Discovery of a novel relationship between two proteins by a chemogenomics analysis

Lina Humbeck, Jette Pretzel, Oliver Koch

Faculty of Chemistry and Chemical Biology, TU Dortmund University, Germany

The term “privileged scaffolds” was coined for the collective core structure of multiple molecules exhibiting bioactivity on different targets [1]. Within proteins, conserved structural elements are similarly common. A recently discovered level of conservatism, the ligand-sensing core, is a similar spatial composition of secondary structure elements around the ligand binding site in proteins with distinct folding patterns that can bind similar scaffolds [2]. Knowledge about ligand-sensing cores facilitates rational identification of new lead structures [3] or prediction of polypharmacology [2].

Compound databases like DrugBank [4] or ChEMBL [5] contain a wealth of data about molecules and their bioactivity on diverse proteins. Hence, a python-based tool for knowledge discovery aiming at new insights into the relationship of privileged scaffolds and ligand-sensing cores was developed. Its main objective is the identification of scaffolds that bind to unrelated proteins for revealing conserved structural elements. In a first step, a command line version of Scaffold Hunter [6] assigns scaffolds to all imported molecules. Afterwards a sequence similarity analysis of proteins whose ligands share a scaffold is performed. Only protein targets with identity below 40 % are considered as unrelated. Finally, the results are visualized for an in depth analysis.

We will present the overall workflow and the result of a chemogenomics analysis of the DrugBank. Around 1500 scaffolds were identified that bind to different proteins. An analysis of one of these scaffolds already ended up in a new ligand-sensing core that is shared between five different proteins. Based on this information an enriched library of molecules that show a similarity to known inhibitors of four of these proteins was selected. Testing this library for inhibitory activity against the fifth protein led to IC_{50} values down to the nanomolar range and to an initial hit rate of ~11 % within the molecule series that was selected based on known inhibitors of one of the proteins. This clearly indicates a relationship and similar ligand binding of one pair of these proteins sharing a similar ligand-sensing core and proves the usefulness of this approach. Currently, we investigate the hits using orthogonal assays and crystallization experiments to solve complex structures with the most promising hits.

[1] M. E. Welsch, S. A. Snyder, B. R. Stockwell, *Curr Opin Chem Bio*, **2010**, *14*, 347-361.
[2] O. Koch, *Fut Med Chem*, **2011**, *3*, 699-708.
[3] D. Willmann, S. Lim, S. Wetzel, E. Metzger, A. Jandausch, W. Wilk, M. Jung, I. Forne, A. Imhof, A. Janzer, J. Kirfel, H. Waldmann, R. Schüle, R. Buettner, *Int J Cancer*, **2012**, 131:2704-2709.
[4] V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, A. Tang, G. Gabriel, C. Ly, S. Adamjee, Z. T. Dame, B. Han, Y. Zhou, D. S. Wishart, *Nucleic Acids Res*, **2014**, *42*, D1091-1097.
[5] A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, J. P. Overington, *Nucleic Acids Res*, **2014**, *42*, 1083-1090.
[6] K. Klein, O. Koch, N. Kriege, P. Mutzel, T. Schäfer, *Mol Inf*, **2013**, *32*, 964–975.