13].

Compound collection. The 900.000 compounds Chembridge screening compound collection was retrieved from www.hit2lead.com. After an ADME-tox filtering using FAF-drugs2 ^[14], 700.000 compounds were selected to constitute our commercially available drug-like compound collection.

Structure-based virtual screening. Molecular docking was performed using Surflex-dock version 2.5 ^[15]. Surflex-dock is based on a modified Hammerhead fragmentation / reconstruction algorithm to dock compounds flexibly into the binding site. The query molecule is decomposed into rigid fragments that are superimposed to the Surflex-protomol *i.e.*, molecular fragments covering the entire binding site. The docking poses are evaluated by an empirical scoring function.

Ligand-based virtual screening. Ligand-based virtual screening was performed using the 2D/3D similarity search methods implemented in the webservice provided by www.hit2lead.com. Analogues identifications were performed using 2D and 3D similarity

search with a 60% 2D or 3D similarity cut-off. Compounds were selected for experimental tests after a careful visual inspection of the retrieved analogues.

Measurement of TNF- α intrinsic fluorescence.

All samples were brought to 10 mM phosphate buffer, 0.05% Tween20, 1% DMSO. Fluorescence readings were made with a Quantamaster QM4CW luminescence spectrometer, (Photon Technology International) by exciting TNF α at 290 nm and measuring the emission peaks at 322 nm. Compound 1 inner filter effects were corrected using absorbance measurements of compound alone solutions at 290 and 322 nm.

Neutralization of cellular TNF α induced apoptosis.

The biological activity of TNF α was assessed by its apoptotic activity on L929 cells. The neutralization assay consisted in mixing TNF α with the potential inhibitor, incubating the mixture for 2 hours, and then testing its apoptotic activity.

80 % confluent L929 cells were plated in flat bottom plates at 4×10^5 cells per well in 100 μ l of DMEM medium containing 10 % FBS, 2 mM L-Glutamine, 100 U/ml Penicillin - 100 μ g/ml Streptomycin and incubated for one night at 37°C, 5 % CO₂. TNF α (150 pg/ml), Actinomycin D (4 μ g/ml) and the compounds at different concentrations (ranging from 100 μ M to 0,1 μ M) were mixed in 150 μ l complete DMEM medium in U-bottom plates. After two hours incubation at 37°C, 5 % CO₂, 100 μ l of the mix was added to the plated cells and incubated at 37°C, 5 % CO₂ for 24 hours. A negative control was also used by incubating TNF α with no compound. Each point was done in duplicate (the test was highly reproducible).

Supernatants were discarded and 100 μ l of MTT at 0,5 mg/ml were added to wells. After two hours, supernatants were discarded and 200 μ l of DMSO were then added. Plates were read at 570 nm with a spectrophotometer providing the optical density (OD) of each well. The percentage of neutralization (%Neutra) by a compound was calculated as follows :

$$\%Neutra = \frac{OD_{compound} - OD_{TNF\alpha}}{OD_{cells} - OD_{TNF\alpha}} \times 100$$

An IC₅₀ could be computed from the percentage of neutralization for each compound.

LPS/D-Galactosamine induced lethal septic shock

7 weeks-old Balb/C mice were injected intraperitoneally with 100 µl of a solution containing A. 25 % Cremophor® EL, 25 % absolute ethanol, 50 % PBS and various doses of compound 1 (1 mg to 0.25 mg) or B. 25 % Cremophor® EL, 25 % absolute ethanol, 50 % PBS alone (compound 1 vehicle) or C. 0.6mg of etanercept or D. PBS alone (etanercept vehicle). Eight hours after, the mice received an intraperitoneal injection of 0,1 µg of LPS and 20 mg of D-Galactosamine in 200 µl of PBS. For each condition, groups of 8 mice were used. Mice survival was monitored during 48 hours after the injections.

A similar test was done per os, by force-feeding 7 weeks-old mice with 100 µl of a solution of DMSO containing various doses of compound 1 (5 mg down to 0.5 mg) 8 hours before receiving an intraperitoneal injection of 200 µl of PBS containing 0,1 µg of LPS and 20 mg of D-Galactosamine. A control group (vehicle) was force-fed with 100 µl of a sonicated solution of DMSO alone 8 hours before to be injected with 200 µl of a LPS/D-Galactosamine solution. Mice survival was monitored during 48 hours after the injection of LPS/D-Galactosamine.

Colitis induced by Dextran Sulfate Sodium (DSS)

18-22 g female C57BL/6 mice received 3% DSS in water for 5 days followed by 7 days of water. At Day 1, Day 4, Day 7 and Day 11, mice received an intraperitoneal injection of 100 µl of A. 1 mg of compound 1 in 25 % Cremophor® EL, 25 % absolute ethanol, 50 % PBS or B. 25 % Cremophor, 25 % absolute ethanol, 50 % PBS alone (compound 1 vehicle). A control group (water) did not drink any DSS and only received PBS i.p.

At Day 13, mice were sacrificed, and their colons were harvested and measured. For each condition, groups of 8 mice were used.

Statistics

For the LPS-DGal septic shock assay, one-tailed p-values were calculated using Fisher's exact test. For the DSS induced colitis assay, the equality of variances was investigated by a F-test and was rejected for compound 1 vs PBS at 0.02 level. Thus, to compare the difference between the means of each sample accounting for unequal variance, we used a two-sided Welch t-test at a 0.05 significance level. For Compound 1 vs PBS the t-test is statistically significant (observed value $t=3.07$, $df=23.4$) with $p < 0.005$ (95% Confidence Interval [0.166; 0.848]).

REFERENCES

- [1] K. Pfeffer, *Cytokine Growth Factor Rev* **2003**, *14*, 185.
- [2] B. B. Aggarwal, *Nat Rev Immunol* **2003**, *3*, 745.
- [3] R. M. Fleischmann, I. Iqbal, R. L. Stern, *Expert Opin Drug Saf* **2004**, *3*, 391.
- [4] P. P. Sfikakis, G. C. Tsokos, *Clin Immunol* **2011**, *141*, 231.
- [5] H. Choi, Y. Lee, H. Park, D. S. Oh, *Bioorg Med Chem Lett* **2010**, *20*, 6195.
- [6] C. H. Leung, D. S. Chan, M. H. Kwan, Z. Cheng, C. Y. Wong, G. Y. Zhu, W. F. Fong, D. L. Ma, *ChemMedChem* **2011**, *6*, 765.
- [7] F. Mancini, C. M. Toro, M. Mabilia, M. Giannangeli, M. Pinza, C. Milanese, *Biochem Pharmacol* **1999**, *58*, 851.
- [8] R. Alzani, A. Corti, L. Grazioli, E. Cozzi, P. Ghezzi, F. Marcucci, *J Biol Chem* **1993**, *268*, 12526.
- [9] M. M. He, A. S. Smith, J. D. Oslob, W. M. Flanagan, A. C. Braisted, A. Whitty, M. T. Cancilla, J. Wang, A. A. Lugovskoy, J. C. Yoburn, A. D. Fung, G. Farrington, J. K. Eldredge, E. S. Day, L. A. Cruz, T. G. Cachero, S. K. Miller, J. E. Friedman, I. C. Choong, B. C. Cunningham, *Science* **2005**, *310*, 1022.
- [10] D. S. Chan, H. M. Lee, F. Yang, C. M. Che, C. C. Wong, R. Abagyan, C. H. Leung, D. L. Ma, *Angew Chem Int Ed Engl* **2010**, *49*, 2860.
- [11] M. A. Freudenberg, C. Galanos, *Infect Immun* **1991**, *59*, 2110.
- [12] S. Melgar, L. Karlsson, E. Rehnstrom, A. Karlsson, H. Utkovic, L. Jansson, E. Michaelsson, *Int Immunopharmacol* **2008**, *8*, 836.
- [13] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J Comput Chem* **2004**, *25*, 1605.
- [14] D. Lagorce, O. Sperandio, H. Galons, M. A. Miteva, B. O. Villoutreix, *BMC Bioinformatics* **2008**, *9*, 396.
- [15] A. N. Jain, *J Med Chem* **2003**, *46*, 499.

FIGURES

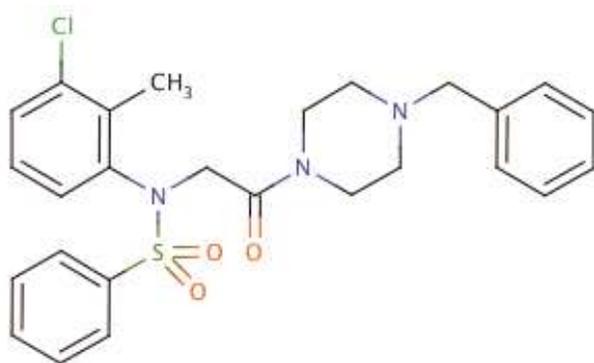


Figure 1: Structure of compound 1.

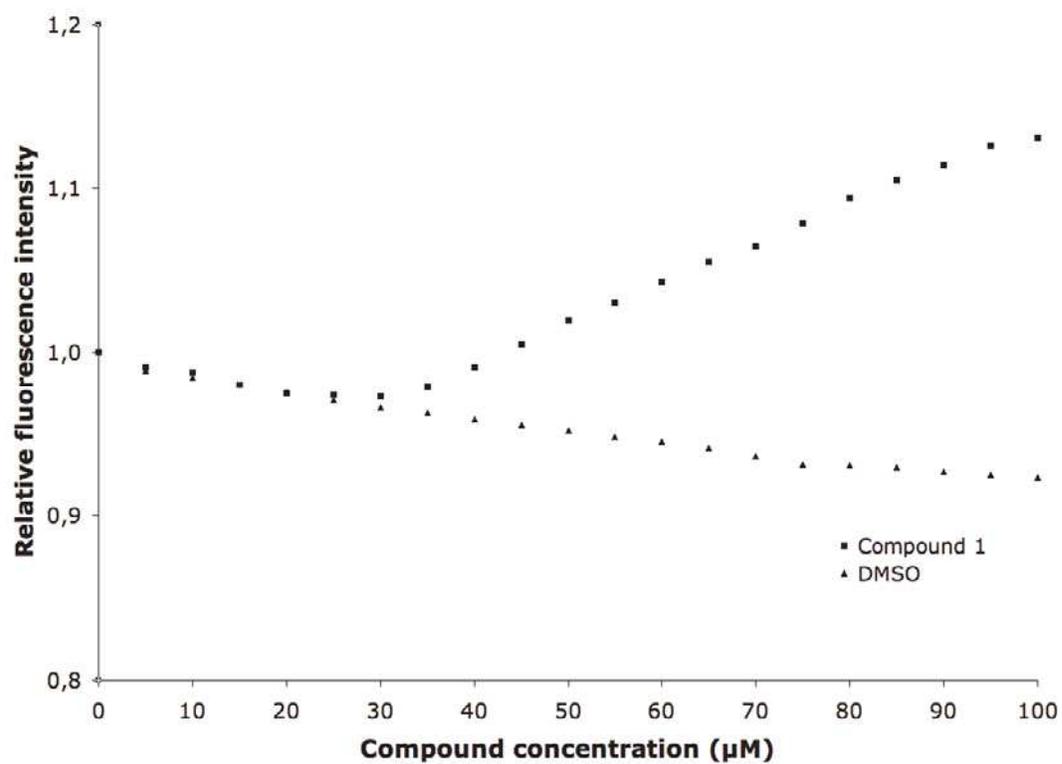


Figure 2: Intrinsic Tryptophan Fluorescence of 0,5 µM TNF α diluted in Phosphate Buffered Saline in the presence of DMSO alone or compound 1 (5 - 100µM) in DMSO.

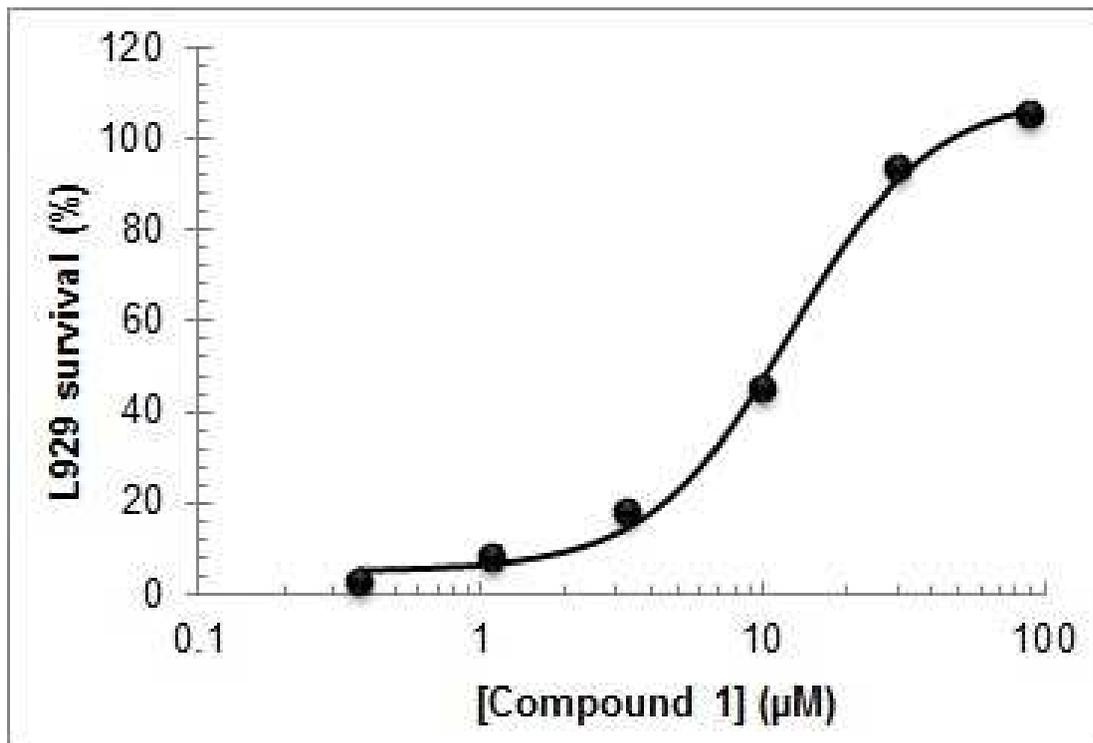


Figure 3: Compound 1 inhibition of TNF α induced apoptosis in L929 cell line. Data represent cell survival in presence of different concentrations of compound 1. IC₅₀=12µM.

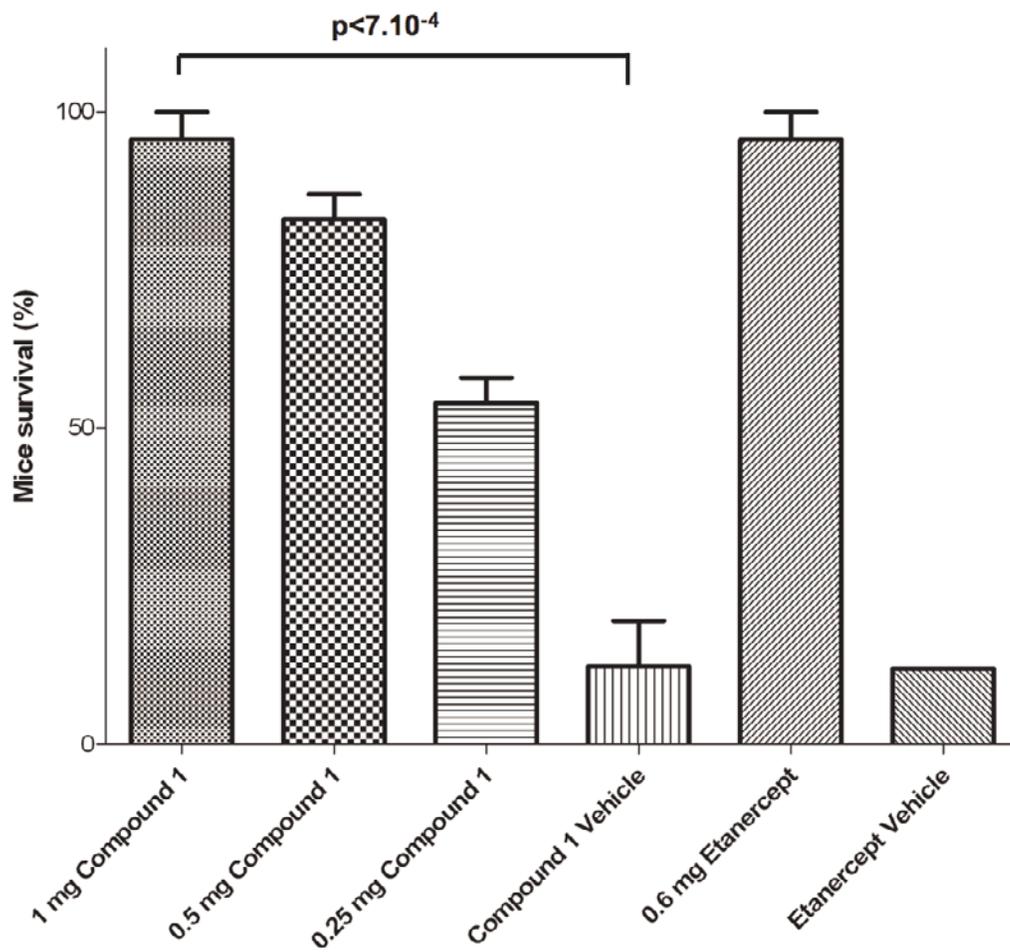


Figure 4: Effect of compound 1 in an *in vivo* murine model of LPS/D-Galactosamine induced septic shock. Mice survival after an intraperitoneal injection of LPS/D-Galactosamine and different doses of compound 1. Groups of eight mice were used. Values are mean \pm s.e.m.

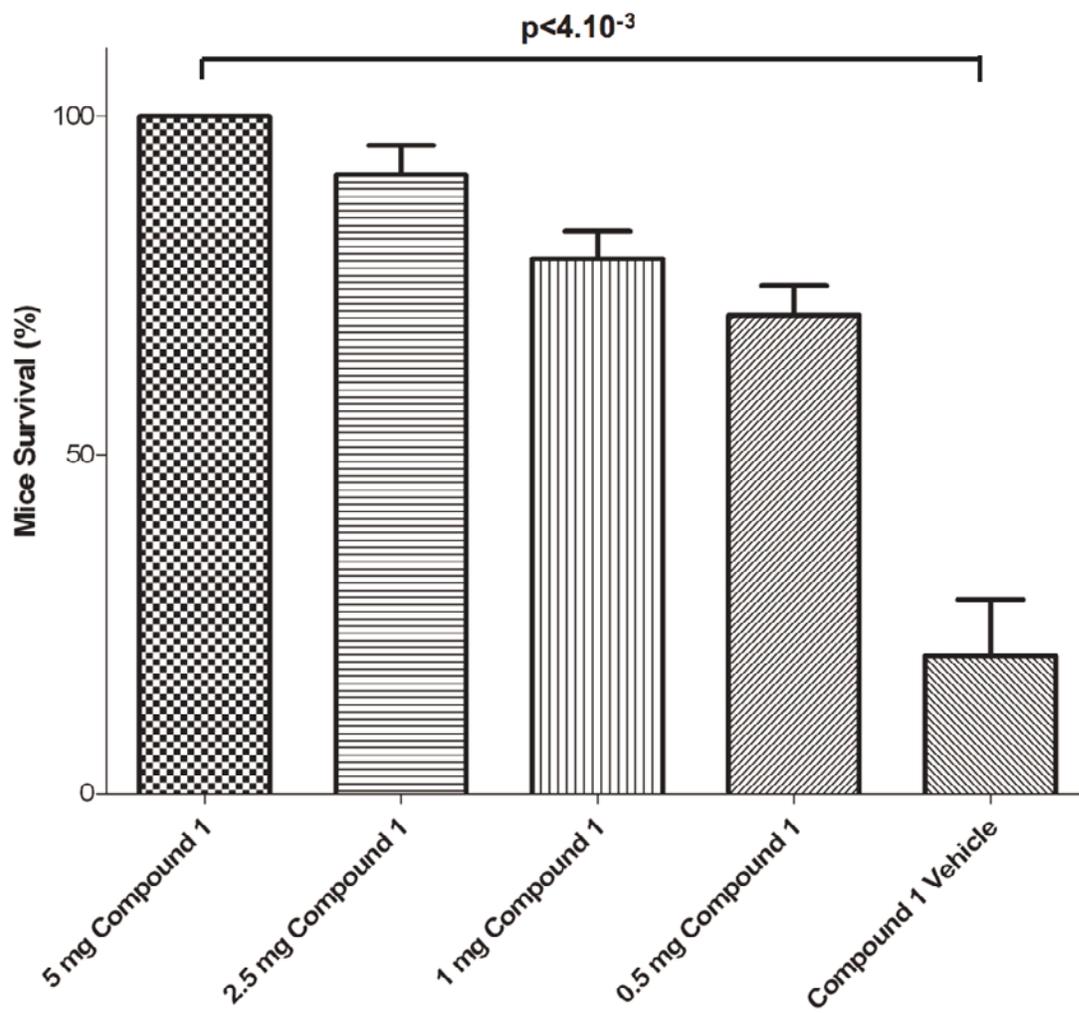


Figure 5: Effect of compound 1 in an *in vivo* murine model of LPS/D-Galactosamine induced septic shock. Mice survival after force-feeding with different doses of compound 1 and an intraperitoneal injection of LPS/D-Galactosamine. Groups of eight mice were used. Values are mean \pm s.e.m.

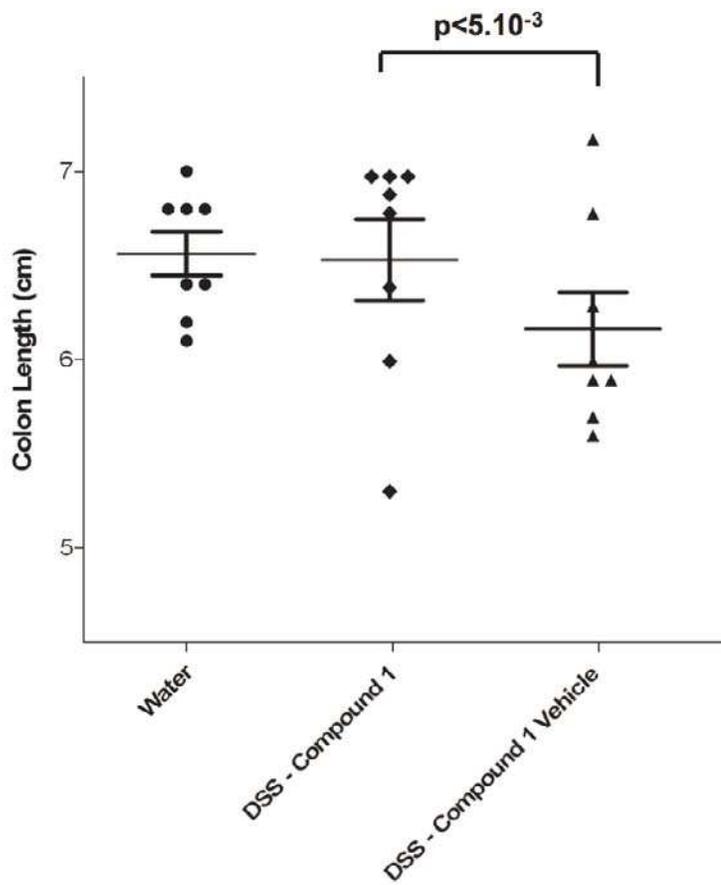
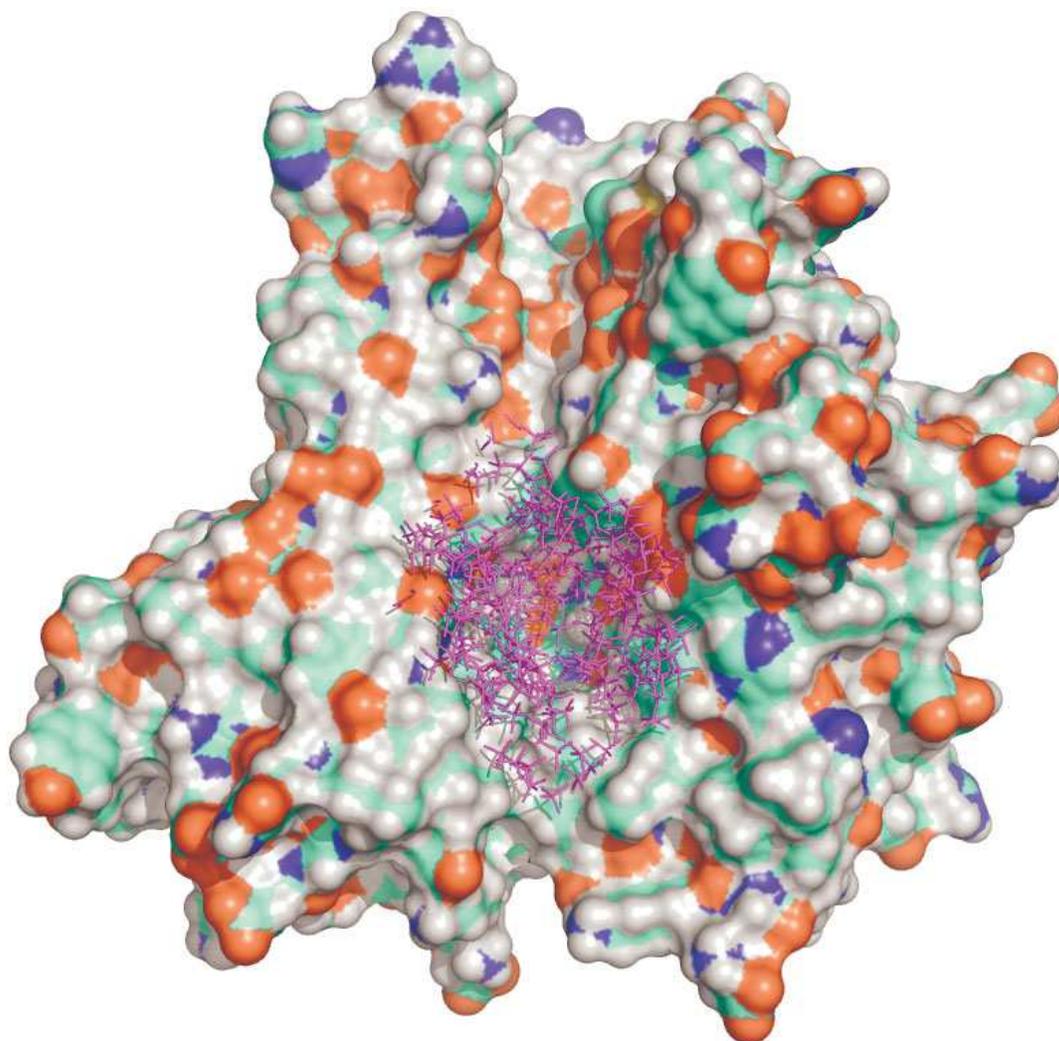
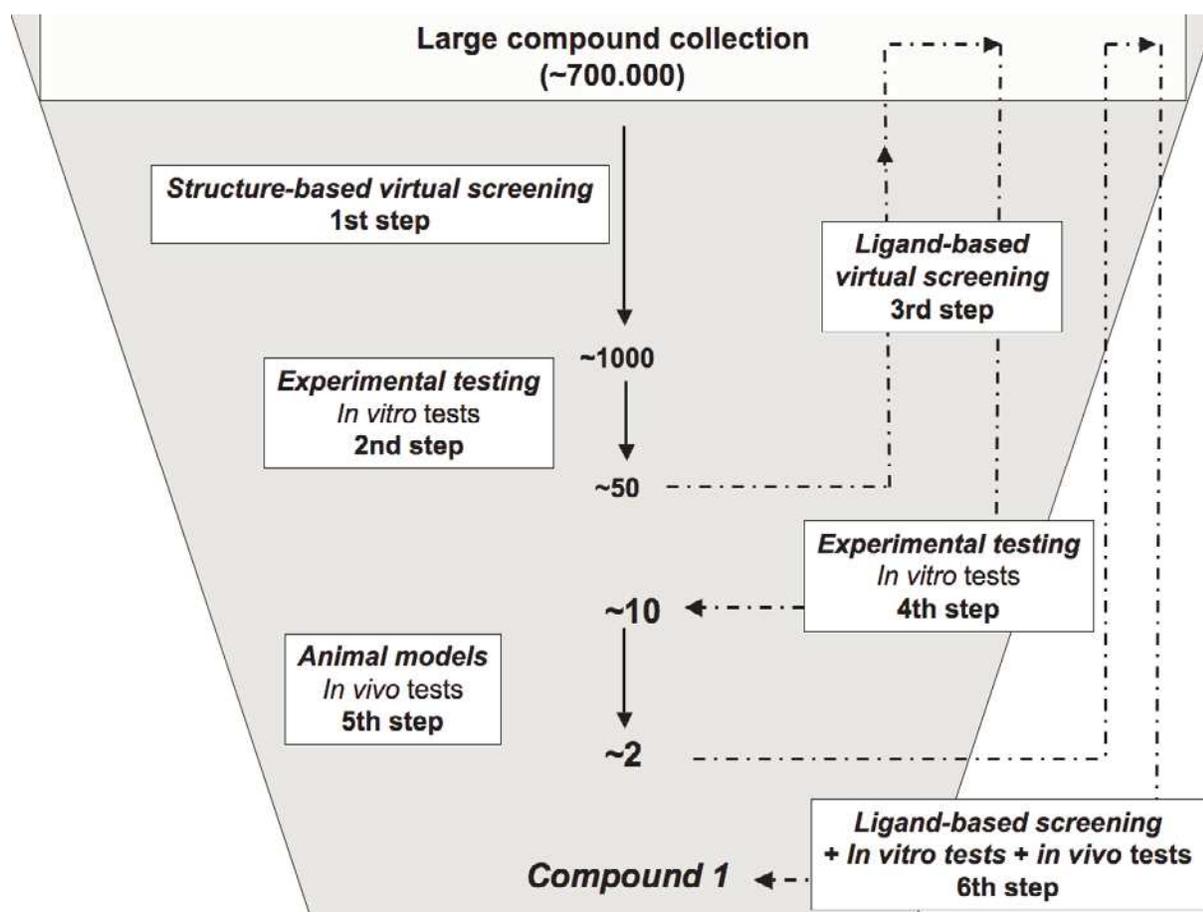


Figure 6: Effect of compound 1 in an *in vivo* murine model of DSS induced acute colitis. Colon length after 5 days of 3% DSS followed by 7 days of water. Compound 1 was injected intraperitoneally every three days as control groups. Groups of eight mice were used. Values are mean \pm s.e.m.

SUPPLEMENTARY ONLINE MATERIAL



Supplementary Figure 1 : Binding pocket defined for the structure-based virtual screening procedure. The binding pocket was defined as the region 4 Å around the co-crystallized ligand SPD304 in the TNF α structure identified by He et al (PDBid =2AZ5). The « protomol » that constitutes the conformational space, that was explored with surflex-dock is displayed in magenta.



Supplementary Figure 2 : compound 1 identification protocol. *1st step* : A collection of 700,000 drug-like commercially available compounds was screened *in silico*. Molecular docking was performed using Surflex-dock version 2.5. A visual inspection of the top scoring compounds was performed in order to select a 1000 compound hit list for experimental testing. *2nd step*: The compounds composing the hit list selected were purchased from the chemical supplier. Their inhibitory activity was evaluated *in vitro* on human TNF α . These *in vitro* experiments led to the identification of top hits: active compounds with an IC₅₀ between 1 and 100 μ M. *3rd step*: Using ligand-based virtual screening methods, we searched in our compound collection for analogues of the top hits identified after step 2. Up to 100 analogues were found per top hit. *4th step*: The analogues were purchased from the chemical supplier. As in step 2, their inhibitory activity was evaluated *in vitro* on human TNF α . Their activity was also confirmed on murine TNF α in view of a further evaluation on mice models. The 10 best compounds after these 4 steps were selected as drug candidates for *in vivo* evaluation on mice models. *5th step*: The *in vivo* evaluation of the drug candidates was performed in the TNF α -dependent septic shock model triggered with LPS-galactosamine using two different administration modes: intraperitoneal injection and force feeding. After this step, 2 compounds sharing the same structural scaffold were selected. *6th Step*: Using ligand-based screening methods, we searched in our large compound collection for new analogues of

the two best compounds identified after step 5. Up to 1000 analogues were identified and purchased from the chemical supplier. As in step 2, their inhibitory activity was evaluated *in vitro* on human and murine TNF α . The 9 best compounds were evaluated *in vivo* in our murine septic shock assay by force feeding as described in step 5. The best compound identified after the 6th step is compound 1.

DISCUSSION

Le composé 1, trouvé après le premier après la recherche d'analogues, peut se lier directement au TNF α sous forme trimérique, bien que son mode de liaison semble être différent de celui de SPD304. Le profil de fluorescence intrinsèque du TNF α n'est, en effet, pas modifié de manière équivalente par les deux composés. Il a été montré par les essais *in vitro* qu'il parvient à inhiber l'induction de l'apoptose sans toxicité cellulaire pour une IC₅₀ de 12 μ M avec le TNF α humain. La même chose est observée avec le TNF α murin, pour une IC₅₀ similaire.

In vivo, une protection complète au choc septique a été observée chez la souris après injection d'1 mg par voie intra-péritonéale (contre 600 μ g pour l'etanercept). Cette protection a aussi été observée avec l'administration de 5 mg par voie orale. Le composé 1 a donc une activité hautement protectrice et il semble posséder un bon potentiel de biodisponibilité orale. Un effet protecteur existe aussi sur les modèles de colite.

Grâce à la recherche d'analogues, un nouveau composé a été trouvé. Bien que d'autres petites molécules ciblant le TNF α aient été trouvées jusqu'à maintenant, cela semble être le premier inhibiteur du TNF actif par voie orale, avec un effet comparable à un anti-TNF de référence (etanercept) pour le choc septique. Il est aussi efficace dans les modèles de colite, bien que cela soit moins significatif que pour le choc septique.

Conclusion

En conclusion, l'évaluation de méthodes de criblage *in silico* tient une place importante dans la mise en place du processus de conception d'un médicament. Il est en effet important de connaître les points forts et faibles des méthodes de criblage utilisées afin d'obtenir des résultats optimaux lors de l'application sur une cible d'intérêt. Il est apparu que toutes les méthodes étudiées pouvaient faire des alignements ou des positionnements corrects. De plus, les méthodes basées sur la structure de la protéine cible permettent d'obtenir un meilleur enrichissement global. Plus surprenant, les deux types de méthodes, qu'elles soient basées sur la structure de la cible ou du ligand de référence, permettent de retrouver des actifs ayant des chemotypes divers de manière précoce. Cependant, malgré ces bons résultats, certains problèmes persistent. Ainsi, le paradoxe des actifs mal positionnés dans le site de liaison mais obtenant des scores parmi les meilleurs a été illustré, mettant en évidence la limite des fonctions de score à retrouver la bonne pose. Le problème de la prise en compte du site de liaison comme entité rigide par une grande partie des méthodes de docking est aussi apparu, notamment en raison du mode de construction de la DUD. Les conformations des sites de liaison de certains récepteurs n'étant pas adaptés à tous les actifs, il est donc nécessaire de prendre en compte l'adaptation du récepteur pendant la liaison avec son ligand.

Le docking classique sur un ensemble de structures d'un même récepteur est fréquemment utilisé. Du fait du coût en temps de calcul d'un tel protocole et du nombre de structures disponibles pour certaines cibles, tous les ensembles possibles ne peuvent pas être testés. Il était donc nécessaire de trouver des pistes pour choisir les structures à intégrer à l'ensemble sans devoir faire des criblages préalables. Les résultats préliminaires donnent des tendances intéressantes. En effet, parmi les systèmes étudiés, les conformations dont le site de liaison a le volume le plus faible sont généralement associées aux meilleurs enrichissements, et inversement pour les sites de liaison ayant le plus grand volume. En ce qui concerne les ensembles, ceux qui ne contiennent que deux structures (l'une d'entre elles étant presque toujours la meilleure structure seule) semblent donner les meilleurs enrichissements. Ces résultats restent toutefois à approfondir. De plus, une nouvelle version de la DUD, la DUD-E, venant d'être rendue publique, nos conclusions restent à être vérifiées sur celle-ci.

L'utilisation des métriques d'évaluation est un autre aspect problématique de l'évaluation des méthodes de criblage. Le facteur d'enrichissement reste une métrique très utilisée. Cependant, depuis 2005, la courbe de ROC, indépendante du ratio d'actifs dans la banque d'évaluation,

contrairement au facteur d'enrichissement est recommandée. Elle n'est pas toujours facilement interprétable et d'autres alternatives ont depuis été proposées, telles que BEDROC ou RIE. Il semble aussi intéressant d'évaluer la capacité d'une fonction de score à prédire le risque pour une molécule d'être active. Bien que dépendante du ratio d'actifs de la banque, la courbe de prédictibilité a l'avantage de fournir une échelle commune aux différentes fonctions de score et de représenter les écarts de score entre les composés, en plus de leurs rangs. Les résultats obtenus sont encore en cours d'analyse.

Alors qu'aucune petite molécule inhibitrice du TNF α n'a été mise sur le marché jusqu'à présent, l'application de ces méthodes de docking a permis la mise en évidence d'un composé inhibiteur, ayant une activité *in vitro* et *in vivo* par voie orale, dans des modèles de choc septique et de colite chez la souris. Ce composé pourrait constituer les bases de nouveaux traitements des maladies inflammatoires chroniques actifs par voie orale. En accord avec la littérature²²¹⁻²²³, cette application montre également que, malgré le fait que les méthodes nécessitent d'être améliorées, le criblage *in silico* peut permettre d'apporter un soutien important à l'identification de nouvelles molécules à visée thérapeutique.

Bibliographie

1. MORGAN, S ; GROOTENDORST, P ; LEXCHIN, J ET AL. The cost of drug development: a systematic review. *Health Policy*, 2011, 100, 4-17.
2. GORDON, EM ; BARRETT, RW ; DOWER, WJ ET AL. Applications of combinatorial technologies to drug discovery. 2. Combinatorial organic synthesis, library screening strategies, and future directions. *Journal of Medicinal Chemistry*, 1994, 37, 1385-401.
3. YU, H ; ADEDOYIN, A. ADME-Tox in drug discovery: integration of experimental and computational technologies. *Drug Discovery Today*, 2003, 8, 852-61.
4. TOBINICK, EL. The value of drug repositioning in the current pharmaceutical market. *Drug News & Perspectives*, 2009, 22, 119-25.
5. HUGHES, JP ; REES, S ; KALINDJIAN, SB ET AL. Principles of early drug discovery. *British Journal of Pharmacology*, 2011, 162, 1239-49.
6. PEREIRA, DA ; WILLIAMS, JA. Origin and evolution of high throughput screening. *British Journal of Pharmacology*, 2007, 152, 53-61.
7. ZHANG, S. Computer-aided drug discovery and development. *Methods in Molecular Biology (Clifton, N.J.)*, 2011, 716, 23-38.
8. MCGANN, M. FRED pose prediction and virtual screening accuracy. *Journal of Chemical Information and Modeling*, 2011, 51, 578-96.
9. PAN, SY ; PAN, S ; YU, Z ET AL. New perspectives on innovative drug discovery: an overview. *Journal of Pharmacy & Pharmaceutical Sciences : a Publication of the Canadian Society for Pharmaceutical Sciences, Societe Canadienne des Sciences Pharmaceutiques*, 2010, 13, 450-71.
10. LIPINSKI, CA. Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods*, 2000, 44, 235-49.
11. BOHACEK, RS ; McMARTIN, C ; GUIDA, WC. The art and practice of structure-based drug design: a molecular modeling perspective. *Medicinal Research Reviews*, 1996, 16, 3-50.
12. MONGE, A. *Création et utilisation de chimiothèques optimisées pour la recherche "in silico" de nouveaux composés bioactifs*. Thèse de Chimie Informatique et Théorique, 2006, 185 p.
13. BRADLEY, MP. An overview of the diversity represented in commercially-available databases. *Journal of Computer-Aided Molecular Design*, 2002, 16, 301-9.
14. BOSTRÖM, J ; GREENWOOD, JR ; GOTTFRIES, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *Journal of Molecular Graphics & Modelling*, 2003, 21, 449-62.
15. CORINA. [référence du 08 août 2012], www.molecular-networks.com/node/1
16. Concord. [référence du 15 juillet 2012], www.tripos.com
17. ICM. [référence du 15 juillet 2012], www.molsoft.com
18. Chemical Abstract Service. [référence du 16 juillet 2012], www.cas.org
19. CAS Registry. [référence du 16 juillet 2012], <http://www.cas.org/expertise/cascontent/ataglance/index.html>
20. CAplus. [référence du 16 juillet 2012], <http://www.cas.org/expertise/cascontent/ataglance/index.html>
21. SPRESIweb. [référence du 16 juillet 2012], www.spresi.com
22. ACD Chemicals. [référence du 16 juillet 2012], <http://accelrys.com/products/databases/sourcing/available-chemicals-directory.html>
23. Sigma-Aldrich. [référence du 16 juillet 2012], <http://www.sigmaaldrich.com>
24. Specs. [référence du 16 juillet 2012], <http://www.specs.net>
25. MDL Drug Data Report. [référence du 17 juillet 2012], <http://accelrys.com/products/databases/bioactivity/mddr.html>
26. JÓNSDÓTTIR, SO ; JØRGENSEN, FS ; BRUNAK, S. Prediction methods and databases within chemoinformatics: emphasis on drugs and drug candidates. *Bioinformatics (Oxford, England)*, 2005, 21, 2145-60.
27. Derwent World Drug Index. [référence du 17 juillet 2012], <http://www.daylight.com/products/wdi.html>
28. National Cancer Institute. [référence du 17 juillet 2012], <http://cactus.nci.nih.gov/download/nci/>

29. ChEMBLdb. [référence du 17 juillet 2012], <https://www.ebi.ac.uk/chembl/db/>
30. NEWMAN, DJ ; CRAGG, GM. Natural products as sources of new drugs over the last 25 years. *Journal of Natural Products*, 2007, 70, 461-77.
31. BUTLER, MS. Natural products to drugs: natural product-derived compounds in clinical trials. *Natural Product Reports*, 2008, 25, 475-516.
32. Tripos LeadScreen. [référence du 17 juillet 2012], <http://www.tripos.com>
33. Chemdiv CombiLab. [référence du 17 juillet 2012], <http://www.combichemlab.com/website/files/Combichem/Services/chemdiv.htm>
34. Chembridge diverset. [référence du 17 juillet 2012], www.chembridge.com/screening_libraries/diversity_libraries/
35. ChemBridge. [référence du 17 juillet 2012], <http://www.chembridge.com>
36. JENCKS, WP. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America*, 1981, 78, 4046-50.
37. FARMER, P ; ARÉNS, E. Speculations on design of nonpeptidic peptidomimetics. *Trends in Pharmacological Sciences*, 1982, 2, 362-365.
38. HANN, MM ; LEACH, AR ; HARPER, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *Journal of Chemical Information and Computer Sciences*, 2001, 41, 856-64.
39. REES, DC ; CONGREVE, M ; MURRAY, CW ET AL. Fragment-based lead discovery. *Nature Reviews. Drug Discovery*, 2004, 3, 660-72.
40. ChemBridge fragments. [référence du 17 juillet 2012], (http://www.chembridge.com/screening_libraries/fragment_library/index.php)
41. chemdiv fragments. [référence du 17 juillet 2012], <http://eu.chemdiv.com/index.php/discovery-tools-chemistry/screening-libraries/drug-like/special-edition>
42. WEININGER, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *Journal of Chemical Information and Computer Sciences*, 1988, 28, 31-36.
43. WEININGER, D ; WEININGER, A ; WEININGER, JL. SMILES, 2. Algorithm for Generation of Unique SMILES Notation. *Journal of Chemical Information and Computer Sciences*, 1989, 29, 97-101.
44. SMILES. [référence du 17 juillet 2012], <http://www.daylight.com/dayhtml/doc/theory/index.pdf>
45. PDB. [référence du 17 juillet 2012], <http://www.pdb.org>
46. BERNSTEIN, FC ; KOETZLE, TF ; WILLIAMS, GJ ET AL. The Protein Data Bank: a computer-based archival file for macromolecular structures. *Journal of Molecular Biology*, 1977, 112, 535-42.
47. DALBY, A ; NOURSE, JG ; HOUNSHELL, WD ET AL. Description of Several Chemical Structure File Formats Used by Computer Program Developed at Molecular Design Limited. *Journal of Chemical Information and Computer Sciences*, 1992, 32, 244-255.
48. MOL2. [référence du 17 juillet 2012], <http://www.tripos.com/data/support/mol2.pdf>
49. Ligprep. [référence du 17 juillet 2012], www.schrodinger.com
50. Quacpac. [référence du 17 juillet 2012], <http://www.eyesopen.com/quacpac>
51. LI, AP. Screening for human ADME/Tox drug properties in drug discovery. *Drug Discovery Today*, 2001, 6, 357-366.
52. SCHNEIDER, G. *Prediction of Drug-like Properties.*, Madame Curie Bioscience Database : 1 octobre 2002, [référence du 22 juillet 2012]. <http://www.landesbioscience.com/curie/chapter/447/>
53. LIPINSKI, CA ; LOMBARDO, F ; DOMINY, BW ET AL. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 2001, 46, 3-26.
54. GHOSE, AK ; VISWANADHAN, VN ; WENDOLOSKI, JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of Combinatorial Chemistry*, 1999, 1, 55-68.
55. DROR, O ; SHULMAN-PELEG, A ; NUSSINOV, R ET AL. Predicting molecular interactions in silico: I. A guide to pharmacophore identification and its applications to drug design. *Current Medicinal Chemistry*, 2004, 11, 71-90.

56. RIPPHAUSEN, P ; NISIUS, B ; PELTASON, L ET AL. Quo vadis, virtual screening? A comprehensive survey of prospective applications. *Journal of Medicinal Chemistry*, 2010, 53, 8461-7.
57. ICM pharmacophore. [référence du 05 août 2012], <http://www.molsoft.com/gui/ph4.html>
58. KIER, LB. *Molecular orbital theory in drug design*. New York, American Press, 1971, 258 p.
59. GUND, P. Three-dimensional pharmacophoric pattern searching. *Progress in Molecular and Subcellular Biology*, 1977, 5, 117-143.
60. MARTIN, YC, . In : Martin, YC ; PWillett (Eds.), *Designing Bioactive Molecules: Three-Dimensional Techniques and Applications*. Washington, DC : American Chemical Society, 1998, 28 p.
61. WILLETT, P ; BARNARD, JM ; DOWNS, GM. Chemical Similarity Searching. *Journal of Chemical Information and Computer Sciences*, 1998, 38, 983-996.
62. JOHNSON, MA ; MAGGIORA, GM. *Concepts and applications of molecular similarity*. New York : John Wiley & Sons (Ed.), 1990, 393 p.
63. LEACH, AR ; GILLET, VJ. *An Introduction to Chemoinformatics*. : Kluwer Academic Publishers (Ed.), 2003, 276 p.
64. GASTEIGER, J ; ENGEL, T. *Chemoinformatics: A Textbook* : Gasteiger, J ; TEngel (Eds.), 2003, 680 p.
65. SCIOR, T ; BENDER, A ; TRESADERN, G ET AL. Recognizing Pitfalls in Virtual Screening: A Critical Review. *Journal of Chemical Information and Modeling*, 2012.
66. HAWKINS, PCD ; SKILLMAN, AG ; WARREN, GL ET AL. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *Journal of Chemical Information and Modeling*, 2010, 50, 572-84.
67. JAIN, AN. Ligand-based structural hypotheses for virtual screening. *Journal of Medicinal Chemistry*, 2004, 47, 947-61.
68. RUSH, TS3 ; GRANT, JA ; MOSYAK, L ET AL. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *Journal of Medicinal Chemistry*, 2005, 48, 1489-95.
69. ROCS. [référence du 20 juillet 2012], <http://www.eyesopen.com/rocs>
70. KEISER, MJ ; ROTH, BL ; ARMBRUSTER, BN ET AL. Relating protein pharmacology by ligand chemistry. *Nature Biotechnology*, 2007, 25, 197-206.
71. PETRONE, PM ; SIMMS, B ; NIGSCH, F ET AL. Rethinking Molecular Similarity: Comparing Compounds on the Basis of Biological Activity. *ACS Chemical Biology*, 2012, , .
72. KAUVAR, LM ; HIGGINS, DL ; VILLAR, HO ET AL. Predicting ligand binding to proteins by affinity fingerprinting. *Chemistry & Biology*, 1995, 2, 107-18.
73. FLIRI, AF ; LOGING, WT ; THADEIO, PF ET AL. Biospectra analysis: model proteome characterizations for linking molecular structure and biological response. *Journal of Medicinal Chemistry*, 2005, 48, 6918-25.
74. FLIRI, AF ; LOGING, WT ; THADEIO, PF ET AL. Biological spectra analysis: Linking biological activity profiles to molecular structure. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102, 261-6.
75. ABDO, A ; SALIM, N ; AHMED, A. Implementing relevance feedback in ligand-based virtual screening using Bayesian inference network. *Journal of Biomolecular Screening : the Official Journal of the Society for Biomolecular Screening*, 2011, 16, 1081-8.
76. TURTLE, H ; CROFT, BW. Evaluation of an Inference Network-Based Retrieval Model. *ACM Transactions on Information Systems*, 1991, 9, 187-222.
77. HOLLIDAY, JD ; HU, C ; WILLETT, P. Grouping of coefficients for the calculation of inter-molecular similarity and dissimilarity using 2D fragment bit-strings. *Combinatorial Chemistry & High Throughput Screening*, 2002, 5, 155-66.
78. JACKSON, DA ; SOMERS, KM ; HARVEY, HH. Similarity coefficients: Measures of Co-occurrence and Association or Simply Measures of Occurrence ?. *The American Naturalist*, 1989, 133, 136-453.
79. GILLET, VJ ; WILD, DJ ; WILLETT, P ET AL. Similarity and Dissimilarity Methods for Processing Chemical Structure Databases. *The Computer Journal*, 1998, 41, 548-558.
80. HAMMING, R. Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 1950, 29, 147-160.

81. MYINT, K ; MA, C ; WANG, L ET AL. Fragment-similarity-based QSAR (FS-QSAR) Algorithm for Ligand Biological Activity Predictions. *SAR and QSAR in Environmental Research*, 2011, 22, 385-410.
82. CERETO-MASSAGUÉ, A ; GUASCH, L ; VALLS, C ET AL. DecoyFinder: an easy-to-use python GUI application for building target-specific decoy sets. *Bioinformatics (Oxford, England)*, 2012, 28, 1661-1662.
83. BINIASHVILI, T ; SCHREIBER, E ; KLIGER, Y. Improving classical substructure-based virtual screening to handle extrapolation challenges. *Journal of Chemical Information and Modeling*, 2012, 52, 678-85.
84. BORK, P ; JENSEN, LJ ; VON MERING, C ET AL. Protein interaction networks from yeast to human. *Current Opinion in Structural Biology*, 2004, 14, 292-9.
85. PRYMULA, K ; JADCZYK, T ; ROTERMAN, I. Catalytic residues in hydrolases: analysis of methods designed for ligand-binding site prediction. *Journal of Computer-Aided Molecular Design*, 2011, 25, 117-33.
86. WEISEL, M ; PROSCHAK, E ; SCHNEIDER, G. PocketPicker: analysis of ligand binding-sites with shape descriptors. *Chem Cent J*, 2007, 1, 7.
87. LEVITT, DG ; BANASZAK, LJ. POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *Journal of Molecular Graphics*, 1992, 10, 229-34.
88. HENDLICH, M ; RIPPMMANN, F ; BARNICKEL, G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of Molecular Graphics & Modelling*, 1997, 15, 359-63, 389.
89. BRADY, GPJ ; STOUTEN, PF. Fast prediction and visualization of protein binding pockets with PASS. *Journal of Computer-Aided Molecular Design*, 2000, 14, 383-401.
90. BINKOWSKI, TA ; NAGHIBZADEH, S ; LIANG, J. CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Research*, 2003, 31, 3352-5.
91. LAURIE, ATR ; JACKSON, RM. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. *Bioinformatics (Oxford, England)*, 2005, 21, 1908-16.
92. ZHOU, D ; YE, JJ ; LI, Y ET AL. The molecular basis of the interaction between the proline-rich SH3-binding motif of PNRC and estrogen receptor alpha. *Nucleic Acids Research*, 2006, 34, 5974-86.
93. STAHL, M ; TARONI, C ; SCHNEIDER, G. Mapping of protein surface cavities and prediction of enzyme class by a self-organizing neural network. *Protein Engineering*, 2000, 13, 83-8.
94. PUGALENTHI, G ; KUMAR, KK ; SUGANTHAN, PN ET AL. Identification of catalytic residues from protein structure using support vector machine with sequence and structural features. *Biochemical and Biophysical Research Communications*, 2008, 367, 630-4.
95. ZHANG, K ; XU, Y ; CHEN, G. PECB: prediction of enzyme catalytic residues based on Naive Bayes classification. *Int J Bioinform Res Appl*, 2008, 4, 295-305.
96. DAI, T ; LIU, Q ; GAO, J ET AL. A new protein-ligand binding sites prediction method based on the integration of protein sequence conservation information. *BMC Bioinformatics*, 2011, 12 Suppl 14, S9.
97. WALLACE, AC ; BORKAKOTI, N ; THORNTON, JM. TESS: a geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases. Application to enzyme active sites. *Protein Science : a Publication of the Protein Society*, 1997, 6, 2308-23.
98. CAPRA, JA ; SINGH, M. Predicting functionally important residues from sequence conservation. *Bioinformatics (Oxford, England)*, 2007, 23, 1875-82.
99. Poseview. [référence du 05 août 2012], <http://www.poseview.de/>
100. MCGANN, MR ; ALMOND, HR ; NICHOLLS, A ET AL. Gaussian docking functions. *Biopolymers*, 2003, 68, 76-90.
101. KUNTZ, ID ; BLANEY, JM ; OATLEY, SJ ET AL. A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology*, 1982, 161, 269-88.
102. RENNER, S ; SCHWAB, CH ; GASTEIGER, J ET AL. Impact of conformational flexibility on three-dimensional similarity searching using correlation vectors. *Journal of Chemical Information and Modeling*, 2006, 46, 2324-32.
103. MILLER, MD ; KEARSLEY, SK ; UNDERWOOD, DJ ET AL. FLOG: a system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure. *Journal of Computer-Aided Molecular Design*, 1994, 8, 153-74.

104. HIRAYAMA, K ; AOKI, S ; NISHIKAWA, K ET AL. Identification of novel chemical inhibitors for ubiquitin C-terminal hydrolase-L3 by virtual screening. *Bioorganic & Medicinal Chemistry*, 2007, 15, 6810-8.
105. VILLOUTREIX, BO ; LACONDE, G ; LAGORCE, D ET AL. Tyrosine kinase syk non-enzymatic inhibitors and potential anti-allergic drug-like compounds discovered by virtual and in vitro screening. *PLoS ONE*, 2011, 6, e21117.
106. HAMZA, A ; ZHAO, X ; TONG, M ET AL. Novel human mPGES-1 inhibitors identified through structure-based virtual screening. *Bioorganic & Medicinal Chemistry*, 2011, 19, 6077-86.
107. KITCHEN, DB ; DECORNEZ, H ; FURR, JR ET AL. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews. Drug Discovery*, 2004, 3, 935-49.
108. BROOIJMANS, N ; KUNTZ, ID. Molecular recognition and docking algorithms. *Annual Review of Biophysics and Biomolecular Structure*, 2003, 32, 335-73.
109. LEACH, AR. *Molecular Modelling: Principles and Applications*. Harlow (Essex) : Pearson Education Limited (Ed.), 1996, 745 p.
110. DESJARLAIS, RL ; SHERIDAN, RP ; DIXON, JS ET AL. Docking flexible ligands to macromolecular receptors by molecular shape. *Journal of Medicinal Chemistry*, 1986, 29, 2149-53.
111. RAREY, M ; KRAMER, B ; LENGAUER, T ET AL. A fast flexible docking method using an incremental construction algorithm. *Journal of Molecular Biology*, 1996, 261, 470-89.
112. EWING, TJ ; MAKINO, S ; SKILLMAN, AG ET AL. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. *Journal of Computer-Aided Molecular Design*, 2001, 15, 411-28.
113. JAIN, AN. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. *Journal of Medicinal Chemistry*, 2003, 46, 499-511.
114. GOODSSELL, DS ; OLSON, AJ. Automated docking of substrates to proteins by simulated annealing. *Proteins*, 1990, 8, 195-202.
115. ABAGYAN, R ; TOTROV, M ; KUZNETSOV, D. ICM: an efficient technique for structure predictions and design. *Journal of Computational Chemistry*, 1994, 15, 488-506.
116. JONES, G ; WILLETT, P ; GLEN, RC ET AL. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 1997, 267, 727-48.
117. MORRIS, G ; GOODSSELL, D ; HALLIDAY, R ET AL. Algorithm and Empirical Binding Free Energy Function. *Journal of Computational Chemistry*, 1998, 19, 1639-1662.
118. WU, G ; VIETH, M. SDOCKER: a method utilizing existing X-ray structures to improve docking accuracy. *Journal of Medicinal Chemistry*, 2004, 47, 3142-8.
119. NAKAJIMA, N ; NAKAMURA, H ; KIDERA, A. Multicanonical EnsembleGenerated by Molecular Dynamics Simulation for Enhanced Conformational Sampling of Peptides. *The Journal of Physical Chemistry B*, 1997, 101, 817-824.
120. DI NOLA, A ; ROCCATANO, D ; BERENDSEN, HJ. Molecular dynamics simulation of the docking of substrates to proteins. *Proteins*, 1994, 19, 174-82.
121. MANGONI, M ; ROCCATANO, D ; DI NOLA, A. Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation. *Proteins*, 1999, 35, 153-62.
122. KOSHLAND, DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 1958, 44, 98-104.
123. CARLSON, HA ; MCCAMMON, JA. Accommodating protein flexibility in computational drug design. *Molecular Pharmacology*, 2000, 57, 213-8.
124. HART, TN ; READ, RJ. A multiple-start Monte Carlo docking method. *Proteins*, 1992, 13, 206-22.
125. LEACH, AR. Ligand docking to proteins with discrete side-chain flexibility. *Journal of Molecular Biology*, 1994, 235, 345-56.
126. DESMET, J ; DE MAEYER, M ; HAZES, B ET AL. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature*, 1992, 356, 539-42.
127. KNEGTEL, RM ; KUNTZ, ID ; OSHIRO, CM. Molecular docking to ensembles of protein structures. *Journal of Molecular Biology*, 1997, 266, 424-40.

128. GOODFORD, PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *Journal of Medicinal Chemistry*, 1985, 28, 849-57.
129. FISCHER, HE. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Deutschen Chemischen Gesellschaft*, 1894, 27, 2985-93.
130. PAULING, L ; DELBRÜCK, M. THE NATURE OF THE INTERMOLECULAR FORCES OPERATIVE IN BIOLOGICAL PROCESSES. *Science*, 1940, 92, 77-9.
131. KAUZMANN, W. Some factors in the interpretation of protein denaturation. *Advances in Protein Chemistry*, 1959, 14, 1-63.
132. DAVIS, SJ ; DAVIES, EA ; TUCKNOTT, MG ET AL. The role of charged residues mediating low affinity protein-protein recognition at the cell surface by CD2. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95, 5490-4.
133. BRANDSDAL, BO ; OSTERBERG, F ; ALMLÖF, M ET AL. Free energy calculations and ligand binding. *Advances in Protein Chemistry*, 2003, 66, 123-58.
134. HELMS, V ; WADE, RC. Computational Alchemy to Calculate Absolute Protein-Ligand Binding Free Energy. *Journal of American Chemical Society*, 1998, 120, 2710-2713.
135. BÖHM, HJ. The computer program LUDI: a new method for the de novo design of enzyme inhibitors. *Journal of Computer-Aided Molecular Design*, 1992, 6, 61-78.
136. BÖHM, HJ. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. *Journal of Computer-Aided Molecular Design*, 1992, 6, 593-606.
137. ROGNAN, D ; LAUEMOLLER, SL ; HOLM, A ET AL. Predicting binding affinities of protein ligands from three-dimensional models: application to peptide binding to class I major histocompatibility proteins. *Journal of Medicinal Chemistry*, 1999, 42, 4650-8.
138. JONES, G ; WILLETT, P ; GLEN, RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *Journal of Molecular Biology*, 1995, 245, 43-53.
139. WEINER, SJ ; KOLLMAN, PA ; NGUYEN, DT ET AL. An all atom force field for simulations of proteins and nucleic acids. *Journal of Computational Chemistry*, 1986, 7, 230-252.
140. KRAMER, B ; RAREY, M ; LENGAUER, T. Evaluation of the FlexX Incremental Construction Algorithm for Protein-Ligand Docking. *PROTEINS: Structure, Function, and Genetics*, 1999, 37, 228-241.
141. VERDONK, ML ; COLE, JC ; HARTSHORN, MJ ET AL. Improved protein-ligand docking using GOLD. *Proteins*, 2003, 52, 609-23.
142. MUEGGE, I ; MARTIN, YC. A general and fast scoring function for protein-ligand interactions: a simplified potential approach. *Journal of Medicinal Chemistry*, 1999, 42, 791-804.
143. DEWITTE, RS ; SHAKHNOVICH, EI. SMOG: de Novo Design Method Based on Simple, Fast and Accurate Free Energy Estimates. 1. Methodology and Supporting Evidence. *Journal of American Chemical Society*, 1996, 118, 11733-11744.
144. GOHLKE, H ; HENDLICH, M ; KLEBE, G. Knowledge-based scoring function to predict protein-ligand interactions. *Journal of Molecular Biology*, 2000, 295, 337-56.
145. TAYLOR, RD ; JEWsbury, PJ ; ESSEX, JW. A review of protein-small molecule docking methods. *Journal of Computer-Aided Molecular Design*, 2002, 16, 151-66.
146. CORTÉS-CABRERA, A ; GAGO, F ; MORREALE, A. A reverse combination of structure-based and ligand-based strategies for virtual screening. *Journal of Computer-Aided Molecular Design*, 2012, 26, 319-27.
147. WU, M ; DAI, D ; YAN, H. PRL-dock: Protein-ligand docking based on hydrogen bond matching and probabilistic relaxation labeling. *Proteins*, 2012, 80, 2137-53.
148. LEE, HS ; ZHANG, Y. BSP-SLIM: a blind low-resolution ligand-protein docking approach using predicted protein structures. *Proteins*, 2012, 80, 93-110.
149. IRWIN, JJ. Community benchmarks for virtual screening. *Journal of Computer-Aided Molecular Design*, 2008, 22, 193-9.
150. BOSTRÖM, J. Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. *Journal of Computer-Aided Molecular Design*, 2001, 15, 1137-52.

151. VERDONK, ML ; BERDINI, V ; HARTSHORN, MJ ET AL. Virtual screening using protein-ligand docking: avoiding artificial enrichment. *Journal of Chemical Information and Computer Sciences*, 2004, 44, 793-806.
152. TRESADERN, G ; AGRAFIOTIS, DK. Conformational sampling with stochastic proximity embedding and self-organizing superimposition: establishing reasonable parameters for their practical use. *Journal of Chemical Information and Modeling*, 2009, 49, 2786-800.
153. KROEMER, RT ; VULPETTI, A ; McDONALD, JJ ET AL. Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations. *Journal of Chemical Information and Computer Sciences*, 2004, 44, 871-81.
154. COLE, JC ; MURRAY, CW ; NISSINK, JWM ET AL. Comparing protein-ligand docking programs is difficult. *Proteins*, 2005, 60, 325-32.
155. ABAGYAN, RA ; TOTROV, MM. Contact area difference (CAD): a robust measure to evaluate accuracy of protein models. *Journal of Molecular Biology*, 1997, 268, 678-85.
156. KIRCHMAIR, J ; MARKT, P ; DISTINTO, S ET AL. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection--what can we learn from earlier mistakes?. *Journal of Computer-Aided Molecular Design*, 2008, 22, 213-28.
157. BISSANTZ, C ; FOLKERS, G ; ROGNAN, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *Journal of Medicinal Chemistry*, 2000, 43, 4759-67.
158. PHAM, TA ; JAIN, AN. Parameter estimation for scoring protein-ligand interactions using negative training data. *Journal of Medicinal Chemistry*, 2006, 49, 5856-68.
159. MCGOVERN, SL ; SHOICHET, BK. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *Journal of Medicinal Chemistry*, 2003, 46, 2895-907.
160. HUANG, N ; SHOICHET, BK ; IRWIN, JJ. Benchmarking sets for molecular docking. *Journal of Medicinal Chemistry*, 2006, 49, 6789-801.
161. MYSINGER, MM ; CARCHIA, M ; IRWIN, JJ ET AL. Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry*, 2012, 55, 6582-94.
162. TRIBALLEAU, N ; ACHER, F ; BRABET, I ET AL. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *Journal of Medicinal Chemistry*, 2005, 48, 2534-47.
163. CHEN, H ; LYNE, PD ; GIORDANETTO, F ET AL. On evaluating molecular-docking methods for pose prediction and enrichment factors. *Journal of Chemical Information and Modeling*, 2006, 46, 401-15.
164. TRUCHON, J ; BAYLY, CI. Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. *Journal of Chemical Information and Modeling*, 2007, 47, 488-508.
165. THOMAS IV, B ; JOSEPH-McCARTHY, D ; ALVAREZ, J, Pharmacophore-Based Molecular Design. In : Guner, O (Ed.), *Pharmacophore Perception, Development and Use in Drug Design*. La Jolla, California, USA : International University Line, 2000.
166. ZHOU, H ; SKOLNICK, J. FINDSITE(X): a structure-based, small molecule virtual screening approach with application to all identified human GPCRs. *Molecular Pharmaceutics*, 2012, 9, 1775-84.
167. NOVIKOV, FN ; STROYLOV, VS ; ZEIFMAN, AA ET AL. Lead Finder docking and virtual screening evaluation with Astex and DUD test sets. *Journal of Computer-Aided Molecular Design*, 2012, 26, 725-35.
168. HAMZA, A ; WEI, N ; ZHAN, C. Ligand-Based Virtual Screening Approach Using a New Scoring Function. *Journal of Chemical Information and Modeling*, 2012, 52, 963-74.
169. PARK, SH ; GOO, JM ; Jo, C. Receiver operating characteristic (ROC) curve: practical review for radiologists. *Korean Journal of Radiology : Official Journal of the Korean Radiological Society*, 2004, 5, 11-8.
170. SHERIDAN, RP ; SINGH, SB ; FLUDER, EM ET AL. Protocols for bridging the peptide to nonpeptide gap in topological similarity searches. *Journal of Chemical Information and Computer Sciences*, 2001, 41, 1395-406.
171. KIRCHMAIR, J ; DISTINTO, S ; MARKT, P ET AL. How to optimize shape-based virtual screening: choosing the right query and including chemical information. *Journal of Chemical Information and Modeling*, 2009, 49, 678-92.

172. MOFFAT, K ; GILLET, VJ ; WHITTLE, M ET AL. A comparison of field-based similarity searching methods: CatShape, FBSS, and ROCS. *Journal of Chemical Information and Modeling*, 2008, 48, 719-29.
173. LEMMEN, C ; LENGAUER, T. Computational methods for the structural alignment of molecules. *Journal of Computer-Aided Molecular Design*, 2000, 14, 215-32.
174. RAREY, M ; KRAMER, B ; LENGAUER, T. Time-efficient docking of flexible ligands into active sites of proteins. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, 1995, 3, 300-8.
175. LEMMEN, C ; LENGAUER, T ; KLEBE, G. FLEXS: a method for fast flexible ligand superposition. *Journal of Medicinal Chemistry*, 1998, 41, 4502-20.
176. TOTROV, M. Atomic property fields: generalized 3D pharmacophoric potential for automated ligand superposition, pharmacophore elucidation and 3D QSAR. *Chemical Biology & Drug Design*, 2008, 71, 15-27.
177. ROCS & FRED. [référence du 25 juillet 2012], www.eyesopen.com
178. GOOD, AC ; OPREA, TI. Optimization of CAMD techniques 3. Virtual screening enrichment studies: a help or hindrance in tool selection?. *Journal of Computer-Aided Molecular Design*, 2008, 22, 169-78.
179. VON KORFF, M ; FREYSS, J ; SANDER, T. Comparison of ligand- and structure-based virtual screening on the DUD data set. *Journal of Chemical Information and Modeling*, 2009, 49, 209-31.
180. CLEVES, AE ; JAIN, AN. Effects of inductive bias on computational evaluations of ligand-based modeling and on drug discovery. *Journal of Computer-Aided Molecular Design*, 2008, 22, 147-59.
181. BARRIL, X ; MORLEY, SD. Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. *Journal of Medicinal Chemistry*, 2005, 48, 4432-43.
182. BOLSTAD, ESD ; ANDERSON, AC. In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking. *Proteins*, 2009, 75, 62-74.
183. BROUGHTON, HB. A method for including protein flexibility in protein-ligand docking: improving tools for database mining and virtual screening. *Journal of Molecular Graphics & Modelling*, 2000, 18, 247-57, 302-4.
184. CAVASOTTO, CN ; KOVACS, JA ; ABAGYAN, RA. Representing receptor flexibility in ligand docking through relevant normal modes. *Journal of the American Chemical Society*, 2005, 127, 9632-40.
185. TOTROV, M ; ABAGYAN, R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Current Opinion in Structural Biology*, 2008, 18, 178-84.
186. ARMEN, RS ; CHEN, J ; BROOKS, CL3. An Evaluation of Explicit Receptor Flexibility in Molecular Docking Using Molecular Dynamics and Torsion Angle Molecular Dynamics. *J Chem Theory Comput*, 2009, 5, 2909-2923.
187. FERRARI, AM ; WEI, BQ ; COSTANTINO, L ET AL. Soft docking and multiple receptor conformations in virtual screening. *Journal of Medicinal Chemistry*, 2004, 47, 5076-84.
188. GIGANTI, D ; GUILLEMAIN, H ; SPADONI, J ET AL. Comparative evaluation of 3D virtual ligand screening methods: impact of the molecular alignment on enrichment. *Journal of Chemical Information and Modeling*, 2010, 50, 992-1004.
189. SCHNEIDER, N ; HINDLE, S ; LANGE, G ET AL. Substantial improvements in large-scale redocking and screening using the novel HYDE scoring function. *Journal of Computer-Aided Molecular Design*, 2012, 26, 701-23.
190. YOON, S ; WELSH, WJ. Identification of a minimal subset of receptor conformations for improved multiple conformation docking and two-step scoring. *Journal of Chemical Information and Computer Sciences*, 2004, 44, 88-96.
191. CRAIG, IR ; ESSEX, JW ; SPIEGEL, K. Ensemble docking into multiple crystallographically derived protein structures: an evaluation based on the statistical analysis of enrichments. *Journal of Chemical Information and Modeling*, 2010, 50, 511-24.
192. RUEDA, M ; BOTTEGONI, G ; ABAGYAN, R. Recipes for the selection of experimental protein conformations for virtual screening. *Journal of Chemical Information and Modeling*, 2010, 50, 186-93.

193. THOMAS, MP ; McINNES, C ; FISCHER, PM. Protein structures in virtual screening: a case study with CDK2. *Journal of Medicinal Chemistry*, 2006, 49, 92-104.
194. BOTTEGONI, G ; ROCCHIA, W ; RUEDA, M ET AL. Systematic exploitation of multiple receptor conformations for virtual ligand screening. *PLoS ONE*, 2011, 6, e18845.
195. RAO, S ; SANSCHAGRIN, PC ; GREENWOOD, JR ET AL. Improving database enrichment through ensemble docking. *Journal of Computer-Aided Molecular Design*, 2008, 22, 621-7.
196. VERDONK, ML ; MORTENSON, PN ; HALL, RJ ET AL. Protein-ligand docking against non-native protein conformers. *Journal of Chemical Information and Modeling*, 2008, 48, 2214-25.
197. SING, T ; SANDER, O ; BEERENWINKEL, N ET AL. ROCR: visualizing classifier performance in R. *Bioinformatics (Oxford, England)*, 2005, 21, 3940-1.
198. FRACZKIEWICZ, R ; BRAUN, W. Exact and Efficient Analytical Calculation of the Accessible Surface Areas and their Gradients for Macromolecules. *Journal of Computational Chemistry*, 1998, 19, 319-333.
199. HAWKINS, PCD ; WARREN, GL ; SKILLMAN, AG ET AL. How to do an evaluation: pitfalls and traps. *Journal of Computer-Aided Molecular Design*, 2008, 22, 179-90.
200. JAIN, AN ; NICHOLLS, A. Recommendations for evaluation of computational methods. *Journal of Computer-Aided Molecular Design*, 2008, 22, 133-9.
201. JAIN, AN. Bias, reporting, and sharing: computational evaluations of docking methods. *Journal of Computer-Aided Molecular Design*, 2008, 22, 201-12.
202. McCLISH, DK. Analyzing a Portion of the ROC Curve. *Medical Decision Making*, 1989, 9, 190-195.
203. CLARK, RD ; WEBSTER-CLARK, DJ. Managing bias in ROC curves. *Journal of Computer-Aided Molecular Design*, 2008, 22, 141-6.
204. HAND, DJ. Evaluating diagnostic tests: The area under the ROC curve and the balance of errors. *Statistics in Medicine*, 2010, 29, 1502-10.
205. HUANG, Y ; PEPE, MS. A parametric ROC model-based approach for evaluating the predictiveness of continuous markers in case-control studies. *Biometrics*, 2009, 65, 1133-44.
206. BURA, E ; GASTWIRTH, J. The binary regression quantile plot: assessing the importance of predictor in binary regression visually. *Biometrical Journal*, 2001, 43, 5-21.
207. O'MALLEY, WE ; ACHINSTEIN, B ; SHEAR, MJ. Action of Bacterial Polysaccharide on Tumors. I. Reduced Growth of Sarcoma 37, and Induced Refractoriness, by Serrafia Morcescens Polysaccharide. *Journal of the National Cancer Institute*, 1962, 29, 1161-1168.
208. CARSWELL, EA ; OLD, LJ ; KASSEL, RL ET AL. An endotoxin-induced serum factor that causes necrosis of tumors. *Proceedings of the National Academy of Sciences of the United States of America*, 1975, 72, 3666-70.
209. PFEFFER, K. Biological functions of tumor necrosis factor cytokines and their receptors. *Cytokine & Growth Factor Reviews*, 2003, 14, 185-91.
210. AGGARWAL, BB. Signalling pathways of the TNF superfamily: a double-edged sword. *Nature Reviews. Immunology*, 2003, 3, 745-56.
211. FLEISCHMANN, RM ; IQBAL, I ; STERN, RL. Considerations with the use of biological therapy in the treatment of rheumatoid arthritis. *Expert Opinion on Drug Safety*, 2004, 3, 391-403.
212. ALZANI, R ; CORTI, A ; GRAZIOLI, L ET AL. Suramin induces deoligomerization of human tumor necrosis factor alpha. *The Journal of Biological Chemistry*, 1993, 268, 12526-9.
213. MANCINI, F ; TORO, CM ; MABILIA, M ET AL. Inhibition of tumor necrosis factor-alpha (TNF-alpha)/TNF-alpha receptor binding by structural analogues of suramin. *Biochemical Pharmacology*, 1999, 58, 851-9.
214. HE, MM ; SMITH, AS ; OSLOB, JD ET AL. Small-molecule inhibition of TNF-alpha. *Science*, 2005, 310, 1022-5.
215. CHOI, H ; LEE, Y ; PARK, H ET AL. Discovery of the inhibitors of tumor necrosis factor alpha with structure-based virtual screening. *Bioorganic & Medicinal Chemistry Letters*, 2010, 20, 6195-8.
216. CHAN, DS ; LEE, H ; YANG, F ET AL. Structure-based discovery of natural-product-like TNF- α inhibitors. *Angewandte Chemie (International Ed. in English)*, 2010, 49, 2860-4.

217. LEUNG, C ; CHAN, DS ; KWAN, MH ET AL. Structure-based repurposing of FDA-approved drugs as TNF- α inhibitors. *ChemMedChem*, 2011, 6, 765-8.
218. LAGORCE, D ; SPERANDIO, O ; GALONS, H ET AL. FAF-Drugs2: free ADME/tox filtering tool to assist drug discovery and chemical biology projects. *BMC Bioinformatics*, 2008, 9, 396.
219. PETTERSEN, EF ; GODDARD, TD ; HUANG, CC ET AL. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 2004, 25, 1605-12.
220. Hit2Lead. [référence du 01 août 2012], www.hit2lead.com
221. RASTELLI, G ; PACCHIONI, S ; SIRAWARAPORN, W ET AL. Docking and database screening reveal new classes of Plasmodium falciparum dihydrofolate reductase inhibitors. *Journal of Medicinal Chemistry*, 2003, 46, 2834-45.
222. WYSS, PC ; GERBER, P ; HARTMAN, PG ET AL. Novel dihydrofolate reductase inhibitors. Structure-based versus diversity-based library design and high-throughput synthesis and screening. *Journal of Medicinal Chemistry*, 2003, 46, 2304-12.
223. KUBINYI, H, 16. In : Ekins, S (Ed.), *Computer Applications in Pharmaceutical Research and Development*. : John Wiley & Sons, 2006, 48 p.

Liste des publications

1. Giganti D*, **Guillemain H***, Spadoni JL, Nilges M, Zagury JF, Montes M. Comparison evaluation of 3D Virtual Ligand Screening Methods: Impact of the Molecular Alignment on Enrichment. *Journal of Chemical Information and Modeling*, 2010, *50*, 994-1004. (*) contribution égale des co-auteurs
2. Montes M*, Mouhsine H*, **Guillemain H**, Baron B, Desallais L, Moreau G, Latouche A, Ben Nasr N, Ratsimandresy R, Spadoni JL, Do H, England P, Zagury JF. In Vivo Orally Active Small-Molecule Inhibitors of TNF α . Soumis. (*) contribution égale des co-auteurs
3. Ben Nasr N*, **Guillemain H***, Lagarde N*, Zagury JF, Montes M. Multiple structures for Virtual Ligand Screening: defining binding site properties-based criteria to optimize the selection of the query. En préparation. (*) contribution égale des co-auteurs
4. Les courbes de prédictibilité comme métrique d'évaluation des méthodes de criblage *in silico*. En préparation

Liste des communications orales

1. Hélène Guillemain. Identification of chemical small molecules targetting pro-inflammatory cytokines. *4th Ermenonville International Workshop. 2009*
2. Hélène Guillemain. Optimization of Virtual Screening protocols. *5th Ermenonville International Workshop. 2010*
3. Hélène Guillemain. ROC curves – Predictiveness curve. *6th Ermenonville International Workshop. 2011*

Evaluation et application de méthodes de criblage *in silico*

Résumé

Lors de la conception de médicaments, le criblage *in silico* est de plus en plus utilisé et les méthodes disponibles nécessitent d'être évaluées. L'évaluation de 8 méthodes a mis en évidence l'efficacité des méthodes de criblage *in silico* et des problèmes de construction de la banque d'évaluation de référence (DUD), la conformation choisie pour les sites de liaison n'étant pas toujours adaptée à tous les actifs. La puissance informatique actuelle le permettant, plusieurs structures expérimentales ont été choisies pour tenter de mimer la flexibilité des sites de liaison. Un autre problème a été mis en évidence : les métriques d'évaluation des méthodes souffrent de biais. De nouvelles métriques ont donc été proposées, telles que BEDROC et RIE. Une autre alternative est proposée ici, mesurant la capacité prédictive d'une méthode en actifs. Enfin, une petite molécule active sur le TNF α *in vitro* et *in vivo* sur souris a été identifiée par un protocole de criblage *in silico*. Ainsi, malgré le besoin d'amélioration des méthodes, le criblage *in silico* peut être d'un important soutien à l'identification de nouvelles molécules à visée thérapeutique.

Mots clés : criblage *in silico*, docking, similarité 3D, métriques d'évaluation, courbe de prédictibilité, ensemble docking, TNF α

Résumé en anglais

Since the introduction of virtual screening in the drug discovery process, the number of virtual screening methods has been increasing and available methods have to be evaluated.

In this work, eight virtual screening methods were evaluated in the DUD database, showing adequate efficiency. This also revealed some shortcomings of the DUD database as the binding site conformation used in the DUD was not relevant for all the actives.

As computational power now permits to address this issue, classical docking runs have been performed on several X-ray structures, used to represent the binding site flexibility. This also revealed that evaluation metrics show some biases. New evaluation metrics have thus been proposed, e.g. BEDROC and RIE. An alternative method was also proposed using predictiveness curves, based on compound activity probability.

Finally, a virtual screening procedure has been applied to TNF α . A small molecule inhibitor, showing *in vitro* and *in vivo* activity in mice, has been identified. This demonstrated the value of virtual screening for the drug discovery process, although virtual screening methods need to be improved.

Keywords: virtual screening, docking, 3D similarity, evaluation metrics, predictiveness curve, ensemble docking, TNF α