[L8] Adaptive Molecular Design – Learning from Data

Gisbert Schneider

Department of Chemistry and Applied Biosciences, Swiss Federal Institute of Technology (ETH) Zurich, Switzerland

Good compounds are overlooked for various reasons. Computers can help by examining molecular features no chemist can see. Identifying promising candidates (positive design) is equally important as eliminating the bad apples to avoid undesired effects (negative design) as early as possible in the drug discovery process. While medicinal chemists excel in optimizing hits to eventually become lead structures and enter clinical trials, the computer's domain is to rapidly sift through many millions of molecules to discard the bulk before any screening assay is performed with the selected hits that remain after thorough *in silico* scrutiny. "Big Data" in this context means sifting through chemical space while considering available bioactivity data for navigation. "Deep learning" methods may help in this endeavor. In fact, recent technological advances in both computer hardware and software have enabled a renaissance of "de novo" design of molecules with desired pharmacological properties.

De novo design generally faces three challenges: (i) structure generation, (ii) scoring, and (iii) optimization. The problem of innovative structure generation has generally been solved. While some of the earlier design methods constructed overoptimized, synthetically challenging molecular structures, the current modeling toolbox contains receptor- and ligand-based algorithmic approaches for combining molecular building blocks (*e.g.*, atoms, fragments) in a chemically meaningful way to obtain structurally novel lead- and druglike virtual structures.

We will present our current perspective on the concept of automated molecule generation by highlighting chemocentric methods that may capture druglike chemical space, consider ligand promiscuity for hit and lead finding, and provide fresh ideas for the rational design of customized screening compound libraries. We will specifically focus on natural-product-inspired molecular design by computational means. Recent applications of automated *de novo* design methods will be presented that suggest innovative, synthetically accessible small compounds mimicking structurally more complex natural products.

References:

[1] Schneider, P. & Schneider, G. J. Med. Chem., 59, 2016, 4077-4086.
[2] Rodrigues, T., Reker, D., Schneider, P. & Schneider, G. Nature Chem., 2016 doi:10.1038/nchem.2479.
[3] Friedrich, L., Rodrigues, T., Schneider, P. & Schneider, G. Angew. Chem. Int. Ed., 2016 doi: 10.1002/anie.201601941.
[4] Schneider, P., Röthlisberger, M., Reker, D. & Schneider, G. Chem. Commun. 52, 2016, 1135-113
[5] Gawehn, E., Hiss, J. A. & Schneider, G. Mol. Inf, 35, 2016, 3-14.
[6] Reker, D. & Schneider, G., Drug Discovery Today 20, 2015, 458-465.