XIX SBQT
Simpósio Brasileiro de Química Teórica 2017
12 a 17/Nov, 2017, Águas de Lindóia/SP, Brasil

DockThor 2.0: a Free Web Server for Protein-Ligand Virtual Screening

Isabella A. Guedes[1], Eduardo Krempser[2], Laurent E. Dardenne[1]

1 Laboratório Nacional de Computação Científica – LNCC/MCTI, Av. Getúlio Vargas, 333, Petrópolis – RJ, Brazil
2 Fundação Oswaldo Cruz – Fiocruz – Rio de Janeiro – RJ - Brazil

Abstract: INTRODUCTION. Receptor-ligand molecular docking is a structure-based approach widely used by the scientific community in Medicinal Chemistry. The main objective is to assist the process of drug discovery, searching for new lead compounds against relevant therapeutic targets with known three-dimensional structures [1]. The program DockThor [2,3], developed by the group GMMSB/LNCC, has obtained promising results in comparative studies with other well-established docking programs for predicting experimental binding modes, considering several molecular targets and chemical classes of ligands. The first version of the DockThor portal was firstly available for binding mode prediction and was strict to the docking of a single ligand at a time. Despite useful for pose prediction, the scoring function implemented in the first version of the DockThor portal was not suitable for predicting binding affinities, which is crucial for virtual screening studies. Recently, our research group developed new empirical scoring functions to predict protein-ligand binding affinities that demonstrated to be competitive with the best scoring functions reported in the literature [4]. Such encouraging results motivated the development of a new version of the Web server for virtual screening. The DockThor 2.0 portal utilizes the computational facilities provided by the SINAPAD Brazilian high-performance platform (www.lncc.br/sinapad) and the Santos Dumont supercomputer. **MATERIALS AND METHODS.** The DockThor program has implemented a grid-based method that employs a steady-state genetic algorithm for multiple solutions as the search engine and the MMFF94S force field as the scoring function for pose evaluation. The web server provides the major steps of ligand and protein preparation, being possible to change the residues protonation states and to define the degree of flexibility of the ligand. It is possible to perform virtual screening experiments with a maximum of 100 compounds as a guest user or 1000 compounds for registered users with an approved project. The main parameters of the grid box and the genetic algorithm can also be customized. We recently developed general and specific scoring functions for target-classes, the last to account for binding characteristics associated with a target class of interest, focusing on proteases, kinases and protein-protein interactions complexes (PPIs). The scoring functions were derived using linear regression (MLR) and more sophisticated machine learning techniques for nonlinear problems using the PDBbind refined set 2013 (N = 2959) for training and

testing. We also trained and evaluated general scoring functions using docking results obtained with DockThor. **DISCUSSION AND RESULTS.** The DockThor program has obtained very satisfactory results in redocking experiments using benchmarking datasets considering diverse RMSD values as success criterion, achieving performances of 78%, 83.33% and 78% in the Astext diverse (N = 85), Iridium-HT (N = 120) and PDBbind 2013 core set (N = 195), respectively, for the top-ranked energy pose. Furthermore, our scoring functions obtained promising performances when evaluated in both experimental and docking structures, indicating that they are reliable to be applied in both cases, with the best one achieving a high correlation with measured binding data (R = 0.705) from the PDBbind core set v2013. In the DockThor portal, docking results are automatically clustered and ordered by an internal analyses tool. The parameters of the analysis step may also be customized by the user, as the number of different binding modes and compare them with a reference conformation of the ligand through the RMSD calculation. Compounds docked in virtual screening studies are ranked according to the score provided by the general and linear scoring function developed by our research group specifically for the DockThor program. Such scoring function was also validated for virtual screening using the DUD-E dataset for several protein targets with active and decoy compounds. For example, the scoring function developed for the DockThor program obtained an AUC = 0.722 when evaluated on a set of 976 active and 25980 inactive compounds for the protease Trypsin [5]. **CONCLUSION.** The new version of the DockThor Portal is a free Web server for protein-ligand docking and virtual screening experiments available for the scientific community at the address www.dockthor.lncc.br/v2.

Key-words: virtual screening, protein-ligand docking, drug design.
Support: This work has been supported by FAPERJ, CNPq and CAPES.

References:

[1] I. A. Guedes, C. S. de Magalhães, and L. E. Dardenne, “Receptor–ligand molecular docking,” *Biophys. Rev.*, vol. 6, no. 1, pp. 75–87, Mar. 2014.
[2] C. S. de Magalhães, H. J. C. Barbosa, and L. E. Dardenne, “Selection-Insertion Schemes in Genetic Algorithms for the Flexible Ligand Docking Problem,” in *Genetic and Evolutionary Computation – GECCO 2004*, vol. 3102, K. Deb, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 368–379.
[3] C. S. de Magalhães, D. M. Almeida, H. J. C. Barbosa, and L. E. Dardenne, “A dynamic niching genetic algorithm strategy for docking highly flexible ligands,” *Inf. Sci.*, vol. 289, pp. 206–224, Dec. 2014.
[4] I. A. Guedes, A. M. S. Barreto, M. A. Miteva, and L. E. Dardenne, “Development of Empirical Scoring Functions for Predicting Protein-Ligand Binding Affinity,” Laboratório Nacional de Computação Científica, Petrópolis - RJ, 2016.
[5] F. S. S. Pereira, I. A. Guedes, and L. E. Dardenne, “Avaliação de Funções Empíricas de Scoring em Experimentos de Triagem Virtual em Larga Escala,” 2017.