

TopoHERG – A highly selective Pharmacophoric Classifier for hERG-channel active Compounds

B. Nisius¹

A. H. Göller²

¹ *B-IT Life Science Informatics, University of Bonn, Germany, and Bayer Healthcare AG, Bayer Schering Pharma Global Drug Discovery, Wuppertal, Germany.*

² *Bayer Healthcare AG, Bayer Schering Pharma, Global Drug Discovery, Wuppertal, Germany*

Cardiac arrhythmia as a side-effect of many drugs has become a major pharmacological safety concern and led to the withdrawal of many drugs through recent years. Virtually every case of prolonged duration of cardiac action potential related to drug exposure (long QT syndrome) can be traced to one specific mechanism: blockade of I_{kr} current in the heart, conducted by an ion channel encoded by the human ether-a-gogo-related gene (hERG).

Experimental and modelling results provide clear evidence that ligand binding is rather unspecific and in-homogenous, and there will exist multiple binding SAR's. It is therefore wise to combine multiple models based on different structure and property descriptors.

Our highly specific TopoHERG approach uses the Tripos Topomer Search [1] to compare any query molecule via topomer similarity to a database of known actives and inactives [2] using the fact that similar compounds tend to show similar biological effects. The classifier with strict topological distance classifies about 54 % of the compounds with a specificity of 95 % and a sensitivity of 76 % on the classified compounds.

The unclassified compounds were then run through a decision tree-based classifier based on clogP , CMR and the availability of a ionizable nitrogen we adapted from literature [3]. Since the classifier was created on 490 Johnson&Johnson in-house compounds, our data are a real validation set. The outcome of the combined TopoHERG and decision-tree classifiers on the complete dataset is a sensitivity of 54 %, a specificity of 94 % and overall 87 % of the compounds predicted correctly.

TopoHERG, has a high predictivity for any compound it is applicable to due to the fact that regardless of the binding mode of any chemotype the model will classify by inherent pharmacophoric similarity. By design, the model will improve with any new compound added to the database which can grow on a daily basis without re-training. Additionally, no split into training and test set is necessary, broadening the domain of applicability of the method. Combined with one or multiple orthogonal classifiers the overall performance of the model can fulfill the needs of pharmaceutical industry.

1. Cramer, R. D.; Jilek, R. J.; Andrews, K. M. dbtop: Topomer similarity searching of conventional structure databases, *J. Molec. Graph. Mod.*, **2002**, 20, 447-462.
2. Database: The dataset consists of 475 compounds. 232 compounds were collected from publicly available hERG blockade IC_{50} data, the other compounds are in-house patch clamp measurements on HEK and CHO cells. 276 compounds in the dataset don't exhibit hERG blockade ($\text{pIC}_{50} < 5$), 118 compounds show medium hERG activity ($5 \leq \text{pIC}_{50} < 6$) and 82 compounds are highly hERG active ($\text{pIC}_{50} \geq 6$).
3. (a) Buyck, C.; Tollenaere, J.; Engels, M.; Clerck, F. D.; An in silico model for detecting potential hERG blocking; *Poster presentation, Euro-QSAR 2002, Bournemouth, 2002*; (b) Aronov, A. M.; Goldman, B. B.; A model for identifying HERG K⁺ channel blockers; *Bioorg. & Med. Chem.* **2004**, 12, 2307-2315.