International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 2– No.7, May 2012 – www.ijais.org

Comparative Data on Docking Algorithms: Keeping the Update in the Field Knowledge

Hioual K.S, Chikhi A, Bensegueni A,Merzoug A, Boucherit H
Laboratory of Applied Biology and Health,
Department of Biochemistry and Microbiology,
Faculty of natural and life Sciences,
Mentouri University, Constantine,
Algeria

ABSTRACT
Two docking programs FlexX and GOLD that can be used for either single-ligand docking or database screening have been compared for their propensity to recover the X-ray pose of 153 pharmaceutically relevant protein-ligand complexes and for their capacity to discriminate known inhibitors of an enzyme from randomly chosen "druglike" molecules. Unfortunately, both properties are not found to be correlated since GOLD showing the best docking accuracy is the less successful in ranking known inhibitors in docking experiment. A speed comparison demonstrated that FlexX was thefastest. On the other hand, the best known docking algorithms often fail to position the ligand in anorientation close to the experimental binding mode this is what we call false positives, GOLD was shown to be the worst in ranking the top ten solutions. Moreover, the current study pinpoints one physicochemical descriptor of the ligand which is flexibility that generally lead to docking/scoring inaccuracies.

General Terms
Comparative data, docking algorithms.

Keywords
Docking programs; FlexX; GOLD; database screening; compared; druglike; docking accuracy; speed; docking/scoring

1. INTRODUCTION
The research for new drugs was always the major concern of health researchers this is why docking small-molecular-weight ligands to the appropriate macromolecules has become a major computational method for predicting protein–ligand interactions and for drug design [1]. Because of the large number of docking tools available, before we decide which docking engine we have to use in our research work, several questions are often asked by us: (1) is the docking algorithm able to reproduce the X-ray pose of the selected small-molecular-weight ligands?; (2) are the fast-scoring functions able to predict binding free energies from the best-scored pose?; (3) is the scoring function able to discriminate known binders from randomly chosen molecules in virtual screening tests ? [2]. So, analyzing all these data for a comparative study of available docking tools seems to be very difficult. Consequently, we decided to explore which of the two more used and efficient docking programs GOLD and FlexX do indeed find experimental solutions for target-ligand complexes. We also wished to explore which docking algorithms perform best in ranking the experimental binding

mode of the ligand and what is the rate of failures in ranking the top ten solutions for each one, speed and one of the physicochemical properties of the ligand was also a part of this study.

The accessibility and the use of conventional file formats as input (e.g., pdb, sdf, mol2), and the easy application to virtual screening (database docking) are the essential motivations in choosing the two programs [3].

2. MATERIAL AND METHODS
2.1 Algorithms
The algorithms implemented by the two programs FlexX and GOLD are so different, the subsubsections below summarize the working principle in each one.

2.1.1 FlexX
FlexX [4] uses an incremental reconstruction algorithm. In this latter, base fragments are identified first, after that the selected fragment is placed into the active site of the receptor using a hashing technique. The complete ligand is constructed by adding the remaining components one after the other. At each time of reconstruction a specified number of optimal partial solutions are selected for the next extension time. In FlexX the scoring is done using a modified Böhm scoring function, which includes the following terms: entropic; hydrogen bonding; ionic; aromatic; and lipophilic.

2.1.2 GOLD
A genetic algorithm is used by GOLD [5]. The ligand's state is encoded by a chromosome, representing its conformation and hydrogen bonding. The conformation of the ligand is represented by a binary string, in which every byte encodes for one torsional angle. Each torsion is allowed to vary between −180◦ and 180◦ in step-sizes of 1.4 Å. Two integer strings encode mappings suggesting possible hydrogen bonds between the protein and the ligand. The first of these strings encodes a mapping of acceptors in the ligand to the donor hydrogens in the protein. The second string encodes a mapping of donor hydrogens in the ligand to the acceptor atoms in the protein. On decoding a chromosome, GOLD utilizes least-squares fitting to form as many of these hydrogen bonds as possible. In the evolutionary development of the ligand conformations the program employs an island model, in which several subpopulations of chromosomes are created at the beginning

instead of one large population. The genetic operations include migration of individual chromosomes between the subpopulations, crossover and mutation. In order to preserve diversity within the population GOLD employs a niching technique, namely, when adding a new individual to the population, the number of individuals in the population that inhabit the same niche as the new chromosome is determined. If there are more than a specified number of individuals in the niche, then the new individual replaces the worst member of the niche rather than the worst member of the total population. Two individuals share the same niche if the RMSD between their donor and acceptor coordinates is less than 1.0 Å. The fitness of a new individual is assessed using a scoring function, which includes energy terms accounting for hydrogen bonding, short-ranged van der Waals interaction between the ligand and protein, and the ligand internal energy. The latter is a sum of ligand steric and torsional energies.

2.2 Setting up a data set of 153 protein–ligand complexes

The crystal structure of 153 protein–ligand complexes from the PDB [6] were used to create a separate set of coordinates for the protein, its ligand, and the corresponding active site, the input conformation of the ligand was extracted from the X-ray structure. The protein active site was defined as the collection of amino acids for which at least one atom is nearer than 6.5 Å to any nonhydrogen atom of the bound ligand. Important metal ions and cofactors were included in binding sites. all crystallographic water molecules were removed from the active site. Hydrogen atoms were added using SYBYL 6.9 (TRIPOS Associates; St. Louis, MO) standard geometries.

2.3 Docking protocols

The receptor is treated as a rigid body in the docking process. In this situation docking will be more efficient, which is especially crucial in database screening [7].

2.3.1 FlexX 1.3.0

FlexX needs a MOL2 format file for the ligand and a PDB format file for the receptor. The conformational flexibility of the ligand is modeled by a discrete set of preferred torsional angles for acyclic single bonds. The rings were considered rigid, since the program *CORINA* for treating multiple conformations of the rings was not included in the distribution. The active site and the interaction surface of the receptor were defined by using a reference ligand and a 6.5 Å cutoff distance. Base fragments were selected automatically. The maximum number of base fragments was 4. The base fragment was placed into the active site using two algorithms. The first one superimposes triples of interaction centers of a base fragment with triples of compatible interactions in the active site. The second algorithm, called matching, is used when the base fragment had fewer than three interaction centers.

2.3.2 GOLD 5.0.1

GOLD uses the receptor and ligand in any of the following formats: PDB, MOL, SDF or MOL2 format. The active site that has the radius of 10 Å was defined by the reference ligand. The default parameters used were: number of islands was 5, population size was 100, number of genetic operations was 100,000 and niche size was 2.

3. RESULTS AND DISCUSSION

Docking results are discussed in the light of the 3 major issues in the application of docking programs to virtual screening: docking accuracy, ranking accuracy, and speed. These criteria were assessed on a data set of 153 diverse protein–ligand complexes from the PDB.

3.1Docking accuracy

The main criterion of a qualified docking program is its ability to reproduce the experimental binding modes of ligands. To test this, a ligand is taken out of the X-ray structure of its protein–ligand complex and docked back into its binding site. The docked binding mode is then compared with the experimental binding mode, and a root-mean-square distance (RMSD) between the two is calculated; a prediction of a binding mode is considered successful if the RMSD is below a certain value (usually 2.0 Å) [8]. Recently, Nissink et al. pointed out that to establish the success rate of a docking program, a large and carefully constructed set of protein–ligand complexes is required. [9] From here on, the "best pose" is defined as the docking solution that is the nearest to the experimental binding mode, whereas the "top pose" is defined as the docking solution that is ranked first. The ability to predict the correct binding of a ligand into its active site was thus evaluated by comparing the best pose and the experimentally determined solution. The ability to predict the correct binding of a ligand into its active site was thus evaluated by comparing the best pose and the experimentally determined solution. Most good RMSD is in the range] 0.5 Å-1, 0A] for GOLD and FlexX (see Fig1). Fig 1 shows also that within 2 Å of the X-ray pose , docking is successful for 64,05% of the cases using GOLD .At this cutoff, FlexX only achieve successful docking in 55,56% of the cases, respectively.

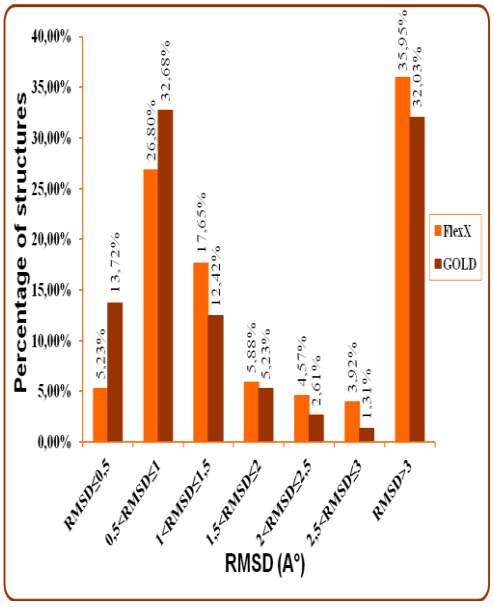

Fig 1 the performance of both programs according to the best docking pose generated

Our result confirms the results obtained by Zaheer et al. In 2010 [10] where six docking programs were used: FRED, GOLD, MOE, AutoDock, and FlexX SURFLEX-Dock for a

comparative study to determine their ability to reproduce poses via the experimental RMSD using 26 complex of Acetyl cholinesterase, FRED was the best followed SURFLEX-Dock and GOLD, other programs such as FlexX, AutoDock and MOE showed a slightly lower performance in the generation of poses.

Michael et al. [11] evaluated in the same year the performance of the four programs GOLD, AutoDock, Dock-SURFLEX FRED by calculating the RMSD using inhibitors of the sarcoplasmic endoplasmic reticulum calcium ATPase, the best results were obtained by GOLD and FRED.

3.2 Ranking accuracy

The ability of both programs GOLD and FlexX to rank various poses generated for each ligand docked into the active site of the target protein is also studied for the 153 complexes of protein-ligand . It is obtained by comparing the solution ranked first (top-posed) by the two programs and that determined by crystallography appreciated by the RMSD [12]. (see Fig 2)

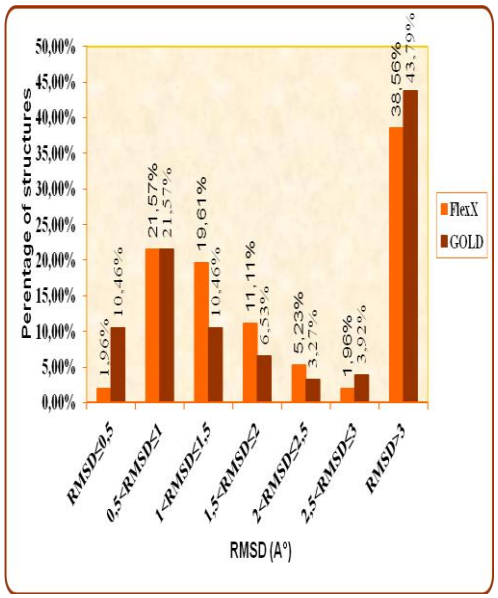

Fig 2 the performance of the two docking programs according to the best pose classified.

Despite that there are equal proportions in the ability of classification of predicted solutions between the two programs GOLD and FlexX in the range of RMSD] 0.5 Å-1, 0A] with a value of 21.57% and although GOLD FlexX exceeds a value of 10.46% in the RMSD between 0 and 0.5 Å, FlexX was able to locate 54.25% of ligands to values of RMSD ≤ 2.0 Å.

If we compare FlexX to GOLD, it could classify only 49.02% of the ligands to RMSD ≤ 2.0 Å, this allows us to say that GOLD didn't well in the classification of solutions, such result is confirmed by earlier studies by Kontoyianni et al. in 2004 [13] and an observation already made by the developers of GOLD, Jones et al. in 1997 [14] stating that GOLD is more efficient in docking than in ranking.

3.3 Rotatable bounds and the performance of the docking process

The influence of the ligand characteristic's on the accuracy of the docking process was also a part of our study (see Fig 3). It is well known that when the number of rotatable bonds of the ligand increases docking accuracy decreases since a much larger conformational space must be sampled [12].

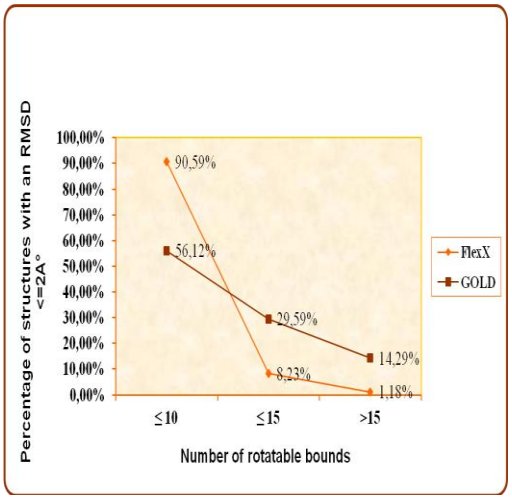

Fig 3 rotatable bounds of the ligand in relation with the precision of the docking process

The complexes in this study were divided into three groups, the ligands with rotatable bounds ≤ 10, the ligands with rotatable bounds ≤ 15 and those with rotatable connections > 15 [12]. The results confirm previous work. Indeed all docking procedures usually fail in the placement of the ligand in the active site when the flexibility of the ligand increases [10].

Taking into account the complexity of the ligands, GOLD seems to be less sensitive, because FlexX reproduced up to 90.59% of experimental poses when the number of rotable bonds of the ligands is less than or equal to 10, this percentage increases even more and reached to 98.82% when the ligand has 15 rotables links or less. Our finding is consistent with that of Kramer et al. [4] who have already made with previous analyzes of FlexX in 1999.

3.4 Speed and docking process performance

The analysis of the execution time of the process of docking or CPU time according to three time intervals: minimum, average and maximum indicates clearly significant differences (see Table1)

Table1 Execution Time docking process for both programs GOLD and FlexX

CPU time (secondes)	FlexX	11,88	109,16	323,91
	GOLD	150,0012	2248,0016	12010,0005

The results in Table 1, shows that FlexX is remarkably faster because it is able to dock a ligand in less than a minute, while GOLD achieved a remarkable distribution of execution time (150s-12010s) yet it is much less sensitive than FlexX in terms of flexibility of the ligand , these results agree well with those obtained by Badry et al. in 2003 [7] and Kellenberge et al. in 2004 [3] showing that GOLD is much slower than FlexX in docking process.

According to the results discussed above, FlexX is a good program for virtual screening and to participate in the process of discovery of new bioactive molecules.

3.5 False positives

This parameter is analyzed taking into account only the best ten solutions classified with the two programs GOLD and FlexX. showed no or few structures with large rmsd values. It is however not rare that the nearly optimal solutions are commonly lost among an ensemble of high-scored solutions with large rmsd values; these are called false positives. [15]. for each program, we try to identify in the the list of the best 10 solutions for each ligand docked, the percentage of solutions with high score and high RMSD value. (see Fig 4)

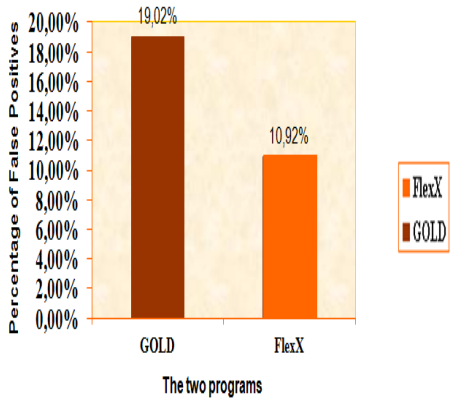

Fig 4 Percentage of False Positives in the ten best solutions for all docked ligands for each program.

The results clearly show that GOLD produced more errors than FlexX in the classement of the top ten best poses because of the difference between the two programs in the scoring functions. The expression of the GOLD fitness score is unable to distinguish between the complexes obtained and the structure of the complex in its native form [15], this is explained by the fact that in some cases, it was observed that the ligands were placed in the active sites of the protein, but did not adopt the conformation that allows them to be correctly oriented in the active site resulting in high values of RMSD, thus resulting installation has been incorrectly classified [16].

3.6 Role of water molecules in the docking accuracy

Water molecules are important in the performance of ligand-protein docking predictions. It can be involved in protein ligand recognition either by forming mediating hydrogen bonds between the protein and the ligand or by being displaced by the ligand; both of these mechanisms have been shown to be of importance to drug discovery [17].

In order to find out the role water molecule plays in the docking pose prediction, water molecules were added in the active site of proteins then the ligands were docked.

FlexX 1.3.0 gave the results summarized in Fig 5

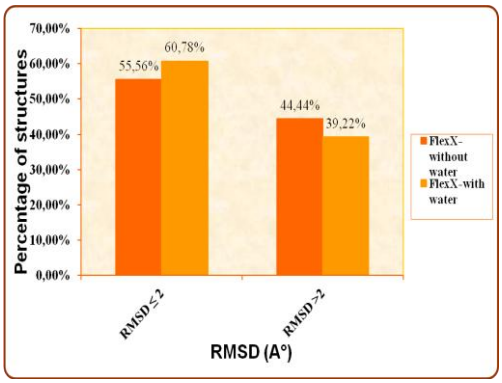

Fig5 results of docking with FlexX 1.3.0 in the absence and presence of water molecules according to the best pose generated

GOLD was found to be similar in identification of good and fair poses either with or without water molecules.

The ability of FlexX to generate correct poses was reduced in the absence of interactions with water molecules, this was the fact that in the absence of water molecules can't complete the good way of construction .

4. CONCLUSION

Two popular docking tools have been compared on common data sets for both docking accuracy. It is not our intention in the current study to propose a hierarchy of available docking programs but to notice advantages and drawbacks of selected tools in different contexts. The findings of our comparative study revealed substantial differences in the performance of commonly used programs for docking of 153 complexes. This observation underscored the need for an individual evaluation of available software for a given inhibitor class and receptor type. The best overall results were obtained with GOLD, as far as docking accuracy and reproducibility were concerned. However, it is important to note that good docking accuracy is necessary but not sufficient for accurate screening utility. As a matter of fact, we made the choice of examining docking tools

from a virtual screening perspective, which means using settings compatible with fast docking. Hence, we believe that speed is nowadays an important aspect of computational drug discovery techniques, FlexX was faster than GOLD and the best in ranking accuracy and the more sensitive in term of rotatable bounds of ligand but the less in doing errors in the classement of the top ten best poses. Additionally, role of water molecule in the docking experiments were also discussed in the detailed. In conclusion this set up will definitely aid in our ongoing projects and to the community having same research interests.

5. ACKNOWLEDGMENTS

We wish to thank the Laboratory of Applied Biology and Health, Department of Biochemistry and Microbiology, Faculty of natural and life Sciences, Mentouri University, Constantine, Algeria for the accomplishment of this work.

6. REFERENCES

[1] Brooijmans N, Kuntz ID. 2003, Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct; 32:335–373.

[2] Zavodszky MI, Sanschagrin PC, Korde RS, Kuhn LA. 2002, Distilling the essential features of a protein surface for improving protein– igand docking, scoring, and virtual screening. J Comput Aided Mol Des; 16:883–902.

[3] Kellenberger E., Rodrigo J., Muller P., Rognan D., 2004, Comparative evaluation of eight docking tools for docking and virtual screening accuracy, Proteins, 57(2), 225–242.

[4] Kramer B, Rarey M, Lengauer T. 1999, Evaluation of the FLEXX incremental construction algorithm for protein–ligand docking. Proteins; 37:228–241.

[5] Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., 1998, J. Comput. Chem., 19; 1639.

[6] Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C. 2002, The Protein Data Bank. Acta Crystallogr D Biol Crystallogr ;58:899 –907.

[7] Badry D. B., Maxim T., Ruben A., Brooks III. C.L. 2003, Comparative study of several algorithms for flexible ligand docking. J.Comput-Aided. Mol. Des, 17, 755–763.

[8] Verdonk M.L., Cole J.C., Hartshorn M.J., Murray C.W., Taylor R.D., 2003, Improved protein–ligand docking using GOLD. Proteins, 52, 609–623.

[9] Nissink JWM, Murray CW, Hartshorn MJ, Verdonk ML, Cole JC, Taylor R. 2002, A new test set for validating predictions of protein– ligand interaction. Proteins; 49:457–471.

[10] Zaheer U.H., Sobia A.H, Reaz U., Jeffry D., Madurab B., 2010, Benchmarking docking and scoring protocol for the identification of potential acetyl cholinesterase inhibitors, J. Mol. Graph. Model, 28, 870–882.

[11] Michael L., Christopher E., Stefan P, 2010, Comparison of current docking tools for the simulation of inhibitor binding by the transmembrane domain of the sarco/endoplasmic reticulum calcium ATPase. Biophys. Chem, 150, 88-97.

[12] Chikhi A., Bensegueni A, 2008, Docking Efficiency Comparison of Surflex, a Commercial Package and Arguslab, a Licensable Freeware. J Comput Sci Syst Biol, 1, 081-086.

[13] Kontoyianni M., McClellan L.M., Sokol G.S, 2004, Evaluation of docking performance: comparative data on docking algorithms, J. Med. Chem, 47, 558–565.

[14] Jones G., Willett P., Glen R.C., Leach A.R., Taylor R, 1997, Development and validation of genetic algorithm for flexible docking. J. Mol. Biol, 267, 727-748.

[15] Becue A., Meurice N., Leherte L., Vercauteren D.P. 2008, Protein-Protein Docking Using Three-Dimensional Reduced Representations and Based on a Genetic Algorithm.. Boeyens J.C.A and Ogilvie J.F.Models, Mysteries, and Magic of Molecules, Springer: Guildford, 301-324.

[16] Verkhivker G.M., Bouzida D., Gehlhaar D.K., Rejto P.A, Arthurs S., Colson A.B., Freer S.T., Larson V., Luty B.A., Marrone T., Rose P.W., 2000 , Deciphering common failures in molecular docking of ligand–protein complexes, J. Comput.-Aided. Mol. Des,14, 731–751.

[17] Roberts B.C., Mancera R.L, 2008, Ligand-protein docking with water molecules, J.Chem. Inf. Model, 48, 397–408.