P-33: A machine learning-based protocol for docking results analysis

Sabina Smusz[1,2], Stefan Mordalski[1], Jagna Witek[1], Krzysztof Rataj[1], Andrzej J. Bojarski[1]

[1]Department of Medicinal Chemistry, Institute of Pharmacology Polish Academy of Sciences, Kraków, Poland, [2]Faculty of Chemistry, Jagiellonian University, Kraków, Poland

Docking is an important part of virtual screening campaigns and belongs to the group of the most popular chem- and bioinformatics procedures [1]. Its aim is to predict the conformation of the ligand and the receptor in their complex. A major challenge, however, is still connected with the analysis of the results. Although advanced scoring schemes were developed for the prediction of interaction energies between ligand and the protein, they still do not fully solve the problem, nor the visual inspection, which is very time-consuming and subjective, especially in case of large number of diverse compounds.

In this study, a novel protocol for automatic evaluation of massive docking results is proposed. It is a combination of the description of docking results in the form of a string with machine learning approach [2].

The docking results are described by means of Structural Interaction Fingerprints (SIFts) [3] and Spectrophores [4]. SIFts provide information about the interactions between ligand and each of the amino acids of the receptors, whereas Spectrophores are the source of information about the conformation of the docked compound, as they consist of atomic properties values calculated in a way that is dependent on the actual spatial orientation of a molecule. Such prepared representation of ligand-receptor complexes constitutes an input for machine learning experiments, with the use of 5 different classification algorithms, followed by multi-step results analysis taking into account the quality of the model of the receptor structure, various conformations of the docked compound and the performance of particular docking algorithm.

The pilot studies were performed for the serotonin receptors 5-HT_6 and 5-HT_7, however the tool is constructed in a way enabling its application for any target.

Acknowledgments: The study is supported by a grant PRELUDIUM 2013/09/N/NZ2/01917 financed by the National Science Centre.

1. Breda, A.; Basso, L. A; Santos, D. S.; De Azevdo, W. F., Jr; Virtual Screening of Drugs: Score Functions, Docking and Drug Design. *Curr. Comput. Aided Drug. Des.* **2008**, *4*, 265–272.
2. Kotsiantis, S. B. Supervised Machine Learning: A Review of Classification Techniques. *Informatica* **2007**, *31*, 249–268.
3. Deng, Z.; Chuaqui, C.; Singh, J. Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions. *J. Med. Chem.* **2004**, *47*, 337–344.
4. Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Van Alsenoy, C.; Tollenaere, J. P. The Electronegativity Equalization Method II: Applicability of Different Atomic Charge Schemes. *J. Phys. Chem.* **2002**, *106*, 7895–7901.