

Anti-bacterial / Anti-viral Set

Anti-bacterial/anti-viral set

Molecules in this set are based on the literature analysis of both targets and their respective ligands, specifically 5,8K known molecules with reported anti-bacterial and anti-viral activities. In order to assemble the set, we used a) Kohonen self-organizing map, b) pharmacophore analysis, c) bioisosteric replacement and d) internal peptidomimetics in order to select initial library of ca. 50,000 drug-like molecule. These were further prioritized based on relevance to the target/ligand of interest, IP potential, Lipinski-rule-of-five parameters and synthetic feasibility to yield final set of 10,000 molecules. Approximately 10% of these compounds were proven to possess relevant biological activity in our internal screening and via collaborative efforts.

Example: application of Kohonen self-organizing maps to selection of compounds with potential anti-bacterial activity.

Structures in the training set: 5806 antibacterial agents from the Prous Integrity database

Structures in the test set: 10100 cmpds from the focused library

Approach: analysis of multidimensional property space of antibacterial compounds and focused library using Kohonen self-organizing maps (for description of the method, see [K. V. Balakin, Y. A. Ivanenkov, N. P. Savchuk, A. A. Ivashchenko, S. Ekins. Comprehensive Computational Assessment of ADME Properties Using Mapping Techniques // Curr. Drug Disc. Techn. 2005. V. 2. 99-113.])

Computational tool: SmartMining program, ver. 1.01

Description of experiment

1. A total of 90 molecular descriptors were calculated for both training and test sets.
2. 15 descriptors (table 1) were selected using Principal Component Analysis (PCA) and used for generation of Kohonen map.
3. A Kohonen map of the training dataset (5806 antibacterial agents) was generated using 15 descriptors (Fig. 1); this map can be used for assessment of "antibacterial-likeness" of chemical compounds.

Anti-bacterial / Anti-viral Set

4. Compounds from the focused library (test set of 10113 cmpds) were placed on the same map (Fig. 2).

Table 1. 15 Molecular descriptors selected by PCA and used for generation of the Kohonen map.

NN	Descriptor	Definition
1	M ₁	Zagreb index (Gutman index)
2	SS	Electrotopological state index
3	SS(=N-)	E-state index of (=N-) fragment
4	SS(-NH ₂)	E-state index of (-NH ₂) fragment
5	SS(=NH)	E-state index of (=NH) fragment
6	SS(#N)	E-state index of (#N) fragment
7	SS(-OH)	E-state index of (-OH) fragment
8	SS(=O)	E-state index of (=O) fragment
9	RBN	Number of rotatable bonds
10	H _b _Don	Number of hydrogen bond donors
11	H _b _Acc	Number of hydrogen bond acceptors
12	Hy	Hydrophilicity index
13	HB ₂	Hydrogen bonding ability of functional groups
14	VDW(O)	Van der Waals volume of oxygen atoms
15	VDW(N)	Van der Waals volume of nitrogen atoms

Anti-bacterial / Anti-viral Set

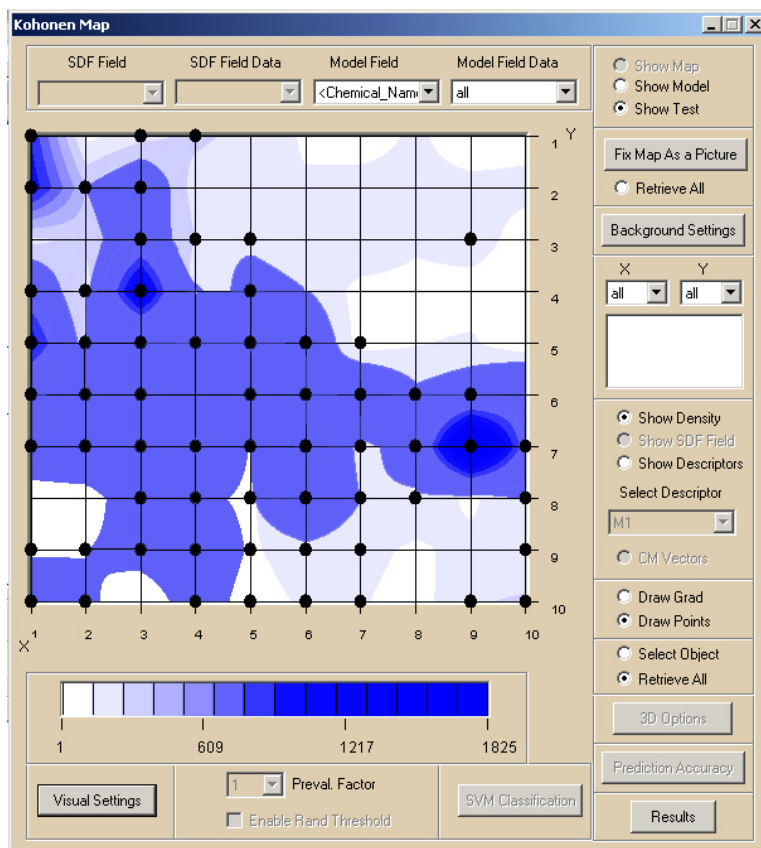


Fig. 1. 2D Kohonen map (10×10); blue area – compounds from the training set, black circles – compounds from the test set.

Anti-bacterial / Anti-viral Set

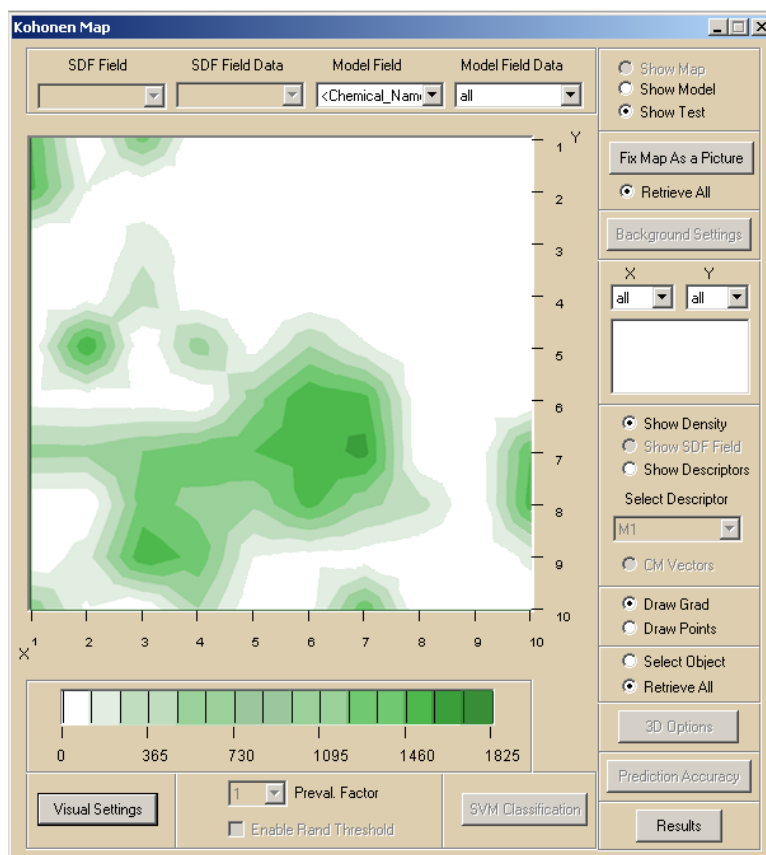


Fig. 2. Kohonen map; green area – localization of compounds from the test set.

Conclusion

Compounds from the focused library (green area, Fig. 2) are located in regions of the map which contain antibacterial agents (blue area, Fig. 1). Therefore, the molecular properties of compounds from the focused library are similar to properties of the antibacterial agents.