

The Freedom space – GTM based analysis of a new chemical space using ChemAtlas tool

Erik Yeghyan¹, Mykola V. Protopopov², Valentyna V. Tararina^{3,4}, Fanny Bonachera¹, Igor M. Dzyuba², Gilles Marcou¹, Olga Klimchuk¹, Dragos Horvath¹, Olena V. Savych², Olga O. Tarkhanova², Alexandre Varnek¹, and Yurii S. Moroz^{3,4},

¹Laboratoire de Chemoinformatique, University of Strasbourg, 4 rue Blaise Pascal, 67000 Strasbourg, France

²Chemspace LLC, Kyiv, Ukraine

³Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

⁴Enamine Ltd., Kyiv, Ukraine

In contemporary drug design, high-performance virtual screening techniques have revolutionized the search for promising candidate compounds, but their success depends critically on the quality and relevance of the screened molecules. To address the challenges of virtual screening and hit discovery, we introduce the Freedom space (version 2.0)^[1], a new dataset of make-on-demand compounds. Inspired by Enamine's REAL space^[2], the Freedom space was developed by combining diverse chemical building blocks and robust transformations, resulting in approximately 200 million diverse compounds with a synthesis success rate exceeding 75%. To demonstrate the uniqueness of the Freedom space, first, a general comparison with the REAL space and the ChEMBL database^[3] was conducted, including scaffold and functional group analyses. These analyses highlighted the distinct differences of the Freedom space. Furthermore, Generative Topological Maps (GTM) analyses using the ChemAtlas tool^{[4][5]} confirmed a clear separation of the chemical space areas. Researchers can use this new dataset to conduct successful discovery projects, enhancing the efficiency and effectiveness of drug discovery efforts, and thereby extending the exploration of synthetically feasible chemical space for hit finding and hit-to-lead campaigns.

Bibliography:

[1] Freedom Space | Chemspace. ChemSpace. <https://chem-space.com/compounds/freedom-space>

[2] REAL Space - Enamine. <https://enamine.net/compound-collections/real-compounds/real-space-navigator>

[3] ChEMBL Database Release 26; EMBL-EBI, 2020. <https://doi.org/10.6019/ChEMBL.database.26>.

[4] Zabolotna, Y.; Bonachera, F.; Horvath, D.; Lin, A.; Marcou, G.; Klimchuk, O.; Varnek, A. Chemspace Atlas: Multiscale Chemography of Ultralarge Libraries for Drug Discovery. *J. Chem. Inf. Model.* 2022, 62 (18), 4537–4548. <https://doi.org/10.1021/acs.jcim.2c00509>.

[5] Ruggiu, F.; Marcou, G.; Varnek, A.; Horvath, D. ISIDA Property-Labelled Fragment Descriptors. *Mol. Inf.* 2010, 29 (12), 855–868. <https://doi.org/10.1002/minf.201000099>.