binding sites. First we extract several most commonly used characteristics of the amino acid residues', and the by using feature selection techniques we reduce the dataset by taking into consideration the most relevant characteristics in order to provide faster decisions in the next stage. By applying feature selection, the accuracy of the models, as well as time complexity and model complexity are improved. After reducing the number of characteristics, next we build prediction models by using various classification method based on the classical and fuzzy set theory. The results confirm that by applying feature selection better models are obtained. Our approach is also compared with several well-known methods for determination of protein binding sites.

References: G. Mirceva, A. Kulakov, Journal of structural biology, **2015**, *189(1)*, 9-19.

P52

APPLICATION OF LINEAR COMBINATION OF PHARMACOPHORE HYPOTHESES INTO SEARCH FOR THE DUAL 5-HT$_{1A}$/SERT LIGANDS

Dawid Warszycki, Rafał Kafel, Andrzej J. Bojarski

Institute of Pharmacology Polish Academy of Sciences, 12 Smetna Street, Kraków, Poland

Pharmacophore modelling is one of the most basic concepts in medicinal chemistry, yet the current approaches face several limitations, such as partial coverage of chemical subspace of ligands of particular target, exploring limited conformational space and resulting in low-feature hypotheses with low selectivity potential. Recently published idea of using linear combination of pharmacophore hypotheses [1], allowed creation of an effective feature mapping protocol neglecting the aforementioned issues.
In this study the linear combination of hypotheses was applied to explore the space of dual 5-HT$_{1A}$/SERT ligands. First, query on the latest version of ChEMBL database [2] resulted in 4427 active ligands for 5-HT1A receptor (with K_i or equivalent equal or less than 100 nM), 3806 actives for SERT and 532 of double activity. This created a strong training base for evaluation whether separate combinations for each target or one combination developed only for common ligands is more efficient in the search of dual ligands. All of the ensembles of models were generated utilizing previously described methodology [1] to reach maximum value of MCC (Mathews Correlation Coefficient) for the test set. Both approaches were tested on validation set of all unused dual ligands and randomly selected ZINC compounds (with typical features for serotonin receptors ligands – aromatic system and polarizable nitrogen atom) Results indicate that a combination for dual ligands are slightly more robust in active/inactive discrimination and less time-consuming, however separate combinations for each target present a valid alternative in case of limited number of dual ligands available for training.

References: 1) Warszycki D et al., PLoS ONE, **2013**, 8(12), e84510; 2) Bento A P et. al., Nucleic Acids Research, **2014**, 42, 1083-1090.

Acknowledgments: The study was partially supported by the Polish-Norwegian Research Programme operated by the National Centre for Research and Development under the Norwegian Financial Mechanism 2009–2014 in the frame of Project PLATFORMex (Pol-Nor/198887/73/2013).