Novel Chemical Space Driven By Reaction Network

Xuejin Zhang [♣] [♠], Alexander Weber [♣], Gisbert Schneider [♠], Jan Kriegl[♣]

[♣]Medicinal Chemistry Department, Computational Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach, Germany,

[♠] Institute of Pharmaceutical Sciences, ETH Zürich

Email: xuejin.zhang@pharma.ethz.ch

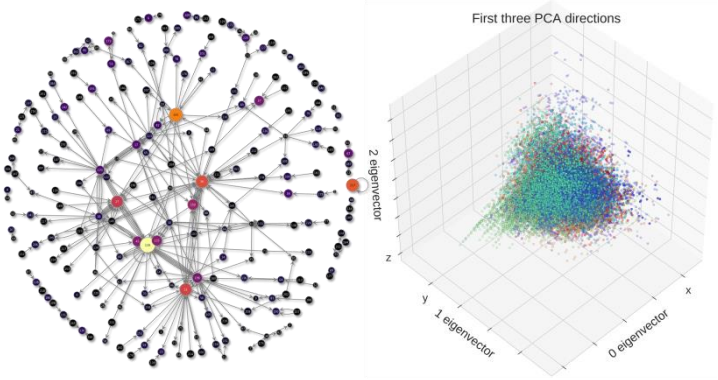

Figure left. The network of chemical reactions.

Figure right. A reaction driven chemical space presented with its first three eigenvectors.

Abstract:

21th century is undoubtedly the "BigData" era[1], especially for chemoinformatics. From reaction record's aspect, there are millions of chemical reactions stored in commercial databases[2]. Figure left exhibits the complexity of chemical reactions via directed network. This number is increasing with text-mining pipelines developed to extract chemical reactions from patents and literatures[3]. From chemical space's view, the cardinality of a typical screening compound collection from a large pharmaceutical company often exceeds one million substances[4], virtual compound pools are even larger. Given these data, one would presume that it should be easy to deliver a real chemical entity with purposed synthetic routes, *i.e.* synthetic route design[5-7]. However, despite the fact that multiple retrosynthetic computational algorithms are available, they are not broadly used by synthetic chemists. Here, we present our naïve synthetic enumeration software guided by easily accessible building blocks and commonly used, *per se*, preferred by chemists, chemical reactions with confined synthetic steps. Through our software, one could deliver a chemical space composed only by theoretically synthesizable compounds. Figure right shows a small novel product pool from ten chosen chemical reactions generated by our software. Automatic property calculation of reactants and products is performed to navigate and analyze enumerated new customized chemical space. We hope that this software could help to establish a rapport between computational and medicinal chemists.

[1] A. Klenner, V. Hähnke, G. Schneider et.al. *Mol. Inf.*, **2012**, 31, 21–26.
[2] S. Roughley, A. Jordan. *J. Med. Chem*, **2011**, 54, 3451–3479.
[3] D. Lowe, *Extraction of chemical structures and reactions from the literature (doctoral thesis)* **2012**.
[4] P. Schneider, G. Schneider, *J. Med. Chem*, **2016**, *59*, 4077-4086.
[5] E. Corey, X. Cheng. *The Logic of Chemical Synthesis, Wiley*, **1989**.
[6] Hartenfeller M, G. Schneider, et al. *PLoS Comput Biol*, **2012**, 8, e1002380.
[7] Wellenzohn, B.; Lessel, U. et.al. *J. Med. Chem*. **2012**, 55, 11031−11041.