

PRELIMINARY STUDIES ON COMPARISON AND EXPERIMENTAL VALIDATION OF TWO VIRTUAL DOCKING PROCEDURES

Roberto Forlani[§], Giulio Dondio[§], Elena Fioravanzo[‡], Massimo Mabilia[‡], Sujen Moiola[§], Marco D. Parenti[‡] and Morena Spreafico[‡]

[§]NiKem Research S.r.l. Via Zambelletti 25, 20021, Baranzate (Milan), Italy,

[‡]S-IN Soluzioni Informatiche Via Salvemini 9, 36100, Vicenza, Italy,

[§]Department of Environmental Health Sciences, M.Negri Institute, Via Eritrea 62, 20157 Milan, Italy.

for more information: roberto.forlani@nikemresearch.com

Introduction

Structure-based drug design was mainly employed in its infancy as a graphics aid to analyze the structural features of a target protein [1]. Today, the pressure for finding new drugs and the large availability of structural information on 3D protein-ligand complexes have drawn much attention on powerful structural based approaches. Indeed, the number of scientific works in which computational tools are used to assist drug design activities has been constantly increasing. In the structure-based design "tool box", some molecular docking algorithms are rapidly emerging as quick, accurate and cost-saving filters for virtual screening of medium to large compounds collections. Comprehensive studies highlighting strengths and limitations of available docking algorithms are still lacking, thus inducing different research groups to develop their own validation procedures [2,3]. Useful and valid predictions of "correct" ligand-receptor bound associations are highly dependent on preliminary procedures, characterized by different methods, settings and parameters, ligand conformational searching algorithms and poses ranking criteria.

Aim of the work

In this communication the analysis of the "bound" and "unbound" issues [4] for two widely used virtual docking softwares: LigandFit [5,6] and Glide [7], is reported.

For this purpose ten complexes (Table 1) were selected (RCSB protein data bank [8]) matching the following criteria:

- relevance of therapeutic class;
- availability of crystal structure;
- structural diversity of ligands;
- diversity of ligand-binding site interactions.

Table 1

PDB ID	Res (Å)	Protein name	Classification	Binding site type	Primary citation
1ASG	2.06	Thrombin	Serine protease	Exposed	J Med Chem 1999, 42, 1376-1383
1BHF	1.80	p38 α (p38 α mitogen-activated protein kinase)	Tyrosine kinase	Exposed	J Biol Chem 1998, 273, 20238-20242
1BX6	2.10	cAMP-dependent protein kinase	Serine/threonine kinase	Buried	Biochemistry, 1999, 38, 2367-2376
1CKP	2.05	Cyclin dependent protein kinase 2	Serine/threonine kinase	Buried	Science, 1998, 281, 533-538
1DHF	2.30	Dihydrofolate reductase	NADPH2 dependent reductase	Exposed	Biochemistry, 1990, 29, 9467-9479
1FKN	1.90	Memapsin 2 (β -secretase)	Aspartic protease	Exposed	Science, 2000, 290, 150-153
1HVR	1.80	HIV protease	Aspartic protease	Buried	Science, 1994, 263, 380-384
1KE5	2.20	Cyclin dependent 2 protein kinase	Serine/threonine kinase	Buried	J Med Chem 2001, 44, 4339-4358
1QCF	2.00	Hck (hematopoietic cell kinase)	Tyrosine kinase	Buried	Mol Cell 1999, 3, 639-648
2QWI	2.00	Neuraminidase	Sialidase	Exposed	Structure, 1998, 6, 735-745

Softwares performances were evaluated according to their ability in:

- reproducing the crystal binding mode by visual inspection and RMSD evaluation of top scored ligand orientations (*bound docking*);
- identifying active molecules [9] within a database of randomly selected drug-like compounds [10] (*unbound docking*) screened over 1KE5.

Results

LigandFit and Glide bound docking results (task i) are showed in Table 2 and Table 3, respectively; while a summary of active molecules identification results (task ii) for both studied algorithms has been reported in Table 4.

Table 2

PDB ID	RMSD _{max} (Å)	SD	RMSD _{min} (Å)	RMSD _{max} (Å)	Atoms	Rot bonds	poses @ max score
1ASG	4.53	1.18	0.66	3.54	83	16	5
1BHF	8.91	3.01	5.47	11.01	81	20	3
1BX6	1.56	0.45	1.08	2.24	65	12	5
1CKP	1.23	0.75	0.25	1.23	54	8	1
1DHF	1.69	0.75	0.25	2.35	49	9	8
1FKN	0.57	-	0.57	1.25	125	36	1
1HVR	1.30	0.35	0.48	1.61	84	10	12
1KE5	0.75	0.01	0.74	0.76	38	4	4
1QCF	0.33	0.17	0.20	0.56	40	2	4
2QWI	1.57	0.28	1.24	2.06	47	7	6

Table 3

PDB ID	RMSD (Å)	GScore	Hbonds	In place GScore	In place Hbonds	Atoms	Rot bonds
1ASG	0.43	-13.63	-5.5	-15.04	-6.1	83	16
1BHF	1.49	-6.44	-5.9	-8.87	-6.9	81	20
1BX6	0.90	-12.71	-6.7	-10.57	-6.1	65	12
1CKP	1.44	-10.49	-1.5	-11.00	-1.72	54	8
1DHF	3.07	-7.55	-3.5	-10.16	-5.6	49	9
1FKN	4.16	-8.41	-3.9	-12.04	-6.8	117	33
1HVR	1.27	-12.75	-3.4	-14.37	-2.9	84	10
1KE5	0.60	-11.29	-4.4	-10.13	-1.77	38	4
1QCF	0.66	-8.88	-1.1	-9.64	-1.7	40	2
2QWI	1.03	-8.27	-5.7	-6.98	-5.9	47	7

Results emerging from this preliminary study seem to be quite interesting in particular for the reproduction of experimentally observed binding modes (task i) with only two unsatisfying results (RMSD > 2 Å [11], highlighted in Table 2 and 3) for each algorithm.

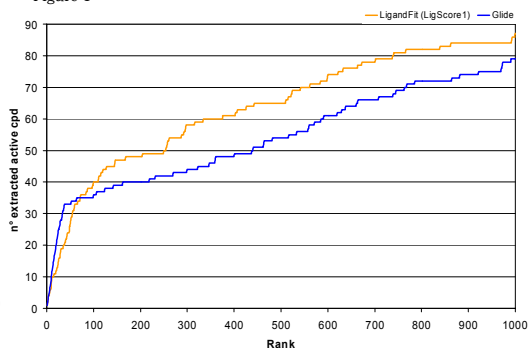
Table 4

	Number of extracted compounds with known activity		
	@ 100*	@ 200*	@ 500*
LigandFit (LigScore1)	39	48	65
Glide (Gscore)	35	40	54

* total number of extracted molecules

Table 4 is a quick overview of the extraction process (task ii), for an overall analysis, the entire profile has been provided (Figure 1).

Figure 1



Conclusions and next steps

At the end of this preliminary study where default working protocols have been taken into consideration for a limited number of complexes, it is possible to conclude that no major differences can be noticed between the two softwares.

Further activities taking into consideration modification to default preliminary procedures and extending the study to other complexes in order to provide further insights on LigandFit and Glide performances are still ongoing.

References

- Luo Z, Wang R, Lai L, RASSE: A new method for structure-based drug design, J. Chem. Inf. Comput. Sci. 1996, 36: 1187-1194.
- Kontoyianni M, McClellan LM, Sokol GS, Evaluation of docking performance: comparative data on docking algorithms, J. Med. Chem. 2004, 47: 558-565.
- Wang R, Lu Y, Wang S, Comparative evaluation of 11 scoring functions for molecular docking, J. Med. Chem. 2003, 46: 2287-2303.
- Halperin I, Boyung M, Wolfson H, Nussinov R, Principles of docking: an overview of search algorithms and a guide to scoring functions, Proteins: Struct., Funct., Genet. 2002, 47: 409-443.
- Cerius² ver. 4.8.1, Accelrys Inc., <http://www.accelrys.com>.
- Venkatachalam CM, Jiang X, Oldfield T, Waldman M LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites, J. Mol. Graphics Modell. 2003, 21:289-307.
- Glide ver. 5.1.020, Schrödinger, L.L.C., <http://www.schrodinger.com>.
- RCSB Protein Data Bank, <http://www.rcsb.org/pdb/>.
- A set of 87 CDK2 known inhibitors (12 nM < IC₅₀ < 25 μM) was selected from Tomimaga Y, Jorgensen WL, J. Med. Chem. 2004, 47: 2534-2549.
- Authors gratefully acknowledge Dr D.Rognan from the University of Strasbourg (CNRS Illkirch, France) for providing a drug-like prefiltered set of 1000 compounds.
- Gohlke H, Hendlich M, Klebe G, Knowledge-based scoring function to predict protein-ligand interaction, J. Mol. Biol. 2000, 295: 337-356.