Escape from ‘availability bias’ in compound design

A. Stracz[1], A. Tarcsay[1], G. Imre[1], I. Solt[1]

[1]ChemAxon

Small molecule design is an information demanding activity, since all relevant knowledge is to be accessible within a single space and requires synchronized application of computational models to assist decision making on synthesis candidates. Our study aims to evaluate a software platform coping with this complexity (Marvin Live[1]). The tool provides central management of innovative ideas and helps triage them based on predicted properties and available knowledge collected from a variety of sources. The calculated properties span phys-chem descriptors, combined metrics like MPO score[2], 3D overlay and modelling results conducted with KNIME. Use cases of rapid freedom to operate analysis by ultra-fast searching (MadFast Similarity Search[3]) of exemplified structures from patents (SureChEMBL[4], ~16M entries) and SAR by catalog via searching large set of synthesizable compounds (Enamine REAL DataBase[5], ~170M entries) real time will be shown to ensure that designers can seamlessly exploit the chemical space around their ideas. The presentation will walk through an example drug design cycle to obtain statistical results regarding performance as well as to demonstrate the suitability of the calculations.

[1] https://www.chemaxon.com/products/marvin/marvin-live/
[2] Wager, Travis T., et al. "Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties." *ACS chemical neuroscience* 1.6 (**2010**): 435-449.
[3] https://www.chemaxon.com/products/madfast/
[4] Papadatos, George, et al. "SureChEMBL: a large-scale, chemically annotated patent document database." *Nucleic acids research* (**2015**): gkv1253.
[5] Shivanyuk, A. N., et al. "Enamine real database: making chemical diversity real."*Chemistry Today* 25 (**2007**): 58-59.