J Comput Aided Mol Des (2017) 31:701–714
DOI 10.1007/s10822-017-0033-6

Predictive cartography of metal binders using generative topographic mapping

Igor I. Baskin[1,2] · **Vitaly P. Solov'ev**[3] · **Alexander A. Bagatur'yants**[4,5] · **Alexandre Varnek**[6]

Received: 2 April 2017 / Accepted: 11 June 2017 / Published online: 7 July 2017

Abstract Generative topographic mapping (GTM) approach is used to visualize the chemical space of organic molecules (L) with respect to binding a wide range of 41 different metal cations (M) and also to build predictive models for stability constants (log*K*) of 1:1 (M:L) complexes using "density maps," "activity landscapes," and "selectivity landscapes" techniques. A two-dimensional map describing the entire set of 2962 metal binders reveals the selectivity and promiscuity zones with respect to individual metals or groups of metals with similar chemical properties (lanthanides, transition metals, etc). The GTM-based global (for entire set) and local (for selected subsets) models demonstrate a good predictive performance in the cross-validation procedure. It is also shown that the data likelihood could be used as a definition of the applicability domain of GTM-based models. Thus, the GTM approach represents an efficient tool for the predictive cartography of metal binders, which can both visualize their chemical space and predict the affinity profile of metals for new ligands.

Keywords Generative topographic mapping · Metal binding · Cartography of chemical space · Activity landscapes

Electronic supplementary material The online version of this article (doi:10.1007/s10822-017-0033-6) contains supplementary material, which is available to authorized users.

✉ Alexandre Varnek
varnek@unistra.fr

[1] Faculty of Physics, M.V. Lomonosov Moscow State University, Leninskie Gory, Moscow, Russian Federation 119991

[2] Laboratory of Chemoinformatics, Butlerov Institute of Chemistry, Kazan Federal University, Kazan, Russian Federation

[3] Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences, Leninskiy Prospect, 31, Moscow, Russian Federation 119071

[4] Federal Research Center Crystallography and Photonics RAS, ul. Novatorov 7a, Moscow, Russian Federation 119421

[5] National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoye Shosse 31, Moscow, Russian Federation 115409

[6] Laboratoire de Chemoinformatique, UMR 7140, Université de Strasbourg, 1 Rue Blaise Pascal, 67000 Strasbourg, France

Introduction

Modern trends in chemoinformatics are characterized by a transition from the building of individual models for selected properties to the assessment of the entire profiles of related properties exhibited by chemicals [1, 2]. Chemogenomics, for example, considers the binding of organic molecules with an ensemble of proteins [3]. The profiling of properties, however, is not limited exclusively to the drug design area; the development of any material requires knowledge of variety of properties [4].

Metal binders constitute an important class of low molecular weight materials, which is crucially important for separating various chemical elements and is widely used in analytical chemistry, chemical catalysis, nuclear chemistry, environment protection, and medical diagnostics. It includes polydentate ligands forming strong coordination bonds with metal cations. Because dozens of different metal cations can simultaneously occur in a solution, the issues of binding potency, selectivity, and promiscuity are of prime importance for analyzing the chemical space of metal binders. It should be noted that ligand binding to