[P1] Druggability Prediction Performances Related To Different Pocket Estimations

Alexandre Borrel[1,2,3], Leslie Regad[1,2], Henri Xhaard[3], Michel Petitjean[1,2] and Anne-Claude Camproux[1,2]

[1] *INSERM, U973, F-75205 Paris, France*
[2] *Univ Paris Diderot, Sorbonne Paris Cité, UMRS 973, MTi, F-75205 Paris, France*
[3] *Centre for Drug Research, Faculty of Pharmacy, FI-00014 University of Helsinki, Finland*

Therapeutical molecules bind to preferred sites of action, which are in the majority of cases pockets located within proteins or at their surface. Therefore, estimation and characterization of pockets is a major issue in drug target discovery. Among the molecules, “drug-like molecules” [1] are small molecules with particular properties as of small size, able to cross the digestive tract. Pocket “druggability”, the ability of a pocket to bind “drug-like” molecules, is essential for drug discovery studies [2] especially for discovering new targets.

Currently, identifying druggable pockets is possible by different statistical models of prediction [3, 4, 5]. These methods differ by methods used to estimated pockets, by descriptors used to characterized pockets and the statistical methods used. Moreover, the quality of these approaches is limited by the few data available, and most of them allow the prediction of the pocket druggability if the structure of the target is complexed to one ligand (holo forms). However of new target discovery, it is important to be able to predict the “druggability” of a pocket in its apo form that means when it is not yet bound to a ligand and deformed by the interaction with one ligand.

Here, we propose a model to predict pocket druggability from holo or apo form. To develop this model, we started from a set of 113 complexes protein ligands [6], with 71 druggable proteins and 41 less druggable proteins. From this set, we used three approaches to estimate pockets by taking, defines pockets as protein atoms less than 4 Å away from the ligand or not the ligand information, based on two algorithms Fpocket [7] and DoGSite [8].

Pockets estimated using three approaches, were then characterized using a set of 57 descriptors. This descriptor set, named pocket profile, allows a characterization of the geometry and the physicochemical properties of pockets. We then built statistical models based on a linear discriminant analysis to predict the pocket druggability from pocket estimated by Fpocket. The construction of this model consisted in the selection of the models with the best accuracy and containing as few descriptors as possible. Finally we used a consensus of 4 best models which present a very good accuracy (close to 80% on average) from pocket set estimated by different pocket estimators and also from apo pockets not complexed with a ligand.

[1] Lipinski, C., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Advanced Drug Delivery Reviews, 46(1-3), 3–26
[2] Nisius, B., Sha, F., & Gohlke, H. (2011). Journal of biotechnology, 159(3), 123–134
[3] Perola, E., Herman, L., & Weiss, J. (2012). Journal of chemical information and modeling, 52(4), 1027–1038
[4] Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., & Rarey, M. (2012). Journal of Chemical Information and Modeling, 52(2), 360–72.
[5] Desaphy, J., Azdimousa, K., Kellenberger, E., & Rognan, D. (2012). Journal of chemical information and modeling, 52(8), 2287–99
[6] Krasowski, A., Muthas, D., Sarkar, A., Schmitt, S., Brenk, R. (2011). Journal of chemical information and modeling, 51(11), 2829–42
[7] Le Guilloux, V., Schmidtke, P., & Tuffery, P. (2009) BMC bioinformatics, 10, 168
[8] Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). Proteins, 81(October)