

ANR-05-CIGC-008 : Projet DOCK

Compte-rendu des réunions du 21 et 25 Avril 2008

OBJET : Etat des lieux de l'avancement de l'algorithme

Lieu : Strasbourg

Version 1 du 29/04/2008

Rapporteur : Loraine Brillet

Participants : Dragos Horvath (UGSF), Loraine Brillet (CEA)

L'ordre du jour était le suivant :

- ⇒ Etat des lieux de l'avancement de l'algorithme et prévisions des étapes suivantes (D. Horvath)
- ⇒ Présentation du jeu de données en cours de préparation pour la comparaison de l'algorithme avec d'autres algorithmes de docking classiques (tels que GOLD) et sa validation sur le site ATP de la protéine CK2 (L. Brillet)

1. Etats des lieux de l'avancement de l'algorithme Docking@Grid (D. Horvath)

Les charges atomiques calculées par Vega ZZ (programme gratuit utilisé pour générer les fichiers d'entrée .car pour l'algorithme, contenant les paramètres topologiques nécessaires aux calculs avec le champ de force) n'ayant pas donné de résultats concluants, plusieurs jeux de charges issues du programme de modélisation MOE (Chemical Computing Group) ont été testés. Le meilleur jeu de charges est pour l'instant celui du champ de force MMFF94, même si l'on ne retrouve pas encore la forme native pour quelques structures. La calibration des paramètres du champ de force se poursuit.

Avant de pouvoir commencer à docker des ligands dans des structures protéiques, il reste à intégrer les degrés de liberté intermoléculaires dans l'algorithme, c'est-à-dire les degrés de rotation et de translation du ligand par rapport à la protéine.

2. Présentation du jeu de données CK2 (L. Brillet)

Un jeu de données de composés actifs et inactifs sur le site ATP de la protéine CK2 alpha est en cours de préparation au CEA. Ce jeu de données comprendra des actifs co-cristallisés dans le site ATP issus de la Protein Data Bank (15 actifs, 4 familles chimiques), des actifs / inactifs issus de la littérature (5-6 familles chimiques décrites) et des nouveaux actifs / inactifs découverts en interne au CEA (~20 actifs et 165 inactifs, 5 familles chimiques).

Ce jeu de données servira à la comparaison de l'algorithme Docking@Grid avec d'autres algorithmes de docking classiques tels que GOLD.

La littérature et les structures cristallographiques existantes attestent de l'importance de la flexibilité de certaines parties de la poche ATP de la CK2 pour la liaison au ligand, ce qui en fait une cible intéressante pour l'algorithme Docking@Grid.

3. Proposition de benchmarks pour la validation

En plus des benchmarks du CCDC dont il a été question lors de la réunion de novembre 2007 à Paris (STIC 2007), donnés sur le site http://www.ccdc.cam.ac.uk/products/life_sciences/validate/downloads/download.php4?,

d'autres jeux de données supplémentaires de benchmarking sont disponibles sur le web :

- Des ensembles de « benchmarking » pour le docking ont été mis à disposition par l'UCSF sur le site <http://dud.docking.org/>. L'« abstract » de la publication associée (*Benchmarking sets for molecular docking, Huang & al., J. Med. Chem., 2006, 49, 6789-6801*) est donnée ci-dessous :

Ligand enrichment among top-ranking hits is a key metric of molecular docking. To avoid bias, decoys should resemble ligands physically, so that enrichment is not simply a separation of gross features, yet be chemically distinct from them, so that they are unlikely to be binders. We have assembled a directory of useful decoys (DUD), with 2950 ligands for 40 different targets. Every ligand has 36 decoy molecules that are physically similar but topologically distinct, leading to a database of 98 266 compounds. For most targets, enrichment was at least half a log better with uncorrected databases such as the MDDR than with DUD, evidence of bias in the former. These calculations also allowed 40 x 40 cross-docking, where the enrichments of each ligand set could be compared for all 40 targets, enabling a specificity metric for the docking screens. DUD is freely available online as a benchmarking set for docking at <http://blaster.docking.org/dud/>.

- La base scPDB du laboratoire de Didier Rognan (Strasbourg), disponible sur le site <http://bioinfo-pharma.u-strasbg.fr/scPDB>, propose plus de 6400 complexes de la PDB préparés pour le docking (état de protonation et rotamères ajustés, résolution <2.5Å, ligands « druglike »).

- Sur le site web du labo de Didier Rognan (<http://cheminfo.u-strasbg.fr/template/d/pages/download/download.php>), on peut également télécharger les 100 complexes utilisés pour la comparaison de 8 algorithmes de docking décrite dans la publication « *Comparative evaluation of eight docking tools for docking and virtual screening accuracy, Kellenberger, Proteins, 2004, 57, 225-242* ».