Using linked open data for assessing multi-target SAR

Barbara Zdrazil,[1] Daniela Digles,[1] Edgar Jacoby,[2] Stefan Senger,[3] Gerhard F Ecker[1]

[1]*University of Vienna, Department of Pharmaceutical Chemistry, Division of Drug Design and Medicinal Chemistry, Althanstraße 14, A-1090 Vienna, Austria*

[2] *Janssen Research & Development, Turnhoutseweg 30, Beerse, Belgium*

[3] *GSK Medicines Research Centre, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, United Kingdom*

Correspondence to: barbara.zdrazil@univie.ac.at

Finding new drug candidates that have the potential to become effective medicines has become increasingly challenging as our understanding of the relevant disease biology has developed. Whereas in the past the 'Holy Grail' was to identify molecules that selectively modulate just one biological target, there is now the realisation that complex diseases require more advanced approaches, for example, investigations that include the extension of traditional single-target structure-activity-relationship (SAR) studies towards delineating approaches, which deliver therapeutically relevant multi-target activity profiles. This strategy offers new opportunities towards a more holistic and mechanistic view of target modulation by small molecules in biological context.

Since open life science data and collaborative drug discovery projects are enabling academic researchers to enter a world of increasing abilities when mining, interpreting, and extrapolating data, concurrently a need for computational approaches that help to manage and process the vast amount of information is emerging.

In an attempt to obtain a convenient, fully flexible workflow for visualizing and retrieving multi-target SAR datasets in an automated fashion, which is able to deliver small molecule information in the context of biological pathways, we designed a KNIME-based workflow [1], integrating data from the Open PHACTS Discovery Platform [2]. An essential step of our workflow is the clustering of the pharmacology data on the basis of Bemis-Murcko scaffold [3] identity. Thus, data dimensionality can be reduced, and SAR profiles of congeneric data series can be inspected. Since integration of data from different providers is an essential element of Open PHACTS, data provenance is of huge importance. As a consequence, our workflow not only delivers congeneric datasets, but it is also capable of conveniently comparing assay descriptions, or underlying literature sources.

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement n° 115191 (Open PHACTS), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

[1] Berthold M, Cebron N, Dill F, Gabriel T, Kötter T, et al.. (2008) KNIME: The Konstanz Information Miner. In: Preisach C, Burkhardt H, Schmidt-Thieme L, Decker R, editors.: Springer Berlin Heidelberg. pp.319–326.

[2] Williams AJ, Harland L, Groth P, Pettifer S, Chichester C, et al. (2012) Open PHACTS: Semantic interoperability for drug discovery. Drug Discov Today 17:1188–1198. doi: 10.1016/j.drudis.2012.05.016

[3] Bemis GW, Murcko MA (1996) The Properties of Known Drugs. 1. Molecular Frameworks. J Med Chem 39: 2887–2893.