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

e-mail: warszyc@if-pan.krakow.pl

Introduction

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, exploration of limited conformational space and, as a result, 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 solving the aforementioned problems.

Methodology

In this study, the linear combination of hypotheses was applied to explore the space of dual 5-HT$_{1A}$/SERT ligands (Figure 1). Query on the latest version of ChEMBL database [2] resulted in 4427 active ligands for 5-HT$_{1A}$ receptor (with K_i or equivalent equal or less than 100 nM), 3806 actives for SERT and 532 compounds with dual activity. Compounds with confirmed inactivity (with K_i or equivalent more than 1000 nM - decoys) were also fetched from the ChEMBL database (1230 for 5-HT$_{1A}$, 1726 for SERT and 200 compounds with confirmed inactivity for at least one of beforementioned targets). MOLPRINT2D fingerprints were generated for all sets of actives and used as input for Hierarchical Clustering Tool implemented in Canvas software [3]. Singletons and small clusters (less than 4 compounds) were rejected. Clustering resulted in 33 clusters for 5-HT$_{1A}$ receptor ligands , 32 clusters for SERT ligands and 27 subgroups for dual compounds.

For representative compounds selected from each cluster, separate pharmacophore hypotheses were generated and tested using Phase software [4]. All of them mapped at least half of the ligands used for their development. The best hypothesis for each cluster was selected on the basis of following criteria: maximal number of features, the highest number of matched representative compounds, and the highest value of selectivity score. Not all clusters returned pharmacophore hypothesis (27 hypotheses out of 33 clusters for 5-HT$_{1A}$R, 21 out of 32 for SERT and 20 out of 24 for dual ligands) Pharmacophore models were tested on three different test sets (actives, decoys and assumed inactives (fetched from DrugBank), each containing 200 compounds), and characterized by MCC coefficient, which was an average from MCC coefficients obtained for combinations of actives/decoys and actives/assumed inactives. Next, the MCC coefficients for all possible linear combinations of hypotheses were calculated by an in-house script until the maximum MCC values were reached (Figure 2). Both approaches were tested on validation set of all unused dual ligands (234 compounds) and randomly selected ZINC compounds (2394 - ratio actives/inactives 1:9; with typical features for serotonin receptors ligands - aromatic system and polarizable nitrogen atom).

Conclusions

Results (Figure 1) indicate that a combination for dual ligands is slightly more robust in active/inactive discrimination and less time-consuming (19 maping procedures for combination of dual ligands vs 22 in total for individual targets), however, separate combinations for each target present a valid alternative in case of limited number of dual ligands available for training (e.g. only 7 dual ligands for 5-HT$_6$/SERT and 24 structures for 5-HT$_7$/SERT). Hypotheses for dual ligands are statistically more complex than hypotheses developed for individual targets (averagely 5.75 features per cluster vs 4.84) which influences on combination's selectivity.

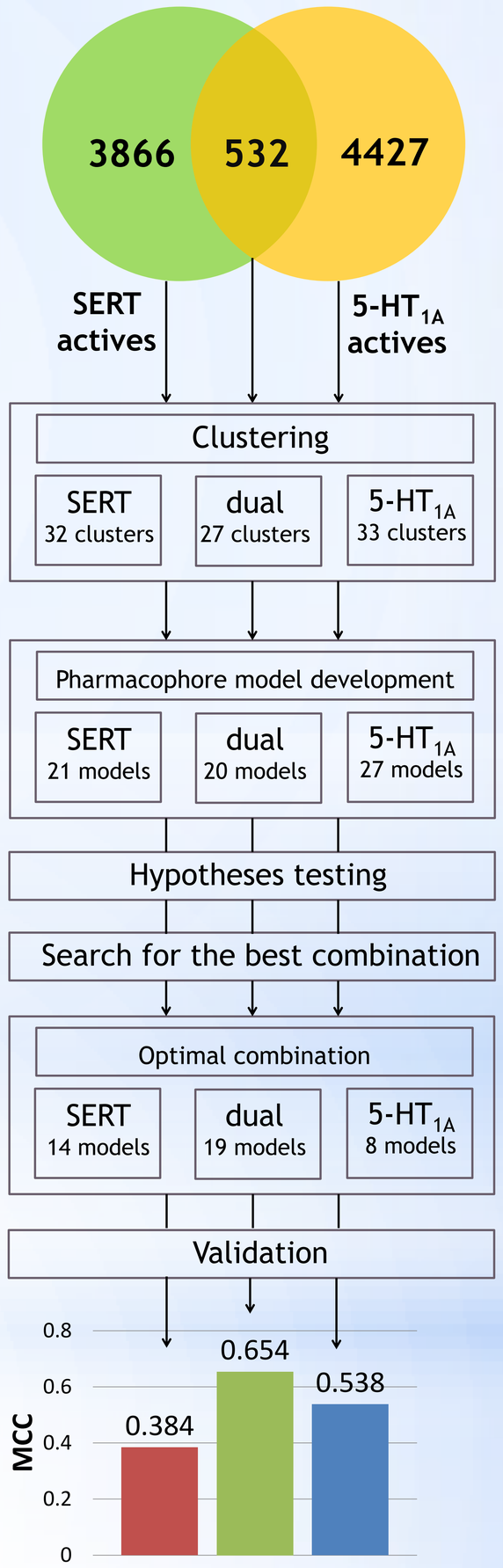

Figure 1. The development of an optimal combinations of pharmacophore models.

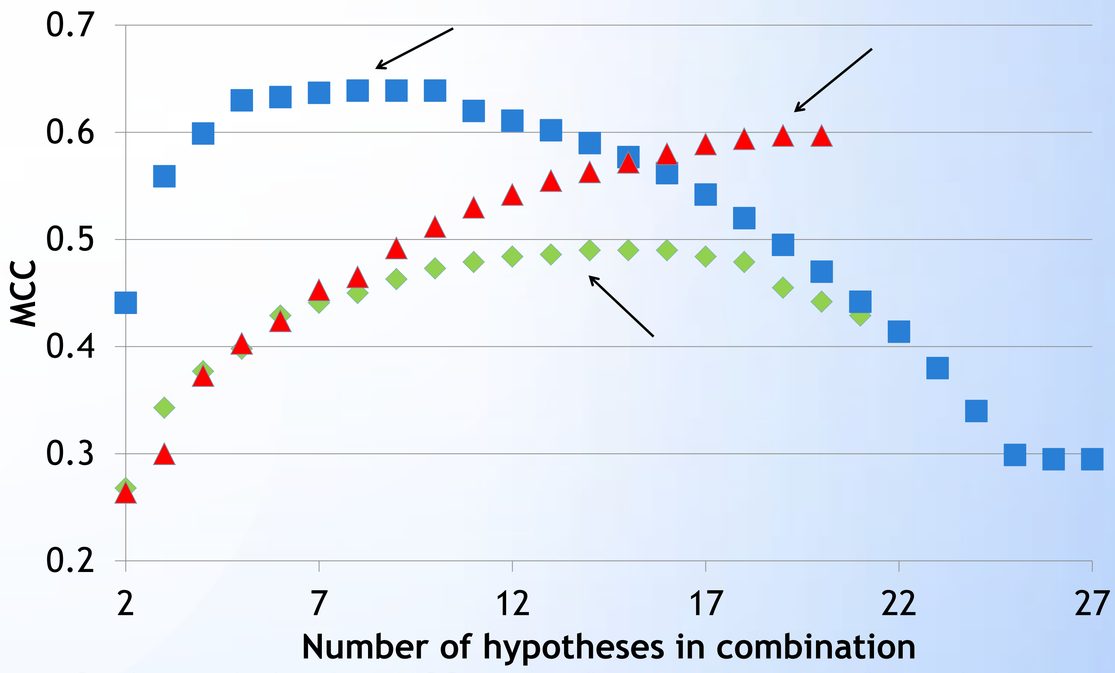

Figure 2. An optimization MCC curve for the ligands of the investigated targets and dual ligands. Arrows indicate the maximum values of MCC reached a rate of 0.490 for SERT for 14 hypotheses, of 0.639 for 5-HT$_{1A}$ for 8 hypotheses and of 0.597 for 19 hypotheses for dual ligands.

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.
[3] Canvas, version 2.5, Schrödinger, LLC, New York, NY, 2015.
[4] Phase, version 4.4, Schrödinger, LLC, New York, NY, 2015.
[5]Huang, M. et al., J. Med. Chem., 2006, 49(23), 6789-6801

Acknowledgements

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).