From chemical space visualization to strategic planning: leveraging chemography in drug discovery at Eli Lilly

Yuliana Zabolotna[1], Eduardo Harguindey[1], Tagir Akhmetshin[2], Juan Perucho[3], Dragos Horvath[2], Alexandre Varnek[2], Juan Felix Espinosa[1]

[1]*Discovery Chemistry Research & Technologies, Lilly Spain, Alcobendas, Madrid, Spain*
[2]*Laboratory of Chemoinformatics, UMR7140, University of Strasbourg, Strasbourg, France*
[3]*Software Product Engineering, Lilly Spain, Alcobendas, Madrid, Spain*

Drug design is a complex and challenging data-driven process. Its success and efficiency largely depend on the speed and quality of methodologies applied to analyze experimental results to inform future decisions in projects. In this context, chemical space (an abstract space with compounds as its core components) proves invaluable for analyzing structural and property information associated with small molecules already explored in projects or available for future explorations.

This is an outline of Eli Lilly's applications of Generative Topographic Mapping (GTM)[1] in various drug discovery tasks associated with chemical space visualizations and analysis. GTM's polyfunctionality[2] extends beyond mere visualization of the similarity relationships in the chemical space by combining chemical space maps with physicochemical and biological property analysis, structure-activity relationships (SAR) visualization, compound collection comparisons, library designs and analogues search.

The diverse functionality of GTM, coupled with frequently used descriptive statistics such as property distribution histograms and Tanimoto distance analysis, offer valuable insights to inform medicinal chemists' decisions across multiple aspects of early stages of drug discovery. This includes: i) assessing project trajectories by visualizing and evaluating chemical space exploration in projects; ii) strategizing future screening approaches through the analysis and comparison of available and previously explored chemical matter; iii) guiding library design as per desired chemical space specifications, and iv) enhancing internal collections by selecting libraries for purchase and cherry-picking compounds from commercially available sources or virtual collections.

Bibliography:

[1] - Bishop, C. M.; Svensén, M.; Williams, C. K. I. GTM: The Generative Topographic Mapping. *Neural Comput.* **1998**, *10* (1), 215– 234, DOI: 10.1162/089976698300017953.
[2] - Horvath D., Marcou G., Varnek A.,Generative topographic mapping in drug design. *Drug Discovery Today: Technologies.* **2019,** 32-33, 99-107, DOI: 10.1016/j.ddtec.2020.06.003.