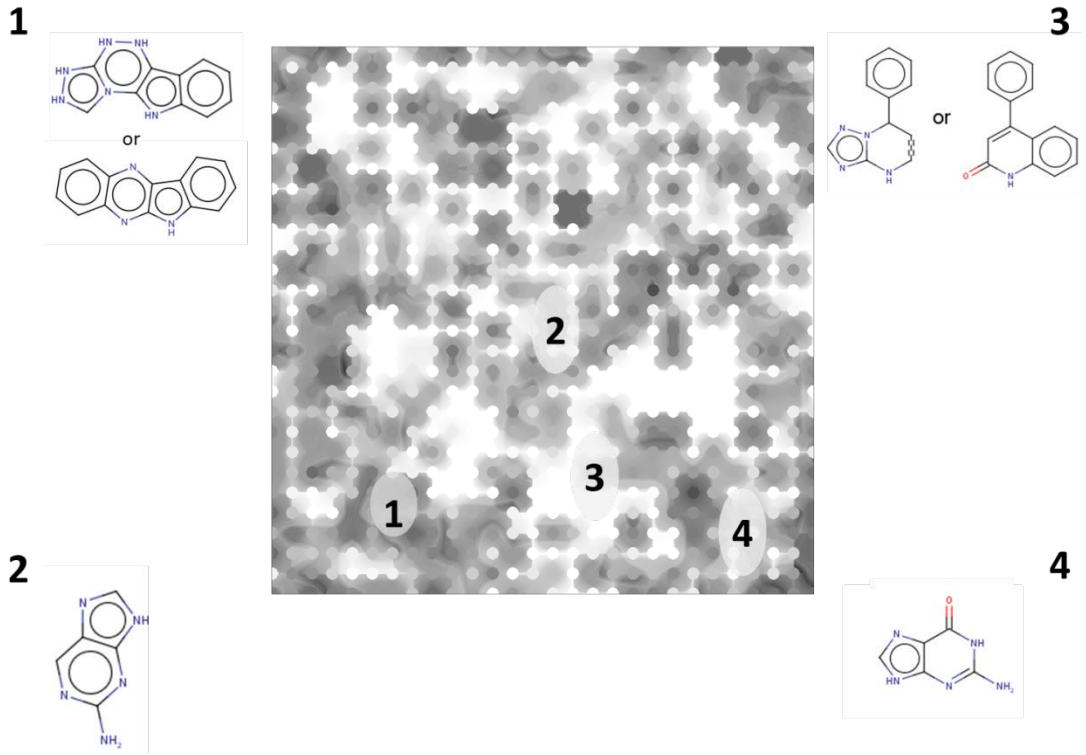
[P32] Chemical space of antiviral compounds: an analysis of the structure-activity information for compounds with antiviral activity from the ChEMBL database.

Kyrylo Klimenko [1,2], Gilles Marcou[1], Dragos Horvath[1], Alexandre Varnek[1]

[1] *Laboratoire de Chemoinformatique, UMR 7140 CNRS/Université de Strasbourg, 1, rue Blaise Pascal, Strasbourg 67000, France*
[2] *A.V. Bogatsky Physico-Chemical Institute of NAS of Ukraine, Lyustdorfskaya doroga, 86, Odessa 65080, Ukraine*

Viral diseases present a real threat to humanity, whereas search for antiviral agents is one of the most challenging fields in drug discovery. Recently significant research efforts provided accumulation of relevant antiviral activity data. With such "big data", it is increasingly more challenging to analyze and use steadily growing experimental information. Even though databases of commercially available antiviral compounds exist, there is still considerable free access structure-activity data, suitable for usage.
Curation, standardization and data fusion of the antiviral data from the ChEMBL database led to a robust data set creation, providing a classification of antiviral compounds to seven broadly defined groups. Generative Topographic Mapping (GTM) [1] enhanced by genetic algorithm, was then used to build maps of the antiviral chemical space, ensuring an optimal separation of compound families associated with the different antiviral classes.



Example of common structural motifs position on the antiviral map:
The ability to define the specific area (responsibility patterns) occupied on a map by various classes of antiviral compounds allowed a GTM-supported search for privileged structural motifs, typical for each antiviral class. The privileged responsibility patterns of class-specific antiviral compounds were analyzed in order to highlight privileged common structural motifs. Responsibility patterns were found to represent underlying c various structural motifs – from very fuzzy (groups of various "interchangeable" similar scaffolds), to the classical scenario in medicinal chemistry (underlying motif actually being the scaffold), to very precisely defined motifs (specifically substituted scaffolds).