Because two is always better than one – towards the search of dual 5-HT$_x$-SERT ligands

Sabina Podlewska,[1,2] Dawid Warszycki,[1] Andrzej J. Bojarski[1]

[1]Department of Medicinal Chemistry, Institute of Pharmacology Polish Academy of Sciences, 12 Smętna Street, 31-343 Kraków
[2]Faculty of Chemistry, Jagiellonian University, 3 Ingardena Street, 30-060 Kraków

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

Serotonin (5-HT) receptors are representatives of class A GPCRs (except for the 5-HT$_3$R, which is a ligand-gated ion channel) and they play important role in the majority of physiological and behavioral processes, such as sleep, appetite, memory and mood control, cardiovascular functions, digestion and muscle contraction.[1] The serotonergic signaling is regulated by the Na^+/Cl^--dependent serotonin transporter (SERT) that pumps the serotonin back into the neurons after completion of the signaling task for its re-utilization. As abnormalities in the serotonergic system function were proved to be responsible for various mental disorders, such as depression, schizophrenia, anxiety and bipolar disorder, we decided to develop a virtual screening protocol for the search of new ligands with dual activity towards selected serotonin receptor subtypes (5-HT$_{1A}$, 5-HT$_6$, 5-HT$_7$) and SERT.

STAGE I

SERT screening

2 660 986 cpds

ENAMINE DATABASE

STAGE II

5-HTx screening

5-HT$_{1A}$, 5-HT$_6$, 5-HT$_7$

Table 1. The number of active and inactive compounds present in particular training set.

TRAINING SET COMPOSITION		
	actives	**inactives**
5-HT$_{1A}$	4167	1156
5-HT$_6$	1519	380
5-HT$_7$	721	371
SERT	3693	1597

ChEMBL

- actives, K_i < 100nM
- inactives, K_i > 1000 nM

ExtFP: 992 954 cpds
KlekFP: 743 804 cpds
MACCSFP: 756 837 cpds

ML EXPERIMENTS

	SERT/ 5-HT$_{1A}$	SERT/ 5-HT$_6$	SERT/ 5-HT$_7$
ExtFP	58 007 cpds	37 972 cpds	39 552 cpds
KlekFP	28 850 cpds	45 143 cpds	33 678 cpds
MACCSFP	87 099 cpds	49 916 cpds	55 0558 cpds

The CheMBL database[2] was used as a data source for training set composition. Due to the limited data available on compounds with dual activity, the retrospective screening strategy was applied to determine the threshold for the active compounds retrieval (the following values in terms of K_i parameter were tested: 100 nM, 200 nM, 500 nM, and 1000 nM). The 10-fold cross-validation studies indicated, that the most effective identification of active ligands is obtained when the K_i of 100 nM threshold for active (less than), and 1000 nM (higher than) for inactive compounds is applied, and the respective training sets were prepared accordingly on the basis of the ChEMBL db data for each target separately (the number of compounds from particular class is presented in Table 1 and the number of dual 5-HT$_x$/SERT ligands is presented in Figure 1). The virtual screening procedure was applied to the Enamine database.

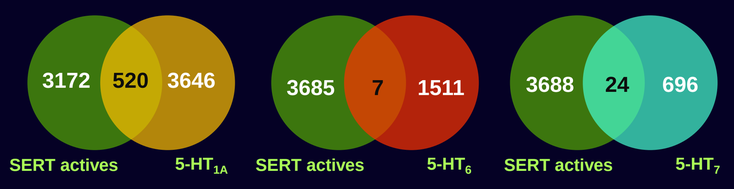


Figure 1. The number of dual 5-HT$_x$/SERT ligands

147 899 cpds

FINGERPRINT-BASED CONSENSUS

SERT/ 5-HT$_{1A}$	SERT/ 5-HT$_6$	SERT/ 5-HT$_7$
9556 cpds	7916 cpds	5361 cpds

All the compounds were represented by three fingerprints from the PaDEL-Descriptor package[3]: Extended Fingerprint (ExtFP), Klekota&Roth Fingerprint (KlekFP) and MACCS Fingerprint (MACCSFP), capturing different pieces of information about the structure and properties of the analyzed molecules. The Enamine compounds were evaluated by the modified version of the Support Vector Machine (SVM) algorithm: Sequential Minimal Optimization (SMO),[4] the WEKA package implementation was used.
To the next step of the screening protocol, only those compounds passed, for which the potential activity was indicated for all three representation of molecules.[5] The high number of compounds rejected in this step revealed that the set of molecules selected by individual fingerprints were rather disjoint.

WEKA
The University of Waikato

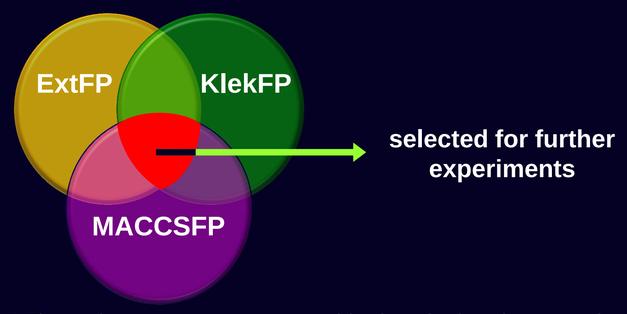


Figure 2. Fingerprint-consensus strategy used for the selection of compounds for further experiments.

SIMILARITY ANALYSIS

SERT/ 5-HT$_{1A}$	SERT/ 5-HT$_6$	SERT/ 5-HT$_7$
8922 cpds	7752 cpds	5124 cpds

The Tanimoto coefficients towards known 5-HT$_x$ and SERT binders were calculated and all structures with the similarity coefficient values higher than 0.7 were rejected from further consideration. .

Table 2. Templates used for homology modeling purposes.

	5-HT$_x$	SERT
TEMPLATES	5-HT$_{1B}$	Leucine transporter
	5-HT$_{2B}$	
	A_{2A}	
	beta1	
	beta2	
	CXCR4	Dopamine transporter
	D_3	
	H_1	
	M_2	
	M_3	

DOCKING

SERT/ 5-HT$_{1A}$	SERT/ 5-HT$_6$	SERT/ 5-HT$_7$
5679 cpds	5818 cpds	2287 cpds

Homology models of 5-HT$_{1A}$, 5-HT$_6$, 5-HT$_7$ and SERT were constructed on all available templates (Table 2). For each template, 100 models were generated with the use of Modeller v9.13 (sequence alignments were performed manually). They were evaluated on the basis of 3-step retrospective screening strategy. On the basis of AUROC values, the best models of 5-HT$_x$ for each template were identified and 5 models with the highest actives/inactives discrimination power were selected for further studies (Table 3). In the case of SERT, the models constructed on the dopamine transporter template were taken for further study.

Table 3. The AUROC values obtained in the retrospective screening strategy applied for 5-HT$_x$ homology models evaluation. Green background indicate models selected for further studies.

Target/ Template	5-HT$_{1A}$	5-HT$_6$	5-HT$_7$
5-HT$_{1B}$	-*	0.499	0.441
5-HT$_{2B}$	-	-	0.478
A_{2A}	0.573	0.693	0.709
beta1AR	0.392	-	0.787
beta2AR	0.576	0.729	0.757
CXCR4	0.653	0.718	0.669
D_3	0.629	0.689	0.764
H_1	0.641	0.605	0.828
M_2	0.406	0.589	0.717
M_3	0.591	0.643	0.749

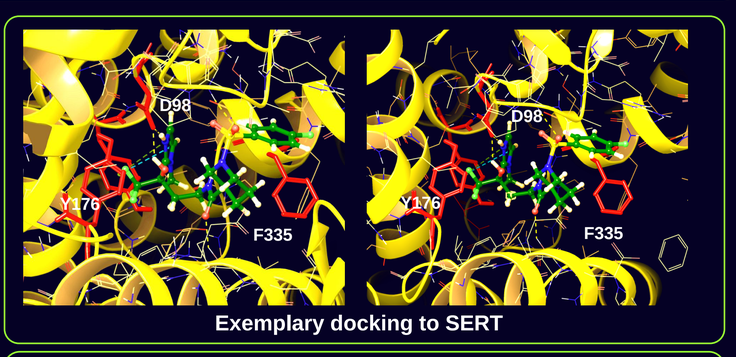
'-' means that the model did not pass to the last evaluation stage

PURCHASING COMPOUNDS

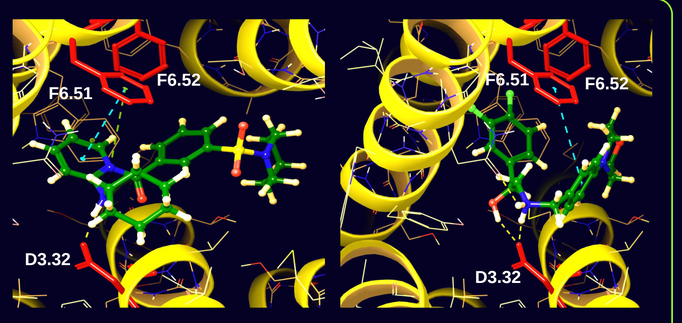
100 cpds

The compounds purchasing was supported by the docking results. To make them independent from the template used for modeling, the docking to 5 best performing models indicated in Table 3 were taken into account. For each compound, the consensus docking score was calculated, taking into account the docking score function and the quality of the homology model constructed. It was expressed as the weighted average of absolute values of the docking score with weights being the AUROC obtained during the homology models evaluation step.
At the end, 100 compounds were selected for purchasing and their activity towards all considered serotonin receptors and SERT will be examined in *in vitro* experiments.

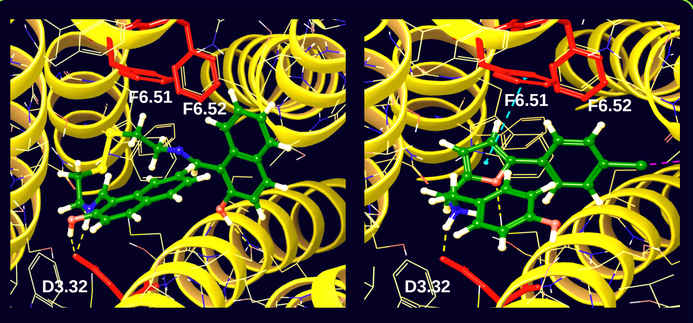
IN VITRO TESTS



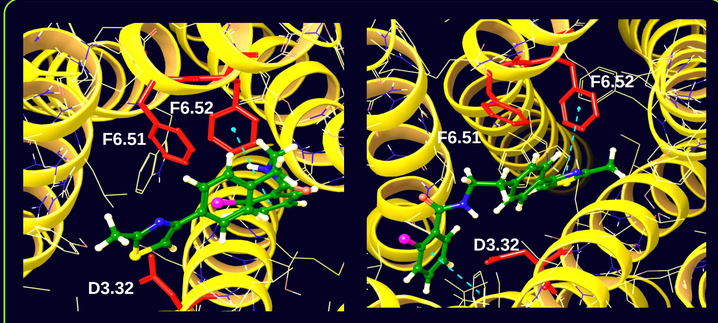
Exemplary docking to SERT



Exemplary docking to 5-HT$_{1A}$R



Exemplary docking to 5-HT$_6$R



Exemplary docking to 5-HT$_7$R

References

[1] Shajib, M. S.; Khan, W. I. The role of serotonin and its receptors in activation of immune response and inflammation, *Acta Physiologica*, **2015**, *213*, 561-574.
[2] Bento et al. The ChEMBL bioactivity database: an update, *Nucleic Acids Research,* **2014**, *42*, 1083-1090.
[3] Yap, C. W. PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints, *Journal of Computational Chemistry,* **2011**, *32*, 1466-1474.
[4] Platt, J. C. Advances in kernel methods: support vector learning, *Cambridge, MA, MIT Press*, **1999**, 185-208.
[5] Smusz, S.; Kurczab, R.; Satała, G.; Bojarski, A. J. Fingerprint-based consensus virtual screening towards structurally new 5-HT6R ligands, *Bioorganic and Medicinal Chemistry Letters,* **2015**, *25*, 1827-1830.

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 the Project PLATFORMex (Pol-Nor/198887/73/2013).