

# Fragment-based approaches in GPCR-targeted drug discovery

Thesis

Ádám Andor Kelemen

PhD candidate

György Miklós Keserű DSc, FRSC

supervisor



Medicinal Chemistry Research Group  
Research Centre for Natural Sciences  
Hungarian Academy of Sciences

Budapest

2018



M Ű E G Y E T E M 1 7 8 2

## Table of contents

Table of contents.....	2
List of abbreviations.....	7
Acknowledgements.....	10
Chapter 1 – Introduction .....	11
1.1. Fragment-based drug discovery .....	11
1.1.1. Concept of Fragment-based drug discovery .....	11
1.1.2. Strategies and examples .....	12
1.1.2.1. Fragment-based drug design workflow.....	12
1.1.2.2. Fragment optimization strategies.....	13
1.2. G-Protein Coupled Receptors .....	14
1.2.1. Signalization and classification .....	14
1.2.2. X-Ray crystal structures of aminergic GPCR's.....	19
1.3. Medicinal Chemistry of Serotonin (5-HT) receptor ligands.....	21
1.3.1. Serotonin Receptor Subtype 6 (5-HT <sub>6</sub> R) <sup>64</sup> .....	21
1.3.1.1. Pharmacological relevance.....	21
1.3.1.2. Known 5-HT <sub>6</sub> R ligands and pharmacophore models .....	21
1.3.2. Serotonin Receptor Subtype 7 (5-HT <sub>7</sub> R) .....	23
1.3.2.1. Pharmacological relevance.....	23
1.3.2.2. Known 5-HT <sub>7</sub> R ligands and pharmacophore models .....	23
1.3.3. Serotonin Receptor Subtype 2B (5-HT <sub>2B</sub> R) .....	24
1.3.3.1. Pharmacological relevance.....	24
1.3.3.2. Known 5-HT <sub>2B</sub> R ligands.....	24

1.4. Synthetic approaches leading to spiro[pyrrolidine-3,3'-oxindoles]-analogues .....	25
Chapter 2 – Results and Discussion .....	29
2.1. Physicochemical-property based scoring method for the design of aminergic fragment libraries <sup>143</sup> .....	29
2.1.1. Development of the scoring scheme .....	29
2.1.2. Retrospective validation and evaluation.....	35
2.1.3. Prospective validation .....	37
2.1.4. Defining the FrAGS cut-off .....	38
2.2. Consensus-scoring based docking method for the design of aminergic fragment libraries <sup>154</sup> .....	40
2.2.1. Consensus ranking .....	41
2.2.1.1. Selection of structures for docking using clustering and self-docking.....	41
2.2.1.2. Optimization of the protein ensemble .....	42
2.2.2. Retrospective validation of FrACS.....	44
2.2.3. Combination of FrAGS with FrACS .....	45
2.2.4. Binding mode analysis of active fragments in 5-HT <sub>6</sub> R homology models .....	48
2.3. Spiro-oxindoles, and -indolines as new 5-HT <sub>6</sub> R antagonists <sup>172</sup> .....	50
2.3.1. Early SAR data on spiro[pyrrolidine-3,3'-oxindole] derivatives .....	50
2.3.2. Hit-to-lead optimization of spiro[pyrrolidine-3,3'-oxindoles] .....	53
2.3.3. Binding mode analysis of the Optimized 5-HT <sub>6</sub> R Ligand (74a) .....	61
2.3.4. Conclusion.....	62
2.4. Spiro-oxindoles as new 5-HT <sub>7</sub> R antagonist .....	64

2.4.1. Identifying novel 5-HT <sub>7</sub> R ligands with spiro-oxindole scaffold .....	64
2.4.2. Optimization of spiro[indoline-3,3'-pyrrolidin]-2-ones as 5-HT <sub>7</sub> R ligands .....	66
2.5. Consensus-scoring based docking for the design of 5-HT <sub>2B</sub> R selective antagonists <sup>198</sup> .....	71
2.5.1. Concept of SCFP-based machine-learning method and combination with an SBP-targeted consensus ranking approach ....	71
2.5.2. Selection of virtual screening hits for GPCR-panel screening.	74
2.5.3. Binding Mode Analysis of the selective 5-HT <sub>2B</sub> R Ligand .....	77
Chapter 3 – Computational Methods and Experimental Section .....	80
3.1. Computational Methods and Procedures .....	80
3.1.1. Computational procedures for physicochemical-property based scoring method (FrAGS) .....	80
3.1.1.1. Compilation of training set .....	80
3.1.1.2. Compilation of ChEMBL retrospective validation set .....	82
3.1.1.3. Compilation of PubChem retrospective validation set .....	82
3.1.2. Computational procedures for the consensus-scoring based docking method (FrACS).....	83
3.1.2.1. Collection of protein structures .....	83
3.1.2.2. Selection of structures for the docking ensemble.....	83
3.1.2.3. Collection of the training set for the optimization of the consensus scoring .....	86
3.1.2.4. In-house fragment library .....	87
3.1.2.5. Molecular docking procedures for the consensus-scoring method .....	87
3.1.2.6. Procedures of PubChem validation of the consensus scoring method .....	87

3.1.2.7. Procedures of consecutive screening by combining FrAGS with FrACS.....	87
3.1.2.8. Docking of fragment hits to the 5-HT <sub>6</sub> R homology model...	88
3.1.3. Computational procedures for the study of spiro-oxindoles and indolines as 5-HT <sub>6</sub> R antagonists .....	88
3.1.4. Computational procedures for the consensus-scoring based docking method targeting 5-HT <sub>2B</sub> R.....	88
3.1.4.1. Compilation of the training set, NSFP-based machine learning model building .....	88
3.1.4.2. Building NSFP-based machine learning model and screening of the MCule database .....	89
3.1.4.3. Ligand and protein preparations for 5-HT <sub>2B</sub> R dockings.....	91
3.2. Biological assay descriptions .....	92
3.2.1. Cell-Based assays for 5-HT <sub>6</sub> R and CB <sub>1</sub> R.....	92
3.2.2. Radioligand binding assays for 5-HT <sub>1A</sub> R, 5-HT <sub>2A</sub> R, 5-HT <sub>6</sub> R, 5-HT <sub>7</sub> R .....	93
3.2.2.1. Cell culture and preparation of cell membranes.....	93
3.2.2.2. Radioligand binding assays .....	93
3.2.3. Radioligand binding assays for 5-HT <sub>1B</sub> R and 5-HT <sub>2B</sub> R.....	94
3.2.3.1. Competition binding in human 5-HT <sub>1B</sub> receptor.....	94
3.2.3.2. Competition binding in human 5-HT <sub>2B</sub> receptor.....	94
3.3. Synthetic procedures.....	95
3.3.1. Synthesis of spiro-oxindoles and indolines as new 5-HT <sub>6</sub> R antagonists.....	95
3.3.2. Synthesis of spiro-oxindoles as new 5-HT <sub>7</sub> R antagonists.....	97
Chapter 4 – Conclusions.....	105
4.1. Conclusions and thesis points .....	105

4.2. Összefoglalás és tézispontok.....	107
References.....	110
APPENDIX.....	125
NYILATKOZAT.....	137

## List of abbreviations

5-HT<sub>x</sub>R – serotonin receptor subtype “x”  
AD – Alzheimer’s Disease  
ADHD - Attention Deficit Hyperactivity Disorder  
ADME – absorption, distribution, metabolism, elimination  
AIBN - azobisisobutyronitrile  
ChEMBL – database cured by the European Bioinformatics Institute (EMBL-EBI)  
CNS – central nervous system  
CTX – cholera toxin  
DAG - diacylglycerol  
DIPEA - diisopropylethylamine  
DMAP – 4-(dimethylamino)pyridine  
DRY-mofif – motif of AspArgTyr conserved amino acids  
ECL – extracellular loop  
EF – enrichment factor  
FBDD – Fragment-based Drug Discovery  
FNR – False Negative Rate  
FPR – False Positive Rate  
FrACS – Fragment Aminergic Consensus Score  
FrAGS – Fragment Aminergic GPCR Score  
FS – fragment screening  
GABA -  $\gamma$ -Aminobutyric acid  
GDP - guanosine diphosphate  
GERD - gastroesophageal reflux disease  
GPCR – G-protein coupled receptor  
GTP- guanosine-5'-triphosphate  
HATU - 1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxide hexafluorophosphate  
HBA – hydrogen bond acceptor  
HBD – hydrogen bond donor  
HCS – high content screening  
HRMS – High-Resolution Mass Spectrometry  
HTS – high-throughput screening  
HYD – hydrophilic pharmacophore feature  
IBS – irritable bowel syndrome  
IC<sub>50</sub> – half-maximal inhibitory concentration

ICL – intracellular loop  
IP3 – triphosphoinositol  
ITC - Isothermal titration calorimetry  
JNK – Janus kinase  
logP – water-octanol partition coefficient  
MAPK ERK - mitogen-activated protein kinase  
MeCN - acetonitrile  
MsCl – mesyl chloride  
NBS – *N*-bromosuccinimide  
NCS – *N*-chlorosuccinimide  
nN – number of nitrogen atoms  
nOx – number of oxygen atoms  
NPxxY – motif of AsnPro(xx)Tyr conserved amino acids  
nRot – number of rotatable bonds  
OBP – orthosteric binding pocket  
PAINS – pan-assay interference compounds  
PAM – positive allosteric modulator  
PD – Parkinson’s Disease  
PDB – Protein Data Bank  
PhSO<sub>2</sub>Cl – benzenesulfonyl chloride  
PI – positive ionisable pharmacophore feature  
PIF – motif of ProIlePhe conserved amino acids  
PSA – polar surface area  
PTX – pertussis toxin  
RMSD – root-mean-square deviation  
RO3 – Rule of Three  
ROC – receiver operating characteristic  
RT – room temperature  
SAR – structure-activity relationship  
SBP – secondary binding pocket  
SCFP – Structure Connectivity Fingerprint  
SILE – Size-Independent Ligand Efficiency  
SMILES – simplified molecular-input line-entry system  
SP – Glide Single Precision docking method  
SPR – Surface Plasmon Resonance  
SSRI - Selective Serotonin Reuptake Inhibitors

TAAR1 – trace-amine receptor subtype 1  
TEA - triethylamine  
TM – transmembrane helix  
TNR – True Negative Rate  
TPR – True Positive Rate  
TPSA – topological polar surface area  
TSA – Thermal Shift Assay  
VSW – Glide Virtual Screening Workflow  
XRD – X-Ray diffraction

## Acknowledgements

Kiemelt köszönetemet szeretném kifejezni témavezetőmnek, Dr. Keserű György Miklósnak a munkám szakmai irányításáért, tanácsaiért, és a szakmai előmenetelben nyújtott támogatásáért.

Külön köszönettel tartozom Dr. Kovács Péternek az éveken át tartó támogatásáért, és hasznos tanácsaiért.

Köszönöm a Gyógyszerkémiai Kutatócsoport egykori és jelenlegi munkatársainak a mindennapi munkában nyújtott segítségükért és támogatásukért: Egyed Attila, Orgován Zoltán, Dr. Ábrányi-Balogh Péter, Petri László, Szilágyi Bence, Dr. Bajusz Dávid, Dr. Ferenczy György, Dr. Balog József András, Hegedűs Katalin, Pappné Borsos Éva, Csányi Dorottya, Rockov Márta, Dr. Takács Daniella, Dr. Bereczki Helga, Dr. Benéné Visy Júlia, Dr. Németh Tamás, Pándy Gáspár, Németh András, Mihalovits Levente.

Köszönettel tartozom Dr. Kiss Róbertnek a munkámban nyújtott segítségéért és tanácsaiért.

Köszönöm Dr. Gömör Ágnesnek a minták nagy felbontású tömegspektrometriás méréseit.

Köszönettel tartozom Dr. Flachner Beátának, és Dr. Lőrincz Zsoltnak az 5-HT<sub>6</sub> és CB<sub>1</sub> receptorokon végzett *in vitro* mérésekért.

I would like to express my gratitude to Prof. Andrzej Bojarski, Krzysztof Rataj, Stefan Mordalski PhD, Grzegorz Satała PhD at the Polish Academy of Sciences for their cooperation in computational studies, and for the *in vitro* measurements on the panel of GPCR targets.

I would like to express my gratitude to Aaron Keeley and Andrea Scarpino for their help and ideas in my every day work.

Külön köszönettel és hálával tartozom szüleimnek, Kelemen Istvánnak és Kelemenné Pataki Mártának, hogy mindenben támogattak és mellettem voltak.

# Chapter 1 – Introduction

## Aims and scope of the thesis

In my PhD work I aimed the development of fragment-based ligand-, and structure-based virtual screening methodologies for the design of aminergic GPCR ligands. This thesis presents the development and applicability of physicochemical property-based filtering and consensus scoring based docking methods in prospective validation studies. Thus we aimed to identify possible new chemotypes offering the chance to optimize potent ligands for certain pharmacologically relevant serotonin receptors, such as 5-HT<sub>2B</sub>R, 5-HT<sub>6</sub>R and 5-HT<sub>7</sub>R.

## 1.1. Fragment-based drug discovery

### 1.1.1. Concept of Fragment-based drug discovery

Fragment-based drug discovery (FBDD) is an alternative approach to conventional lead discovery methodologies<sup>1</sup>. FBDD approaches start with the screening of small, fragment sized compounds (typically containing not more than 22 heavy atoms). Owing to the smaller size of fragments with respect to lead like compounds used in traditional high throughput screening (HTS) campaigns, small fragment libraries (~10<sup>3</sup> compounds) give better coverage of the corresponding chemical space than do more extended lead-like libraries (~10<sup>5</sup>-10<sup>6</sup> compounds) and they provide good quality hits with good hit rates. These hits are small, polar compounds (defined by the Rule of Three, RO3<sup>2</sup>) possessing only a couple of pharmacophoric features/interacting moieties. Fragment screening is able to explore binding hotspots of the binding cavity and to provide fragment hits with good ligand efficiencies<sup>3</sup> that are appropriate starting points for optimizing them into high affinity and selective leads (e.g. by structure-based design, etc.). Whereas (HTS) hits, although provide highly potent binders, they only offer limited knowledge right at the initial steps of optimizing them further, thus necessitate extensive structure-activity relationship (SAR) analyses (significant synthetic efforts). Fragments show binding affinities in the range of mM to 10<sup>1</sup>-10<sup>2</sup> μM<sup>4</sup>, thus screening methods include biophysical technologies, such as Surface Plasmon Resonance (SPR)<sup>5-7</sup> (high-throughput rapid method, immobilized protein), Isothermal Titration Calorimetry (ITC)<sup>8</sup> (label-free method, measuring enthalpy changes during fragment-binding), Thermal Shift Assay (TSA)<sup>9,10</sup> (heat-mediated unfolding). Other screening techniques include X-Ray crystallography-based screening<sup>11-13</sup> (enables direct insight and understanding of the binding event), and Nuclear Magnetic Resonance spectroscopy-based screenings<sup>14</sup> that is either protein-observed NMR<sup>15</sup> (appropriate method for the exploration of the binding site), or

ligand-observed NMR<sup>16</sup> (label-free method, provides solubility and stability information). Mass-Spectrometry techniques<sup>17,18</sup> (high sensitivity method, low amount of target protein necessary) may also be advantageously used in fragment screening. Compared to the HTS, fragment screening (FS) campaigns typically provide improved hit rates and hits with higher ligand efficiency<sup>3</sup> This strategy puts emphasis on higher efficiency hits, whereas HTS screenings are focusing on only providing high potency hits, but limited possibilities and knowledge for optimization. Beyond the mentioned conventional screening methodologies (binding-studies) offering more insight to the binding event at the binding site, high content screenings (HCS)<sup>19</sup> are monitoring changes in cell phenotype using automated image analysis. Compared to the conventional biophysical/biochemical assay-based fragment screenings HCS methods suffer limitations: detection of false positive fragment hits by unspecific binding; offering less insight to the binding event, and the complex evaluation of cellular events remains a challenging task.

### 1.1.2. Strategies and examples

The first reported successful FBDD-driven study was published in 1996 presenting a SAR by NMR approach<sup>20</sup>, followed by several success stories of drug companies Astex, Vernalis and Sunesis around the 2000s. FBDD is a well-appreciated strategy resulting in a significant number of chemical probes, clinical candidates and marketed drugs (see **Table Appendix 1**<sup>21,22</sup>).

#### 1.1.2.1. Fragment-based drug design workflow

Fragment libraries are constructed from compounds with small size and low complexity. According to the RO3 criteria<sup>23-25</sup>:

- Molecular weight  $\leq 300$
- Number of hydrogen bond donors and acceptors  $\leq 3$
- clogP  $\leq 3$
- Number of rotatable bonds  $\leq 3$

Molecular weight criterion is sometimes replaced by a restriction on heavy atom count that is between 8 and 22.

- Polar surface area (PSA) is often restricted to  $PSA \leq 60 \text{ \AA}^2$ .

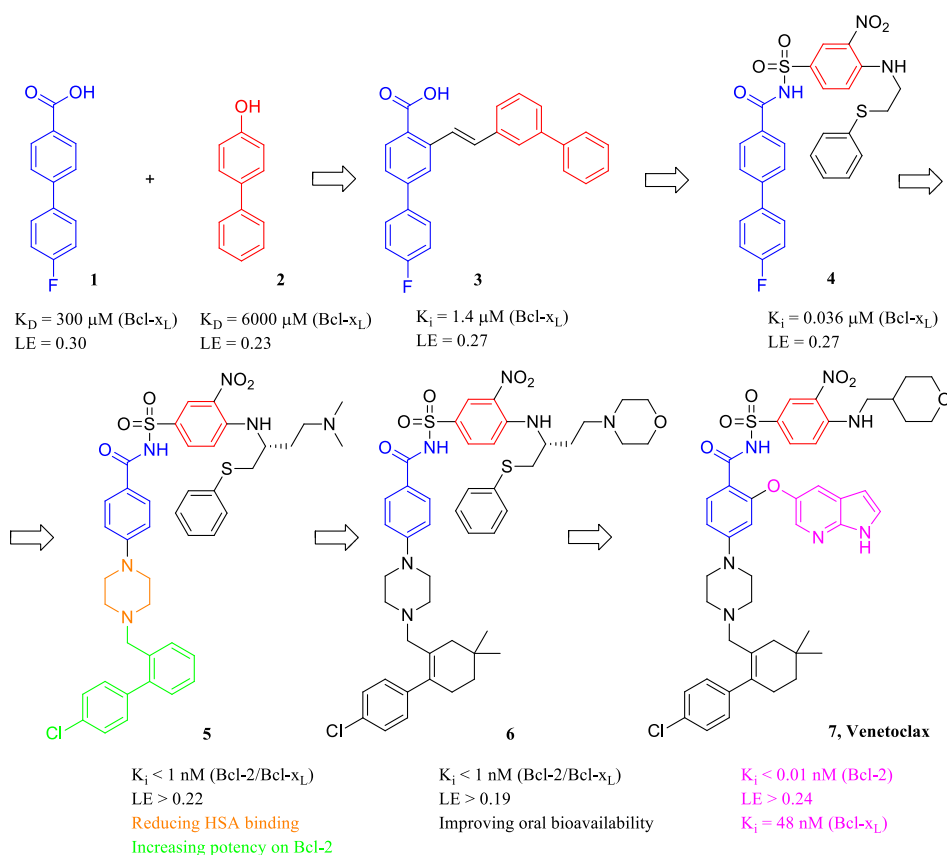
Further steps and consideration for efficient screening include:

- Filtering out reactive functional groups
- Solubility at higher concentrations
- Stability, no aggregation (i.e. lowering the number of false positives)
- Diversity (i.e. better sampling of the chemical universe)

Regarding the different fragment screening techniques, relative sensitivities of the certain methods, solubility and aggregation at the specific assay conditions should be taken into consideration.

### 1.1.2.2. Fragment optimization strategies

Fragment screening hits can be optimized to lead compounds with improved affinities and selectivities across a panel of targets by following different fragment-optimization strategies (see **Table Appendix 1** for a collection of FBDD lead-molecules reached clinics<sup>21,22</sup>). The approved drug Venetoclax (**7**) used for the treatment of chronic lymphocytic leukemia<sup>26</sup> might serve as a general example for presenting several optimization methods (see **Figure 1**).



**Figure 1** – Fragment optimization pipeline presenting fragment linking (1 with 2), growing (4 to 5) and lead progression (6 to 7) on the example of approved drug Venetoclax. – Improvements in selectivities during the optimization steps are marked as colored substructures (orange, green, purple highlightings).

*Fragment linking* <sup>27–36</sup>

The chemical linkage of fragment hits binding at separate subpockets of the binding site may lead to an additive improvement in binding free-energy that corresponds to a dissociation constant equal to the product of the dissociation constants of the constituting fragments<sup>37</sup>. This requires although the binding to be preserved even when introducing rigidity (entropy loss)<sup>38</sup>.

Fragment linking may be preceded by several steps. After the identification of a primary hit by NMR<sup>20</sup>, or MS-screening<sup>39</sup>, it is usually optimized. Consecutively, a second fragment hit is identified by screening in the presence of the first hit binding to the primary binding pocket of the target. The initial fragments **1** and **2** in **Figure 1** were identified as micromolar-milimolar fragment hits on B-cell lymphoma-extra large (Bcl-X<sub>L</sub>). Their linking afforded **compound 3** with a low micromolar binding, followed by the optimization of the structural motif highlighted in red to improve its affinity to the nanomolar range (**4**).

#### *Fragment merging*<sup>40-42</sup>

Merging uses two fragment hits binding at distinct or overlapping positions. Constructing a compound that includes structural elements of both fragments is called merging. This methodology requires structural information most readily obtained with screening by X-Ray<sup>11-13</sup>.

#### *Fragment growing*<sup>43-47</sup>

The growing, or so called fragment-evolution of initial fragment hits using structure-based design, or SAR by catalogue approaches might produce high quality lead-like molecules. This is the most common methodology in fragment-based optimization exemplified by the optimization process leading to vemurafenib<sup>48</sup> and to Venetoclax (**7**), the details of the latter is shown in **Figure 1**; growing **intermediate 4** to **5** by a *N,N*-dimethylethanamine moiety, optimized further to a morpholine ring in **compound 6**.

## 1.2. G-Protein Coupled Receptors

### 1.2.1. Signalization and classification

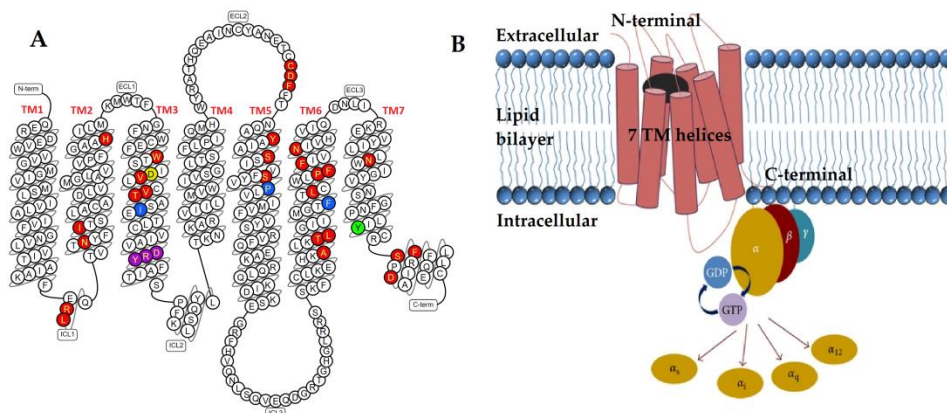
Receptors are microswitches/transmitters of chemical communication between cells. Endogenous ligands (neurotransmitters, hormones, cytokines, mediators, intracellular messengers, etc.) of receptors are specifically enabled at the appropriate location and concentrations, and bind to the orthosteric pockets of the receptor macromolecules. All further secondary sites are considered as allosteric pockets. Receptors are capable of transforming the binding event into signal transmission towards other proteins, messengers.

G-protein-coupled receptors (GPCR, 7TM-receptors) are localized in the cell-membrane (see **Figure 2**), effecting signalization through G-protein, and  $\beta$ -

arrestin. GPCR's are expressed in the peripheral and central-nervous system, functionally coupling to a heterotrimeric ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) G-protein (Figure 2/B).

The approx. 40 % of marketed drugs target GPCR's, thus representing their importance in medicine. Treatment of CNS-related symptoms (Alzheimer's Disease, schizophrenia, depression, cognitive deficit, anxiety, etc.), and other important diseases such as hypertension, diabetes, sepsis, obesity, cancer, etc. are all covered by the indications of GPCR targets.

GPCR's are classified on the basis of sequence similarities (phylogenetic tree), pharmacological function, and signalization pathways (different G-proteins,  $\beta$ -arrestin). The binding site of G-proteins is located at the intra-cellular loop (ICL) region of the receptor. The  $\alpha$ -subunit is variable across the different G-proteins, thus defining:  $\alpha_s$ ,  $\alpha_q$ ,  $\alpha_{12,13}$ ,  $\alpha_{11,14}$ ,  $\alpha_{15,16}$ ,  $\alpha_{i/o}$ ,  $\alpha_{sens}$  subclasses. When signalling through the  $\alpha$ -subunit, cAMP, prostanoid, IP3 and DAG are the secondary messengers, being released by the so-called effector enzymes (adenyl-cyclase, phospholipase A2/C), while the  $\beta\gamma$  dimer modulates ion-channels (potassium-, and calcium release).



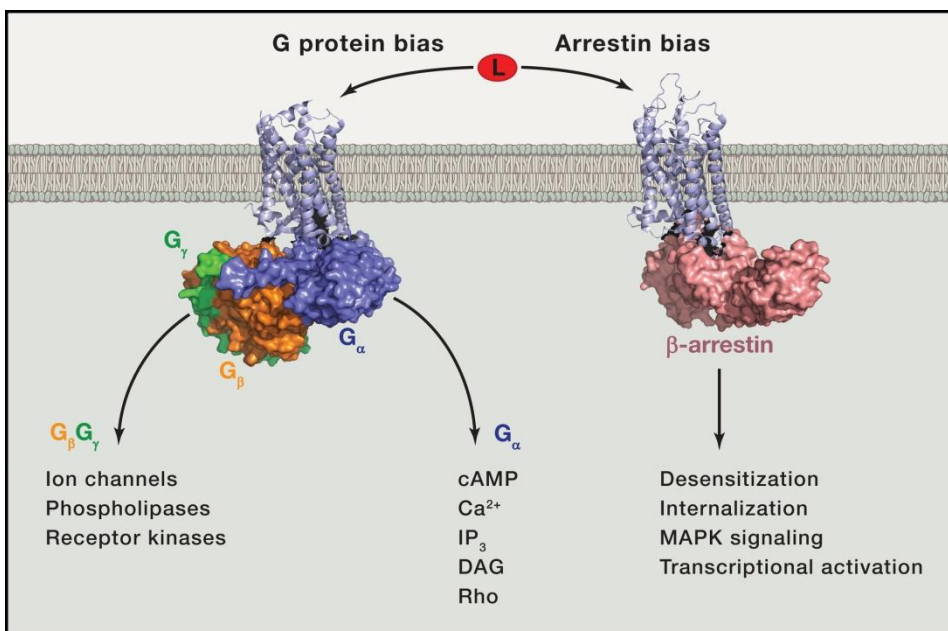
**Figure 2** – Schematic representations of the general GPCR structure and signalization. **2/A**: Structure of the  $\beta_2$ AR receptor (image generated and downloaded at GPCRDB<sup>49</sup>), Key residues for ligand interactions are marked as follows; Asp<sup>3.32</sup> in yellow, DRY-motif (Asp<sup>3.49</sup>, Arg<sup>3.50</sup>, Tyr<sup>3.51</sup>) in purple, Tyr<sup>7.53</sup> of NPxxY-motif in green, P-I-F motif (Pro<sup>5.50</sup>, Ile<sup>3.40</sup>, F<sup>5.44</sup>) in blue, other residues contacting with ligands in red. **2/B**: Schematic representation of a GPCR embedded in the lipid bilayer (image from ref <sup>50</sup>). Red sticks represent transmembrane helices, yellow, brown and blue circles represent G-protein with  $\alpha$ ,  $\beta$ ,  $\gamma$ -subunits. Black circle represents orthosteric ligand embedded in the binding pocket

Among the common structures of GPCR's, all consist of 7 trans-membrane helices and the additional intramembrane helix 8. Helices are connected in a bunch through 3-3 intra-, and extracellular loops. Beyond the highly conserved amino acid residues localized across the helices either in the proximity of the

binding pocket, or participating in the well-established activation micro-switches: DRY<sup>51</sup> and NpxxY<sup>52</sup>, and P-I-F motifs<sup>53</sup>, or in the conserved disulfide bridge<sup>54</sup>, the most conspicuous differences are present in the orthosteric binding pocket (TM3-5-6-7) connected to functional selectivity and in the extended allosteric site near the ECL2 region connected to subtype-selectivity<sup>55</sup>.

Signalling through the G-protein initiates when endogenous agonist binds to the orthosteric pocket (see **Figure 3** – left part), and stabilizes an active conformation. This event is exemplified as a “snapshot” of active-like structures in case of the rhodopsin/opsin<sup>56</sup>, and  $\beta$ 2AR receptors<sup>57</sup>. The local binding event at the binding site triggers amplified intramolecular activation pathways through the outward movements of transmembrane helices TM 5 and TM6, and inward movements of TM3 and TM7, showing the most significant movement distances at the cytoplasmatic face of the receptor (contacting the G-protein). The inward movement of Tyr<sup>7.53</sup> (part of NPxxY, part of TM7) (Ballesteros-Weinstein numbering in superscript)<sup>58</sup> is directed towards the original position of TM6 in the inactive state, whereas the inward movement of TM3 occurs via the breaking of the DRY ionic lock between Asp<sup>3.49</sup> and Arg<sup>3.50</sup> highly conserved residues. Consequently through the caused major conformational changes, the  $\alpha$ -subunit splits from the heterotrimeric protein, via a GDP-GTP energy transfer (mediated by Mg<sup>2+</sup>). Thus the  $\alpha$ -GTP complex binds to different downstream effector proteins (enzymes, i.e. adenylyl-cyclase or phospholipase-C), modulating (increasing or blocking) the production of secondary messengers, such as cAMP, prostanooids, IP3, DAG, thus indirectly affecting on glucose levels, and Ca<sup>2+</sup> levels. In an alternative route GPCR's might affect ion-channels (K<sup>+</sup>, Ca<sup>2+</sup>), through the binding of the  $\beta\gamma$  dimer.

The phosphorylation at the C-terminus of the activated GPCR by a GPCR-kinase makes the GPCR bind with high affinity to multifunctional scaffold proteins (so called  $\beta$ -arrestins)<sup>59,60</sup>. Similar processes may occur through arrestin-mediated signalization (see **Figure 3** – right part). The binding of the  $\beta$ -arrestin prevents the G-proteins from coupling to the GPCR's by steric hindrance mechanism (desensitization)<sup>61</sup>. Also arrestins might act as signal transducers, either by activating mitogen-activated protein kinases (MAPK ERK, JNK), thus regulating cytoskeleton, protein synthesis, cell migration and apoptosis, or modulating tyrosine-kinases (Src), or E3 ubiquitin ligases (AIP4).



**Figure 3** – Functional pathways of GPCR's: Signalization pathways through G-protein subunits and alternatively through β-arrestin. (Image from ref <sup>62</sup>).

GPCRs are divided into six classes based on sequential and functional similarities; Class A (rhodopsin-like receptors), Class B (the secretin receptor family), Class C (metabotropic glutamate receptors), Class D (fungal mating pheromone receptors), Class E (cyclic AMP receptors) and Class F (Frizzled and Smoothed receptors). Hereby class A constitutes the largest family, including aminergic, chemokine, glycoprotein hormone receptors and neuropeptide receptors. Aminergic receptors are the largest group of class A that contain muscarinic achetylcholine, adrenoceptors, dopamine-, histamine-, serotonin-, octopamine-, and trace amine-receptors. Aminergic receptor subfamilies and subtypes together with their pharmacological relevance are summarized in **Table 1**.

5-HT	Serotonin-receptors
5-HT <sub>1A</sub> R	agonists: hypertension, anxiety
5-HT <sub>1B</sub> R	partial agonists, antagonists: antipsychotics
5-HT <sub>1D</sub> R	antagonists: antidepressants
5-HT <sub>1E</sub> R	Not evaluated yet
5-HT <sub>1F</sub> R	Not evaluated yet
5-HT <sub>2A</sub> R	agonists: psychedelics, antagonists: DRD2 antipsychotics add-on
5-HT <sub>2B</sub> R	<b>IBS, chronic heart disease, migraine prophylaxis</b>

5-HT <sub>2C</sub> R	agonists: addictive disorders, drug abuse
5-HT <sub>4</sub> R	Ongoing research
5-HT <sub>5A</sub> R	Ongoing research
5-HT <sub>5B</sub> R	Ongoing research (antipsychotics)
5-HT <sub>6</sub> R	<b>antagonists: cognitive, AD, PD, schizophrenia</b>
5-HT <sub>7</sub> R	<b>antagonists: schizophrenia, depression, epilepsy, migraine, circadian rhythm</b>
mAChR	Muscarinic acetylcholine receptors
m <sub>1</sub> AChR	agonists, PAM's: learning, memory, cognition, AD, schizophrenia, drug addiction
m <sub>2</sub> AChR	adverse side-effects
m <sub>3</sub> AChR	adverse side-effects
m <sub>4</sub> AChR	agonists, PAM's: learning, memory, cognition, AD, schizophrenia, drug addiction
m <sub>5</sub> AChR	agonists, PAM's: learning, memory, cognition, AD, schizophrenia, drug addiction
ADR	Adrenoceptors
α <sub>1A</sub> AR	antagonists: benign prostatic hyperplasia
α <sub>1B</sub> AR	Ongoing research
α <sub>1D</sub> AR	Ongoing research
α <sub>2A</sub> AR	Ongoing research, cognition, agonists: hypertension
α <sub>2B</sub> AR	Ongoing research, agonists: hypertension, post anesthetic care
α <sub>2C</sub> AR	Ongoing research, cognition, agonists: hypertension
β <sub>1</sub> AR	antagonists: beta-blockers (hypertension, coronary heart disease)
β <sub>2</sub> AR	agonists: asthma, antagonists: beta-blockers (hypertension, coronary heart disease)
β <sub>3</sub> AR	agonists: antistress
DRD	Dopamine receptors
DRD <sub>1</sub>	agonists: hypertension, antagonists: antipsychotics
DRD <sub>2</sub>	antagonists: schizophrenia positive symptoms
DRD <sub>3</sub>	agonists, partial agonists: schizophrenia, bipolar mania, major depressive disorder, antagonists: antipsychotics
DRD <sub>4</sub>	Ongoing research, cognition, polymorphism ADHD
DRD <sub>5</sub>	cooperation with M1 and beta2
HR	Histamine receptors
H <sub>1</sub> R	antagonists: anti-inflammatory, anti-allergic
H <sub>2</sub> R	antagonists: peptic ulcer disease, GERD, dyspepsia

H <sub>3</sub> R	antagonists: sleep disorders, Tourette syndrome
H <sub>4</sub> R	antagonists: asthma, anti-inflammatory, anti-allergic
TAAR	Trace-amine receptors
TAAR <sub>1</sub>	antagonists: drug addiction, dimerization with DRD2

**Table 1** - Pharmacological relevance of ClassA/Aminergic GPCR's (based on GPCRDB phylogenetic classification<sup>49</sup>). Indications highlighted in bold are covered by the studies presented in the thesis.

### 1.2.2. X-Ray crystal structures of aminergic GPCR's

Since the first GPCR structure was solved in 2000<sup>63</sup>, an emerging number of class-A aminergic GPCR receptor X-ray structures have been published (**Table 2**). The large number of available structures raises the possibility for using their combinations in docking experiments to identify fragments for aminergic receptors.

PDB ID	Receptor	Ligand (ID, name)	Functional	Res.(Å)	Year
5DSG	m <sub>1</sub> AchR	0HK, tiotropium	antagonist	2.7	2016
3UON	m <sub>2</sub> AchR	QNB, Quinuclidinyl diphenylacetate	antagonist	3.0	2012
4MQS	m <sub>2</sub> AchR	IXO, iperoxo	agonist	3.5	2013
4MQT	m <sub>2</sub> AchR	IXO, iperoxo; 2CU	agonist, allosteric modulator	3.7	2013
5DSG	m <sub>4</sub> AchR	0HK, tiotropium	antagonist	2.6	2016
2RH1	β <sub>2</sub> AR	CAU, carazolol	inverse agonist	2.4	2007
3D4S	β <sub>2</sub> AR	TIM, timolol maleate	agonist	2.8	2008
3NY8	β <sub>2</sub> AR	JRZ, ICI 118,551	inverse agonist	2.8	2010
3NY9	β <sub>2</sub> AR	JSZ, ChEMBL1233771	inverse agonist	2.8	2010
3NYA	β <sub>2</sub> AR	JTZ, alprenolol	neutral antagonist	3.2	2010
3PDS	β <sub>2</sub> AR	ERC	irreversible agonist	3.5	2011
4GBR	β <sub>2</sub> AR	CAU, carazolol	inverse agonist	4.0	2012
4LDE	β <sub>2</sub> AR	P0G, PubChem 49837874	agonist	2.8	2013
4LDL	β <sub>2</sub> AR	XQC, ChEMBL1160696	agonist	3.1	2013
4LDO	β <sub>2</sub> AR	ALE, L-epinephrine	agonist	3.2	2013
3P0G	β <sub>2</sub> AR	P0G, PubChem 49837874	agonist	3.5	2011

<b>2R4R</b>	$\beta_2$ AR	-	-	3.4	2007
<b>2R4S</b>	$\beta_2$ AR	-	-	3.4	2007
<b>3KJ6</b>	$\beta_2$ AR	-	-	3.4	2010
<b>4QKX</b>	$\beta_2$ AR	35V	covalent agonist	3.3	2014
5D5B	$\beta_2$ AR	CAU, carazolol	inverse agonist	3.8	2016
5D5A	$\beta_2$ AR	CAU, carazolol	inverse agonist	2.5	2016
5X7D	$\beta_2$ AR	CAU, carazolol	inverse agonist	2.7	2017
5D6L	$\beta_2$ AR	CAU, carazolol	inverse agonist	3.2	2017
5JQH	$\beta_2$ AR	CAU, carazolol	inverse agonist	3.2	2016
<b>4IAQ</b>	5-HT <sub>1B</sub> R	2GM, dihydroergotamine	agonist	2.8	2013
<b>4IAR</b>	5-HT <sub>1B</sub> R	ERM, ergotamine	agonist	2.7	2013
<b>4IB4</b>	5-HT <sub>2B</sub> R	ERM, ergotamine	agonist	2.7	2013
<b>4NC3</b>	5-HT <sub>2B</sub> R	ERM, ergotamine	agonist	2.8	2013
5TVN	5-HT <sub>2B</sub> R	7LD, lysergic acid diethylamide	agonist	2.9	2017
5TUD	5-HT <sub>2B</sub> R	ERM, ergotamine	agonist	3.0	2017
6BQG	5-HT <sub>2C</sub> R	ERM, ergotamine	agonist	3.0	2018
6BQH	5-HT <sub>2C</sub> R	E2J, Ritanserin	antagonist	2.7	2018
<b>3PBL</b>	DRD <sub>3</sub>	ETQ, eticlopride	antagonist	2.9	2010
5WIU	DRD <sub>4</sub>	AQD, nemonapride	antagonist	2.0	2017
5WIV	DRD <sub>4</sub>	AQD, nemonapride	antagonist	2.1	2017
<b>3RZE</b>	H <sub>1</sub> R	5EH, doxepin	antagonist	3.1	2011

**Table 2.** - Table of 37 available human aminergic X-ray structures (as of 2018). Entries with bolded PDB ID codes were considered for the selection of the FrACS ensemble in **Section 2.2.1.1., page 42** and **3.1.2.1., page 84**.

## 1.3. Medicinal Chemistry of Serotonin (5-HT) receptor ligands

### 1.3.1. Serotonin Receptor Subtype 6 (5-HT<sub>6</sub>R)<sup>64</sup>

#### 1.3.1.1. Pharmacological relevance

The 5-HT<sub>6</sub>R is considered as a current and promising drug target for the treatment of CNS-related indications, such as cognitive, learning and memory deficits related to Alzheimer's disease<sup>65</sup>, Parkinson's disease<sup>66</sup> and schizophrenia<sup>67</sup>. Other CNS indications such as analgesia<sup>68</sup>, obesity<sup>69</sup>, anti-drug-abuse<sup>70</sup>, sleep-wake regulation<sup>71</sup>, self-transcendence<sup>72</sup>, executive cognition<sup>73</sup>, and major depressive disorder<sup>74</sup> have also been suggested.

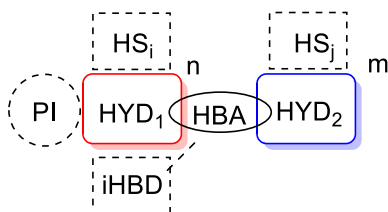
The serotonergic subfamily consists of seven receptor types (5-HT<sub>1-7</sub>). Unlike other serotonin receptors, subtype 6 is almost exclusively present in the (CNS)<sup>75</sup>, with infinitesimal occurrence in peripheral organs. Blockade of the 5-HT<sub>6</sub>R influences the cholinergic, glutamatergic and GABAergic systems by increasing acetylcholine and glutamate release. Initial observations on the 5-HT<sub>6</sub>R pharmacology were related to the finding that many antipsychotics and tricyclic antidepressants display a high affinity for this receptor<sup>76</sup>. Although 5-HT<sub>6</sub>R is considered to be a therapeutically relevant drug target, there are a number of challenges in the discovery of a promising 5-HT<sub>6</sub> antagonist. The main goals and challenges include penetration through the blood-brain barrier (BBB), selectivity towards the other 5-HT subtypes and other targets, ensuring a low number of drug metabolites and overcoming rapid desensitization of the receptor, signalling interference, and the unclear relation between the signalling pathways.

#### 1.3.1.2. Known 5-HT<sub>6</sub>R ligands and pharmacophore models

Chemical similarity to the endogenous neurotransmitter, serotonin (5-hydroxytryptamine) can be recognized in most frequent heteroaromatic ring systems routinely used in 5-HT<sub>6</sub>R ligands that include indoles, indolines, indazoles, pyrrolo[2,3-b]pyridines, pyrazolo[1,5-a]pyridines, [1,2,3]-triazolo[1,5-a]pyrimidines and further 5 + 6 condensed *N*-heterocycles as found in ChEMBL<sup>64</sup>.

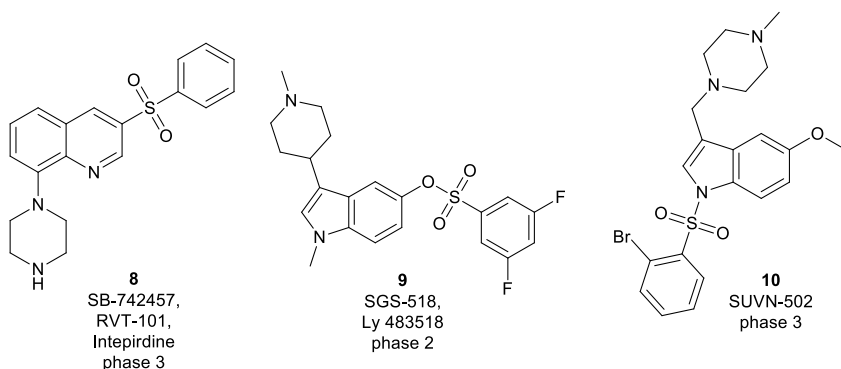
Pharmacophore based approaches on 5-HT<sub>6</sub>R antagonists are typically focused on the canonical pharmacophore model<sup>77</sup> (**Figure 4**), which is defined as having two hydrophobic rings/ring systems (HYD<sub>1</sub> and HYD<sub>2</sub>) connected by a hydrogen bond acceptor (e.g. sulfonyl, sulfonamide linker). Optionally, a positively ionizable group (connected to HYD<sub>1</sub>) offers interactions with the conserved Asp<sup>3.32</sup> aspartate residue of aminergic GPCR's as a key factor of aminergic 7TM-receptor activation<sup>78</sup>. A further pharmacophore feature might

contain an additional intramolecular hydrogen bond donor moiety further stabilizing the binding conformation of the ligands<sup>79</sup>. Selectivity against other aminergic GPCR's was shown<sup>80</sup> to be accessible through omitting the positively ionizable group in the 5-HT<sub>6</sub>R antagonists. Bis(hetero)arylsulfonyl-, and sulfonamide substituents also contribute to 5-HT<sub>6</sub>R affinity and selectivity.



**Figure 4.** General pharmacophore model of 5-HT<sub>6</sub>R antagonists. PI: positive ionisable portion; HYD<sub>1,2</sub>: hydrophobic sites; iHBD: intramolecular hydrogen bond donor site; HBA: hydrogen bond acceptor feature; HS<sub>i,j</sub>: substituents on the hydrophobic sites.

An analysis of by Cummings and colleagues<sup>81</sup> (see also analyses at [clinicaltrials.gov](http://clinicaltrials.gov)<sup>82</sup> and at [ADDF Clinical Trials Report](http://ADDF.ClinicalTrialsReport.com)<sup>83</sup>) concluded that 25 neuroprotective agents were under investigation for AD in phase 1 (2), phase 2 (19) and phase 3 (4) clinical stages. The following four drug candidates (three presented in **Figure 5**) reached clinical phases 2 and 3 as 5-HT<sub>6</sub>R antagonists: Intepirdine/RVT-101/SB-742457 **8** (phase 3 - rejected; GSK, Axovant), Lu-AE58054/idalopirdine/SGS-518 **9** (phase 3 - rejected; Lundbeck), SUVN-502 **10** (phase 2; Seven LifeSciences), PF-0521377/SAM-760 – undisclosed structure (phase 2; Pfizer). The most prominent drug candidates and investigational compounds with high relevance are summarized in **Table Appendix 2**.



**Figure 5** – Structures of the most prominent 5-HT<sub>6</sub>R drug candidates reached clinical phases

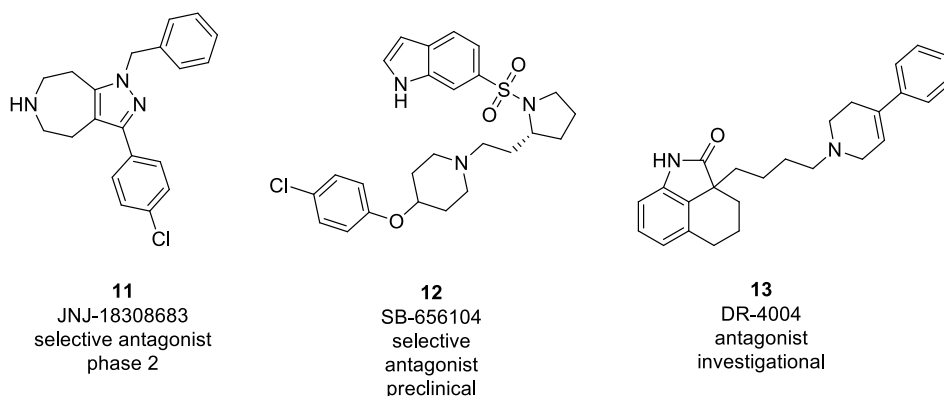
## 1.3.2. Serotonin Receptor Subtype 7 (5-HT<sub>7</sub>R)

### 1.3.2.1. Pharmacological relevance

The 5-HT<sub>7</sub>R was the last identified member of the serotonin family, and its' gene was discovered in 1993<sup>84</sup>. It is expressed both in the CNS and in peripheral tissues, and is coupled positively to adenylyl cyclase (activation raises cAMP levels) through G<sub>αs</sub> and G<sub>α12</sub>. The 5-HT<sub>7</sub>R receptor plays role in the regulation of body temperature, sleep-wake rhythm, circadian rhythm, and mood. Thus the modulation of this receptor has become rapidly an important target for several important CNS-related indications such as depression<sup>85,86</sup>, sleep disorders<sup>87</sup>, stress<sup>88</sup>, anxiety<sup>89</sup>, learning/memory deficits<sup>90</sup>, schizophrenia-like cognitive deficits<sup>91</sup>, epilepsy<sup>92</sup>, nociception<sup>93</sup>, migraine<sup>94</sup>, autism spectrum disorders<sup>95</sup>, and Rett Syndrome<sup>96</sup>. Currently there is a debate whether 5-HT<sub>7</sub>R agonists or antagonists are more beneficial in the treatment of the above listed indications<sup>90,97</sup>.

### 1.3.2.2. Known 5-HT<sub>7</sub>R ligands and pharmacophore models

To date one drug candidate with the pyrazolo[3,4-d]azepine core (JNJ-18308683, **11**) has reached clinics (phase 2 currently recruiting<sup>82</sup>), and a number of investigational compounds has been proved to be selective antagonists (**12**, **13**, **Table Appendix 3**) or agonists (see **Figure 6** and **Table Appendix 3**).



**Figure 6** – Structures of the most prominent 5-HT<sub>7</sub>R selective antagonists

A 3D pharmacophore model was built<sup>98</sup> on the basis of thirty 5-HT<sub>7</sub>R antagonists (DR-4004 (**13**) and analogues, SB-258719 and analogues, SB-269970 analogue) using the Catalyst software<sup>99</sup>. An optimized version<sup>100</sup> of the model was established using extensive SAR analysis, mutagenesis evaluation and structure-based docking experiments (2D depiction of the original model shown in **Figure 7**). The model defines one aromatic structural part (stacking with Phe<sup>3,28</sup> and Tyr<sup>7,43</sup>), and two further hydrophobic regions HYD1 and HYD2 (facing Phe<sup>6,52</sup>), one hydrogen bond acceptor group (binding to Ser<sup>5,42</sup> and Tyr<sup>5,43</sup>)

next to HYD1, and 5-6 Å apart from the HYD1 a positive ionisable moiety contacting Asp<sup>3.32</sup>.



**Figure 7** – General pharmacophore model of 5-HT<sub>2</sub>R antagonists. PI: positive ionisable portion; HYD1-3: hydrophobic sites; HBA: hydrogen bond acceptor feature

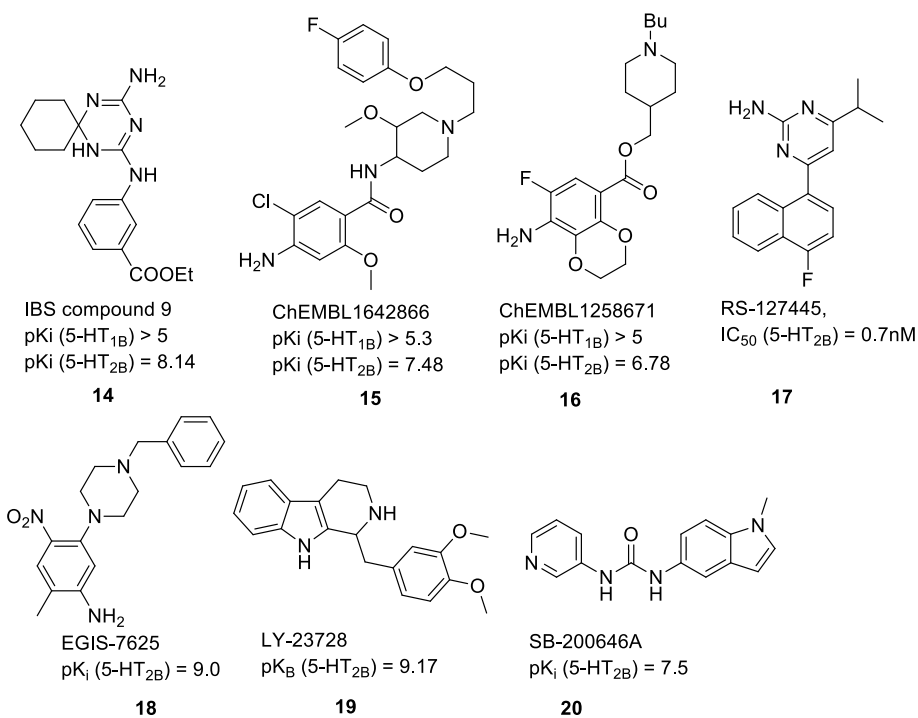
### 1.3.3. Serotonin Receptor Subtype 2B (5-HT<sub>2B</sub>R)

#### 1.3.3.1. Pharmacological relevance

Non-selective agonists of 5-HT<sub>2B</sub>R receptor (e.g. ergolines<sup>101</sup>) are causing acute side-effects of hypertension<sup>102</sup>, cardiac valvulopathy<sup>103</sup> and retroperitoneal fibrosis<sup>104</sup>. In contrast, evidence has shown that inhibition of the receptor is beneficial in the treatment of migraine<sup>105-107</sup>, chronic heart disease<sup>108</sup>, and irritable bowel syndrome (IBS)<sup>109</sup>.

#### 1.3.3.2. Known 5-HT<sub>2B</sub>R ligands

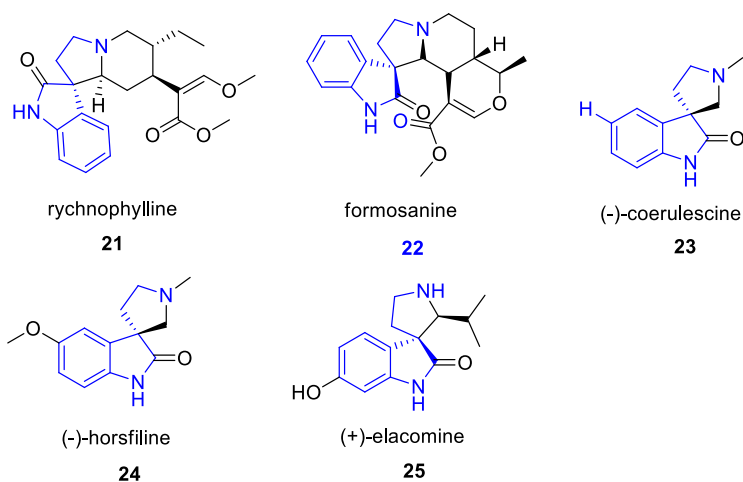
Most prominent 5-HT<sub>2B</sub>R antagonists might be represented by a couple of chemotypes (summarized in **Figure 8**). Examples of high affinity 5-HT<sub>2B</sub>R compounds showing selectivity against 5-HT<sub>1B</sub>R are represented by triazines (**14**)<sup>109</sup>, piperidines (**15**, **16**) pyrimidines (e.g. RS-127445 (**17**)<sup>110</sup>, arylpiperazines (m-CPP<sup>111</sup>, EGIS-7625 (**18**)<sup>112</sup>), tetrahydro-β-carbolines<sup>113</sup> (LY-23728 (**19**), LY-272015, LY-266097 and aryl-ureas (SB-200646A (**20**), SB-204741, SB-215505<sup>114</sup>).



**Figure 8** - Representative 5-HT<sub>2B</sub>R ligands (compounds **14-16** with available 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R binding affinity data, and examples of clinical candidates **17-20**)

## 1.4. Synthetic approaches leading to spiro[pyrrolidine-3,3'-oxindoles]-analogues

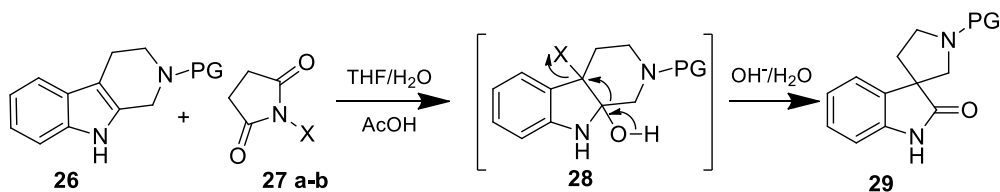
The recent isolation of naturally occurring and biologically active spiropyrrrolidinyl-oxindole alkaloids initiated a significant research on synthetic derivatives<sup>115</sup>, particularly compounds with the spiro[indoline-3,3'-pyrrolidine]-2-one ring system. Spiropyrrrolidinyl-oxindole alkaloids were isolated from different species including *Gelsemium sempervirens*, *Aspidosperma*, *Mitragyna*, *Ourouparia*, *Rauwolfia*, *Vinca* species<sup>116</sup>. The most prominent examples are rynchophylline (**21**), formosanine (**22**), coerulescine (**23**), horsfiline (**24**) and elacomine (**25**) that show diverse biological activity (**Figure 9**). Although these compounds have a common tryptamine derived core, a ChEMBL-analysis<sup>117</sup> showed no representatives tested against serotonergic targets.



**Figure 9** - Representative natural spiro[pyrrolidinyl-oxindole] alkaloids. The tryptamine scaffold is shown in blue.

Derivatization of the spiro[pyrrolidine-3,3'-oxindole] scaffold required a viable synthesis strategy for the preparation of designed analogues. The first, conventional approach is based on an intramolecular Mannich-reaction used in case of several alkaloids including ( $\pm$ )-horsifiline (**24**)<sup>118</sup> and spirotryprostatin B<sup>119</sup>. In another example intermolecular Mannich-reaction afforded a diastereomeric mixture of a spirotryprostatin B precursor, leading to the synthesis of the alkaloid<sup>119</sup>.

An alternative, and well-established approach is the conversion of *N*-protected tryptolines **26** to the corresponding spiro[pyrrolidine-3,3'-oxindoles] **29** by an oxidative spiro-rearrangement reaction induced by an oxidative agent (tert-butyl hypochloride, *N*-chlorosuccinimide (NCS) **27a**, *N*-bromosuccinimide (NBS) **27b**, sodium tungstate, lead tetraacetate, or osmium tetroxide)<sup>120-129</sup>. The reaction is proposed in the literature<sup>129,130</sup> to be completed by the addition of water and bromine forming **28**, and the subsequent elimination of hydrogen-bromide from **28** that finally results in the reorganization of the ring system to spiro[pyrrolidine-3,3'-oxindoles] **29** (Scheme 1).



PG: protecting group, e.g. tert-butyloxycarbonyl (Boc), benzyl (Bn), benzyloxy carbamate (Cbz), etc.  
X: **27a** Cl, **27b** Br

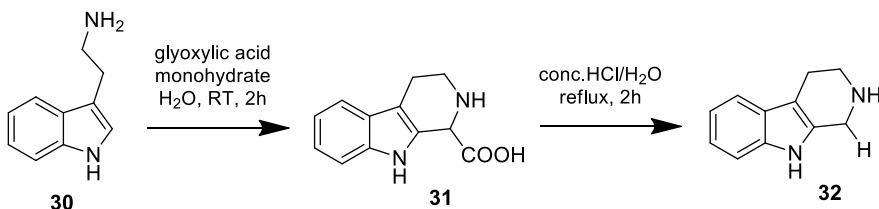
**Scheme 1** - Mechanism of the succinimide assisted oxidative spiro-rearrangement of tryptolines.

Further, more sophisticated approaches are summarized in **Table 3**, such as [1,3]-dipolar cycloaddition reactions<sup>131</sup>, applying azomethine ylides in reaction with oxindolin-3-ylidene, or radical cyclization using AIBN as reagent<sup>132</sup>, or intramolecular Heck-reactions used for the stereoselective synthesis of spirotryprostatin B<sup>133</sup>; and asymmetric nitroolefination of oxindoles presented in Fuji's synthesis of (-)-horsfiline<sup>134</sup>, showing a widespread applicability, and diverse strategies in targeting the synthesis of the spiro[pyrrolidine-3,3'-oxindole] based chemistry.

Number	Reaction/mechanism	Reference
1	Inter and intramolecular Mannich-reactions	119,130
2	Stereoselective [1,3]-dipolar cycloaddition	131
3	Assymmetric Heck-cyclization	133
4	Assymmetric nitroolefination	134
5	N-acyliminium ion spirocyclization	135
6	MgI <sub>2</sub> catalyzed ring expansion reaction	136
7	Reductive amination	137
8	Aldol-reaction	138
9	Radical cyclizations	132
10	Oxidative spiro-rearrangements	120–129

**Table 3** - Synthetic strategies to the spiro[pyrrolidine-3,3'-oxindole] scaffold

The oxidative spiro-rearrangement reactions offer a wide variety of oxidative reagents<sup>120–123,125–127,129,139,140</sup> and the corresponding starting materials are readily accessible through Pictet-Spengler reaction of tryptamines<sup>141</sup> (**Scheme 2**).



**Scheme 2** - Synthesis of tryptoline by Pictet-Spengler condensation.

Following this synthesis strategy, tryptamine (**30**) might be used as a starting material for the synthesis of the common intermediate tryptoline (2,3,4,9-tetrahydro-1H-pyrido[3,4-*b*]indole, or tetrahydro-β-carboline) (**32**). The Pictet-Spengler condensation reaction can be performed by glyoxylic acid-monohydrate in aqueous medium<sup>142</sup>. The crude acid intermediate **31** is decarboxylated in refluxing concentrated hydrochloric acid affording the tetrahydro-β-carboline (**32**), the main synthetic intermediate applied for the synthesis of 5-HT<sub>6</sub>R and 5-HT<sub>7</sub>R antagonists presented in **Sections 2.3., page 51** and **2.4., page 65**.

## Chapter 2 – Results and Discussion

In my thesis I followed an FBDD workflow; using in silico ligand- (Section 2.1., page 29), and structure-based methodologies (Sections 2.2., page 41 and 2.5., page 72) in the identification of chemical starting points and in their optimization into compounds with improved affinities and selectivities (Sections 2.3-2.4.).

### 2.1. Physicochemical-property based scoring method for the design of aminergic fragment libraries<sup>143</sup>

In order to reach beneficial ratios of true positives, and true negatives during virtual screening campaigns, computationally cheap ligand-based methodologies might help to narrow the size of initial datasets of multimillion molecules. These pre-filtered sets consisting of compounds with selected properties, or satisfying certain similarity criteria to known actives, are more likely to end up in improved hit rates during screening campaigns. Our objective was to enrich the screening database with compounds predicted to be active on aminergic GPCRs. To achieve this goal we first collected fragment-like aminergic GPCR ligands from public databases. Their characteristic physicochemical features were analysed to derive a scoring scheme suitable to identify aminergic class A GPCR ligands.

#### 2.1.1. Development of the scoring scheme

The analysis of physicochemical property distributions in chemical databases is a commonly used methodology for deriving design principles of screening libraries<sup>144,145</sup>. A representative training set containing 2183 actives and 5000 reference compounds was collected from the ChEMBL database<sup>117</sup> for the analysis of known aminergic fragment sized molecules (see details of data-mining in Section 3.1.1.1., page 81). Fragments were characterized by widely used physicochemical parameters describing the polarity and topological properties of the molecules: such as calculated logP, logD (at pH = 7.4), polar surface area (PSA at pH = 7.4), by pharmacophoric descriptors such as the number of hydrogen bond acceptors (nHBA), number of hydrogen bond donors (nHBD); pK<sub>a</sub> (acidic, related to the strongest center), pK<sub>a</sub> (basic, related to the strongest center) and by topological descriptors such as the number of rotatable bonds (nRot), number of nitrogen atoms (nN) and number of oxygen atoms (nOx). Descriptors were obtained by ChemAxon's JChem for Office<sup>146</sup>. All of the descriptors were calculated for all of the fragments both for the active and for the reference compounds, followed by calculating the distribution properties (median, mean and standard deviation) for each variable (Table 4).

Active set	Statistical descriptors		
Property	Median	Mean	Standard deviation
PSA (at pH = 7.4)	37.30	39.67	21.04
logD (at pH = 7.4)	0.97	0.90	1.62
nN	2.00	2.28	1.26
nOx	1.00	1.08	0.95
nRot	2.00	2.64	1.89
pK <sub>a</sub> (strongest basic)	8.55	7.92	2.85
logP	2.18	2.04	1.53
nHBD	1.00	0.89	0.96
nHBA	3.00	2.70	1.13
Reference set	Median	Mean	Standard deviation
PSA (at pH = 7.4)	56.49	60.49	29.28
logD (at pH = 7.4)	2.03	1.65	2.24
nN	2.00	2.17	1.48
nOx	2.00	2.14	1.52
nRot	3.00	3.36	2.10
pK <sub>a</sub> (strongest basic)	7.00	5.97	2.60
logP	2.41	2.20	1.80
nHBD	1.00	1.27	1.12
nHBA	3.00	3.23	1.50

**Table 4** - Mean, median, and standard deviation values of the calculated physicochemical properties.

Differences in property-distributions were identified by visual inspection and by the analysis of medians, means and standard deviations, comparing aminergic-fragments, and ChEMBL inactive reference fragments (see **Table 4**). The distributions of the hydrogen bond donor (HBD) and acceptor-counts (HBA) and logP found to be similar for the active and reference sets, however the polar surface area (PSA) and the logD of the actives had a median lower than that of the reference set. We suggested that the lower values of logD are due to the basic character of the aminergic ligands. Fragments of the active and reference sets were also compared regarding their acid-base character (**Table 5**). In fact, 83% of the active fragments showed basic character, whereas 16% were neutral, interestingly, none of them were acidic, and only 1% were zwitter-ionic.

In contrast, 15 % of the randomly sampled reference fragments were basic, 71% neutral (vast majority), 12% acidic, and 3% zwitter-ionic.

	pKa (acidic)	pKa (basic)	actives	inactives
acid	<7	>7 or absent	0	12
base	>7 or absent	>7	83	15
neutral	>7 or absent	<7 or absent	16	71
zwitterion	<7	>7	1	3

**Table 5** - Classification of the acid-base characters of active and inactive fragments

Considering the number of rotatable bonds (nRot), aminergic fragments are typically less flexible (containing generally 0 or 1 rotatable bond), than molecules in the reference set with around 3 rotatable bonds in average. Our analysis revealed that fragment hits of aminergic receptors are generally rigid molecules, containing few heteroatoms, that are most often (~83%) basic nitrogens. Consequently, logD (pH = 7.4), number of nitrogen atoms, number of oxygen atoms, number of rotatable bonds, strongest basic pK<sub>a</sub>, and PSA (at pH = 7.4) can be considered as descriptors discriminating between aminergic-like and general fragments. Statistical significance of the differences between selected property distributions was checked by Mann-Whitney U-tests. All of the compared property-distribution pairs are significantly different with a *p* value lower than 0.05. In the following section we gathered some general statements about aminergic fragments, providing standpoints for the creation of a desirability scoring function.

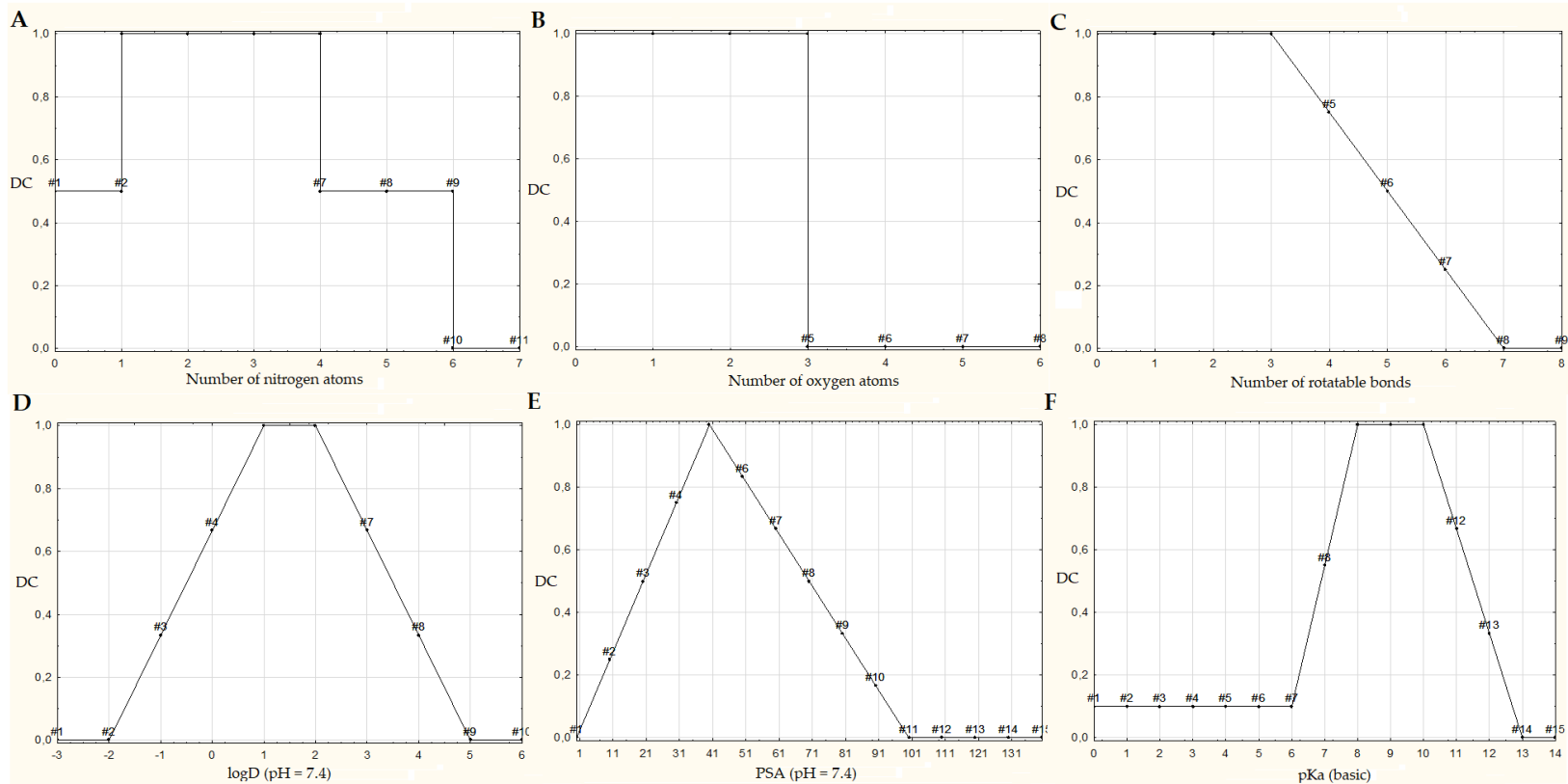
LogD distribution of the active fragments had a mean of 0.90, with a standard deviation of 1.62, which is in line with their predominantly basic character. In contrast, the logD distribution of the randomly selected reference fragments showed a mean = 1.65, and standard deviation = 2.24). Nevertheless, there is no notable difference in the corresponding logP distributions. Counting the number of nitrogen atoms showed that aminergic GPCR-like fragments have at least 1 nitrogen-atom (~40 % contain 1, ~40 % contain 2). The oxygen atoms are not necessarily parts of aminergic fragments (~ 30% do not contain any). The cores of aminergic fragments are typically 1-3 membered (aromatic) ring-systems, thus are less flexible with a mean of 2.64 rotatable bonds, and with 10 % not having any rotatable bonds, unlike random fragments, with a mean of 3.35 rotatable bonds. Only 0.2 % of the aminergic fragments lack any basic-atom whereas 83 % have at least one basic center resulting in a basic pK<sub>a</sub> higher than 7 and having no acidic pK<sub>a</sub> lower than 7. The distribution of basic pK<sub>a</sub> showed a mean value of 7.92 with a standard deviation of 2.84. In contrast, 71% of the reference fragments are neutral (pK<sub>a</sub> mean = 5.96, standard deviation = 2.59). The polar surface area (PSA) of a molecule is defined as the surface sum over all

polar atoms. The widely known Rule of Three<sup>2</sup> filter for fragments limits the polar surface area to be at most 60 Å<sup>2</sup>. The training set had a median value of 37.3 Å<sup>2</sup> and a mean value of 39.67 Å<sup>2</sup>, with a standard deviation of 21.04, thus staying in the range of the criterion.

After determining the characteristic descriptors showing statistically significant differences with respect to the randomly sampled reference fragments we used these properties to derive a desirability scoring scheme.

Approaches for multiparametric filtering<sup>147</sup> include desirability functions<sup>148</sup> that are conveniently used to convert preferred descriptor distributions into interpretable scoring functions. A desirability function maps the value of a variable onto a continuous score in the range of [0; 1]. *Desirable* and *undesirable* property-ranges (x-axis) are defined by the function through series of inflection points<sup>149</sup>, identifying regions of the properties with a certain desirability score of  $y(x)$ . Molecules totally mismatching the criteria of the examined descriptor receive a desirability score of 0. Molecules with *desirable* properties gain a score higher than 0 and at 1 at best. After calculating a desirability score as per properties, the scores may be combined by their summation or multiplication into an overall desirability measure. Desirability functions make it possible to define smooth boundaries for a property, rather than using a rigid cut-off, thus avoiding rejection of compounds based on an uncertain property value close to a criterion boundary (e.g. Lipinski-rules).

The selected properties were next translated to desirability functions (Summarized in **Figure 10**).



**Figure 10** – Desirability functions of six characteristic physicochemical properties of aminergic GPCR ligands. Scatterplots of Desirability scores (DC) of **10/A**: number of nitrogen atoms, **10/B**: number of oxygen atoms, **10/C**: number of rotatable bonds, **10/D**: logD (at pH = 7.4), **10/E**: PSA (at pH = 7.4), **10/F**: basic pK<sub>a</sub> (of the strongest basic center)

The two properties describing the number of heteroatoms (N, O) and the number of rotatable bonds were considered to be discrete variables (**Figure 10/A, B, C**), while logD (at pH = 7.4), PSA (at pH = 7.4) and basic pKa (strongest) are continuous variables (see **Figure 10/D, E, F**). The number of nitrogen atoms is an important feature of aminergic GPCR-fragments due to their mainly basic character that gives a major contribution to the ionic interaction with Asp<sup>3.32</sup> in the orthosteric binding site. However, not all fragments without any nitrogen atom are supposed to be sorted out by the overall desirability score. This is to be underlined since non-basic ligands, such as 5-HT<sub>6</sub>R inhibitors have also been reported<sup>150,151</sup>. The *desirable* property range for the number of nitrogen atoms is between 0 and 5, namely 0 and 4-5 nitrogen atoms are scored with  $y(x) = 0.5$ , and 1-3 nitrogen atoms are scored with  $y(x) = 1$ . A score value of  $y(x) = 0$  was assigned to fragments containing more than 5 nitrogen atoms (**Figure 10/A**). The active fragments of the training set contained mostly 0-3 oxygen atoms supporting the significance of oxygen atoms most probably as hydrogen-bond acceptors. Fragments having 0-3 oxygen atoms are considered to be *desirable* with a score of  $y(x) = 1$  (**Figure 10/B**). It has been suggested<sup>2</sup> that maximum 3 rotatable bonds are advantageous both for binding entropy change and for ADME (absorption, distribution, metabolism, and excretion)/PK (pharmacokinetic) properties. The distribution of the rotatable bond count for the training set showed that ~8 % of the active fragments possessed rigid planar structure, ~21 % contained only 1 and ~27 % only 2 rotatable bonds (most likely standing for 1-2 functional groups). None of the active fragments contained more than 7 rotatable bonds. Consequently, all planar, non-flexible fragments and molecules containing 3 or less rotatable bonds were considered to be *desirable* (score of 1). Less-*desirable* property-ranges of 4-7 are mapped with a monotonic decreasing score function to reach a score of 0 at 7 rotatable bonds. (**Figure 10/C**).

The desirability function for logD (at pH = 7.4) prefers mostly fragments in the range of logD between 1 and 2, and the score decreases monotonously towards -2 and 5 (**Figure 10/D**).

In the case of the polar surface area (at pH = 7.4) we were more permissive than Rule of Three ( $PSA \leq 60 \text{ \AA}^2$ ) and accepted a wider range of ligands with PSA between 0 and  $100 \text{ \AA}^2$ . The desirability function is a hump function with a maximum at  $40 \text{ \AA}^2$  and linearly decreasing towards 0 and 100 (**Figure 10/E**).

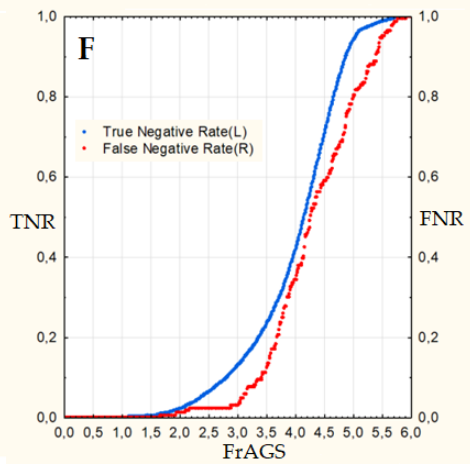
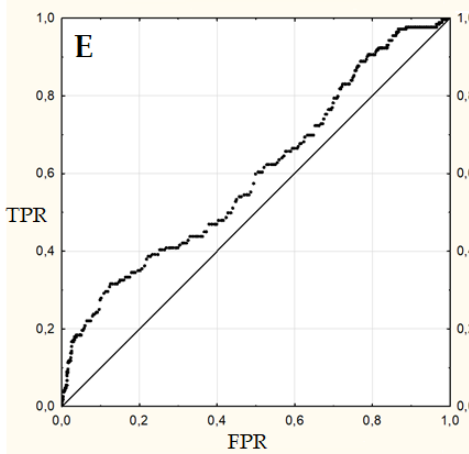
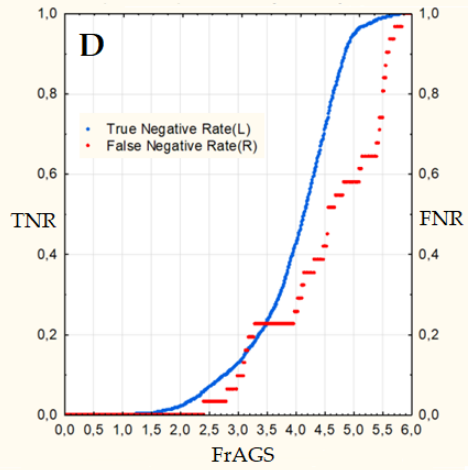
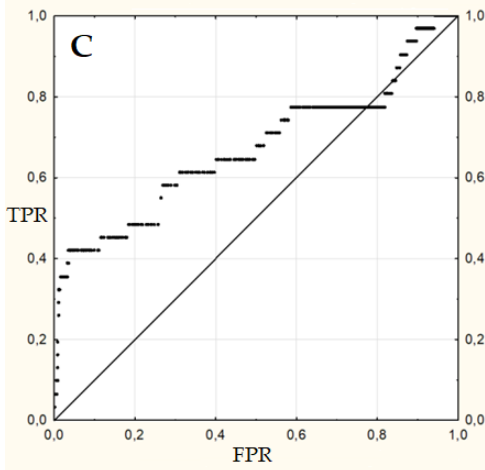
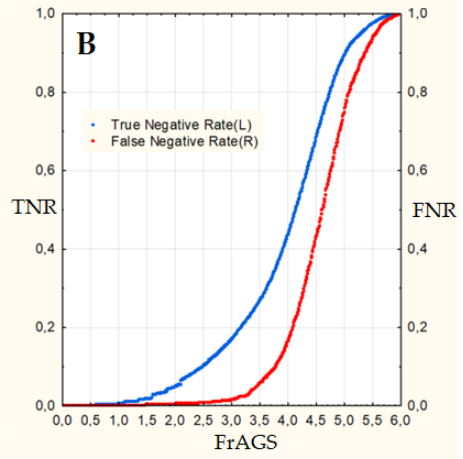
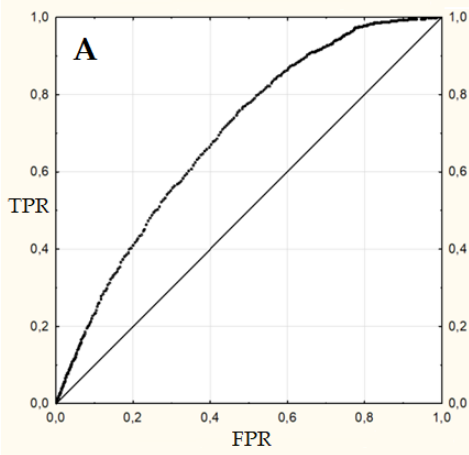
Although 83 % of the fragments of the training set are basic, the desirability function of the strongest basic pKa was mapped with a constant score of 0.1 in the 0-6 pKa range, and with a hump function covering the values of 6 and 13, penalizing fragments with basicity stronger than  $pK_a > 13$ , and preferring fragments with a pKa value between 8 and 10 (**Figure 10/F**).

Desirability scores may be combined either by summation or by multiplication into an overall desirability measure. A disadvantage of the multiplicative approach is that it discards molecules if they receive a desirability score of zero for a single property. In contrast, the additive approach reduces the impact of a single property-score when a large number of properties are involved<sup>147</sup>. Taking these aspects into consideration we defined the “Fragment Aminergic GPCR Score” (FrAGS) as the sum of the individual property-scores, providing a score with a range between 0 and 6.

### 2.1.2. Retrospective validation and evaluation

The validation of the desirability score (FrAGS) was carried out with three complementary approaches. As a first attempt, the active training set was mixed with an order of magnitude larger random reference ChEMBL set to check whether the score is able to enrich aminergic ligands (see details of the dataset assembly in **Section 3.1.1.2., page 83**). The second approach used an independent active set from the PubChem database<sup>152</sup> (see details of the PubChem dataset collection in **Section 3.1.1.3., page 83**). The third validation used data of a high-throughput screening (HTS) and fragment screening (FS) campaigns on aminergic GPCR targets provided by Gedeon Richter Plc. The performance of the FrAGS was measured by the calculation of Enrichment Factors (EF), Receiver Operating Characteristic (ROC)<sup>153</sup> curves, and by the comparison of False Negative Rates (FNR) and True Negative Rates (TNR) as functions of the FrAGS score.

The corresponding ROC curve of the ChEMBL validation (2183 actives/96539 random reference) is shown in **Figure 11/A** (True negative rates and false negative rates as functions of the FrAGS are shown in **Figure 11/B**). In case of the PubChem retrospective validation, FrAGS values were calculated for the 3131 5HT<sub>1</sub>R set of fragments and 17069 TAAR<sub>1</sub> set of fragments, followed by the calculation of TPR, FPR, TNR, FNR, ROC (**Figures 11/C, 11/D and 11/E, 11/F** for 5HT<sub>1</sub>R and TAAR<sub>1</sub>, respectively) and EF (**Figures 13/A-13/E**).

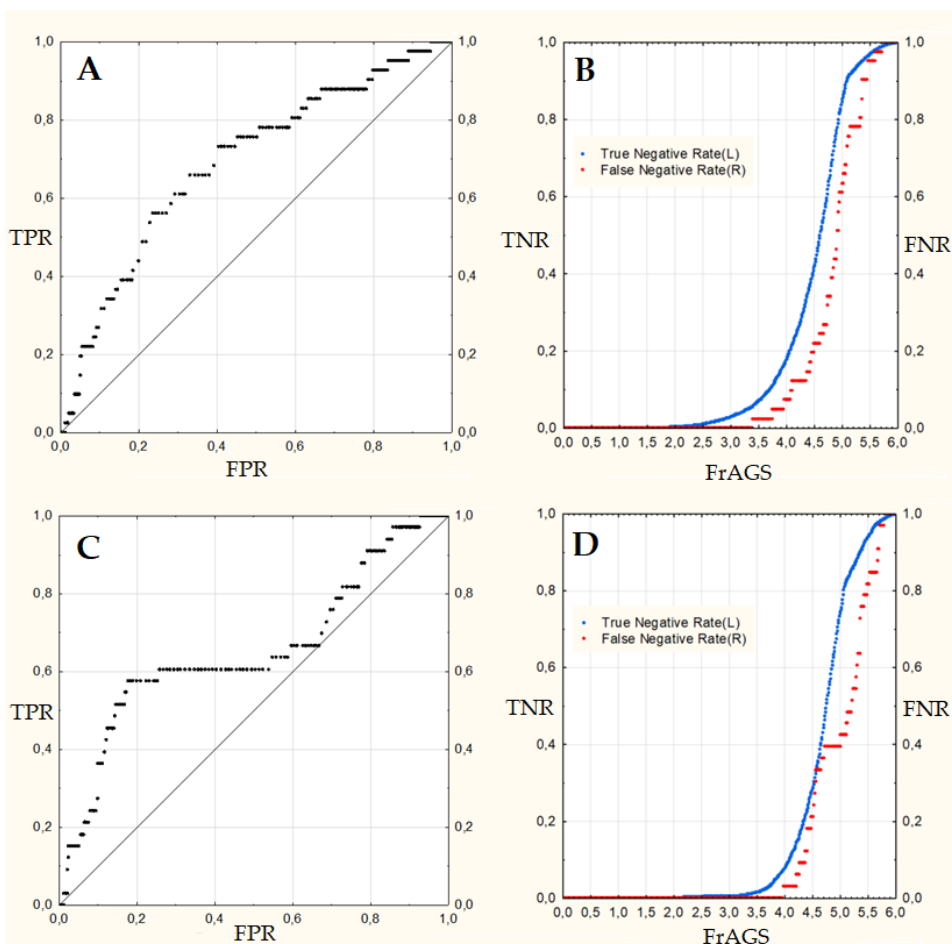


**Figure 11/A** - ROC curve of the ChEMBL validation data (2183 actives/96539 random inactives); **Figure 11/B** - True negative (TNR) and false negative rates (FNR) obtained on the ChEMBL validation dataset. **Figure 11/C** - ROC curve of the PubChem validation data on 5HT<sub>1R</sub>; **Figure 11/D** - True negative (TNR) and false negative rates (FNR) obtained with the PubChem 5-HT<sub>1R</sub> validation dataset; **Figure 11/E** - ROC curve of the PubChem validation data on TAAR<sub>1</sub>; **Figure 11/F** - True negative (TNR) and false negative rates (FNR) obtained with the PubChem TAAR<sub>1</sub> validation dataset

### 2.1.3. Prospective validation

The third validation was carried out on experimental screening data of fragments provided by Gedeon Richter Plc. The targets of the screening campaigns were undisclosed aminergic GPCR receptors. Only fragment sized molecules were examined from the HTS campaign. 9302 fragment size compounds were screened by high-throughput screening, and 3038 fragments were screened with fragment-screening (FS) approach.

In the high-throughput screening dataset 41 fragments were found to be active on an aminergic GPCR target, and 9261 fragments were inactive. Fragment-screening identified 33 active, and 3038 inactive fragments. The ROC curve and the corresponding false negative rates and true negative rates are shown in **Figure 12/A** and **12/B** for the HTS set, and in **Figure 12/C** and **12/D** for the FS set.



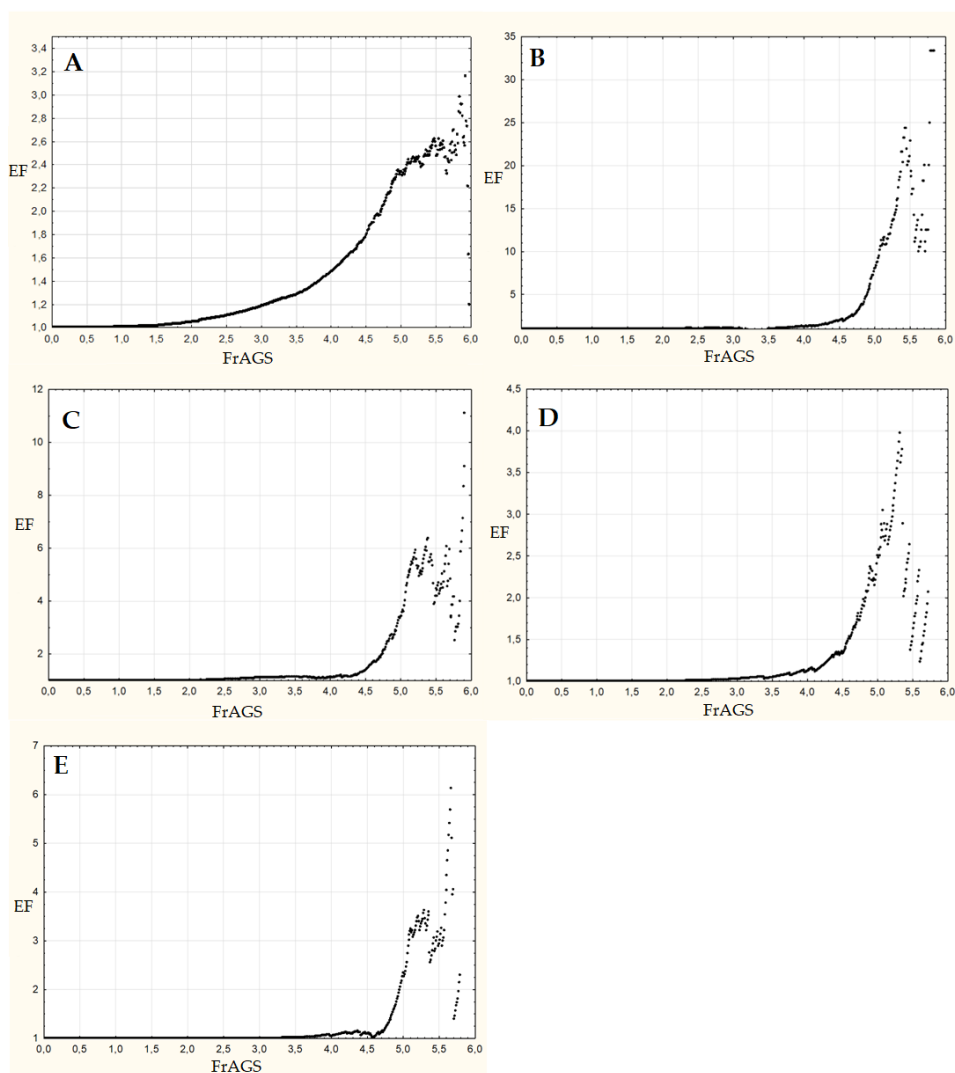
**Figure 12/A** - ROC curve of the HTS validation data; **Figure 12/B** – True negative and false negative rates obtained on the FrAGS - HTS dataset; **Figure 12/C** - ROC curve of the fragment screening (FS) validation data; **Figure 12/D** - True negative and false negative rates obtained on the FrAGS - FS validation dataset

#### 2.1.4. Defining the FrAGS cut-off

Ligand-based filters are typically used for selecting a set of compounds of a large library for further resource intensive computational or experimental studies. For the determination of the appropriate cut-off, the enrichment curves of the validation studies were examined. The plots of the enrichment factors versus FrAGS values are shown in **Figures 13/A-E**.

The validation test using ChEMBL compounds showed, that the enrichment is about 2-3 fold over about FrAGS = 4.5 (**Figure 13/A**). Its' highest value before the oscillatory region is about 2.5. The enrichment factor increases monotonically up to the score of about 5 and oscillates for higher scores (**Figure**

10/A). The enrichments are much higher in the other two validation studies using PubChem (Figure 13/B for 5-HT<sub>1R</sub>, Figure 13/C for TAAR<sub>1</sub>) and the experimental HTS (Figure 13/D) and fragment screening data (Figure 13/E). In the case of the 5HT<sub>1R</sub> validation set the enrichment of true positives over false positives resulted a remarkable ratio of 7.98 ( $EF_{5.0} = 7.98$ ). The TAAR<sub>1</sub> EF curve shows a 3.41 fold enrichment at FrAGS = 5. The enrichment monotonously increases until FrAGS = 5, and oscillates above 5. In case of the HTS validation data the enrichment factor increases monotonically until FrAGS of about 3, and oscillates over this value. Apparently the Fragment Screening validation set also showed that the enrichment factor increases monotonically until about FrAGS = 5.



**Figure 13/A** – Function of the Enrichment Factor against FrAGS – ChEMBL validation set; **Figure 13/B** – Function of the Enrichment Factor against FrAGS – PubChem 5-HT<sub>1R</sub>; **Figure 13/C** – Function of the Enrichment Factor against FrAGS – PubChem TAAR<sub>1</sub>; **Figure 13/D** – Function of the Enrichment Factor against FrAGS – HTS validation data; **Figure 13/E** - Function of the Enrichment Factor against FrAGS – FS validation data

Significant oscillation of the enrichment was observed over FrAGS > 5 that is due to the small number of compounds having high score values. It is proposed that a cut off value of 5.0 is an appropriate choice for the selection of promising aminergic fragments. Our three independent validation studies show enrichments around or above 3 using this cut off value.

FrAGS	EF ChEMBL	EF 5HT <sub>1</sub>	EF TAAR <sub>1</sub>	EF HTS	EF FS
4.5	1.79	2.00	1.39	1.33	1.09
5	2.35	7.98	3.41	2.51	2.34
5.5	2.56	22.86	4.48	1.54	2.94

**Table 6** - Enrichment factors at selected FrAGS values obtained in validation studies

In summary, comparative analysis of the distributions of several physicochemical descriptors for aminergic GPCR ligands and random reference compounds revealed that the following six descriptors are markedly different for the two sets of compounds: **logD (at pH = 7.4), PSA (at pH = 7.4), pK<sub>a</sub> (strongest basic center), number of nitrogen atoms, number of oxygen atoms, and the number of rotatable bonds**. Desirability functions based on the statistically significant differences in the property distributions were defined and were combined to a Fragment Aminergic GPCR Score (FrAGS). This score can take values of 0-6 (continuous variable). The higher the score the more likely the fragment is an aminergic GPCR ligand. The scoring scheme was validated with independent experimental data and it was found that a score cut-off value of 5.0 is appropriate to achieve an enrichment of around, or over 3. FrAGS is suitable option for screening large libraries of fragment-like compounds for aminergic GPCR ligands, and thus it might be a useful tool for compiling focused fragment libraries for drug discovery projects.

## 2.2. Consensus-scoring based docking method for the design of aminergic fragment libraries<sup>154</sup>

Since the first GPCR structure of rhodopsin was solved in 2000<sup>63</sup>, an ever growing number of class-A aminergic GPCR receptor X-ray structures have been published<sup>155</sup> (**1.2.1. Section - Table 1, page 14**). The large number of available aminergic GPCR structures raises the possibility for using their

combinations in docking experiments to identify and assemble aminergic fragments and fragment libraries. In addition to the predicted binding affinity such a structure-based method would provide predictions on binding modes that might support further fragment optimizations<sup>156</sup>.

Multi-target docking is a well-known approach that is capable of achieving higher enrichments of diverse actives compared to single-target docking<sup>157,158</sup>. Considering the similar characteristics of aminergic orthosteric binding sites we propose, that a representative collection of aminergic GPCR structures could provide hits that may also fit to aminergic receptors not included in the screening model. (For this reason, prospective validation of our method will be presented on a receptor with no available X-ray structure; see **Section 2.2.4., page 49**) However, the use of multiple structural data results in increased computational time and large volumes of data for processing<sup>159</sup>. Consequently the nineteen aminergic structures available at the date of the study (**Table 1**, structures with bold PDB ID codes were considered in the study) were clustered by the RMSD values of their binding site residues. In parallel were all subjected to self-docking in order to select an adequate number of representative structures suitable for docking based virtual screening.

## 2.2.1. Consensus ranking

### 2.2.1.1. Selection of structures for docking using clustering and self-docking

After collecting all available aminergic ligand-receptor complexes (description in Computational Methods **Section 3.1.2.1., page 84**) they were investigated in various ways to select a subset of representative structures. Self-docking was performed and the RMSD to their co-crystallized ligands was calculated together with the number of non-covalent interactions recovered. Moreover, conformational clustering of the receptor binding sites was performed. From each resulting cluster the structure best performing in self-docking studies was selected, thus picking an adequate number of structures for the optimization of the scoring method. The resulting self-docking scores, ligand RMSD values and the corresponding cluster trees are shown in Computational Methods **Section 3.1.2.2., page 84**. The first selection criterion was defined by best ligand RMSD value, with a cut-off 0.5 Å followed by the preferable Glide score<sup>160</sup> as second criteria. Altogether nine aminergic receptor structures (PDB codes: 3PBL, 3RZE, 4IB4, 4IAQ, 3UON, 4MQT, 4LDE, 2RH1, 3NY9) were used in the optimization of the screening protocol.

### 2.2.1.2. Optimization of the protein ensemble

The simultaneous docking of the training set (12,608 fragments) into the nine selected structures was performed, and a consensus scoring procedure was used for their evaluation. See details for the collection of the training set in Computational Methods **Section 3.1.2.3., page 87** and **Section 3.1.2.5., page 88** for the procedures of docking.

The FrACS scoring scheme was developed in the following way.

Dockings resulted in nine Glide Scores for each entry. In case of multiple generated states for a single compound, the Glide Scores of the best states were kept. Docking score results for standalone receptor structures were sorted by ascending Glide score values. Rank numbers were assigned to each entry according to its ranking, resulting in nine different ordinal numbers (ranks) for every single entry. These ranks were translated into votes using a ranking cut-off. The Rank(x)% criterion was defined as a percent of upper ranking boundary. A compound receives a vote number of 1 if its docking score is in the top (x) % and 0 otherwise. The resulting votes were summed into a voting score (FrACS, Fragment Aminergic Consensus Score) with a range between 0 and 9. The developed scoring scheme includes the following steps.

1. Docking into 9 structures, obtaining 9 Glide Scores for each compound, sorting compounds by Glide Scores for each protein structure,
2. Ranking docking scores for each structure,
3. Voting rank numbers for each compound by Rank(x)% criterion per each structure,
4. Calculating the sum of the votes (FrACS) for each compound,
5. Sorting by FrACS applying certain cut-offs (ranging between 0, 1, 2, ...9).

Retrospective validations were performed by calculating enrichment factors for a set of FrACS cut-off values. 95% confidence intervals for the enrichment factors were also calculated using an analytical error definition<sup>161,162</sup>. A benefit of applying multiple receptor structures is to achieve higher enrichment and diversity of actives relative to single structure protocols. Enrichments achieved with standalone receptor structures were also calculated as reference (see **Table Appendix 4**).

Enrichments (and corresponding statistical errors) calculated on Rank(x)% values of  $x = 1, 5, 10, 20$  and 30% and summed vote cut-off values of 1-9 are shown in **Table 7**. Top 6-7-fold enrichments were found at Rank(5)% criterion and sum-vote cut-off 5 to 8 (highlighted as bolded entries).

F	Rank(1)%			Rank(5)%			Rank(10)%			Rank(20)%			Rank(30)%		
	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)
9	0	0	n.a.	0	0	n.a.	15	10	5.5 (2.6E-02)	62	41	5.4 (9.5E-03)	138	81	4.8 (6.4E-03)
8	0	0	n.a.	7	6	<b>7.1 (1.5E-02)</b>	37	28	6.2 (1.2E-02)	164	102	5.1 (5.5E-03)	361	200	4.6 (3.5E-03)
7	0	0	n.a.	<b>19</b>	<b>14</b>	<b>6.0 (2.0E-02)</b>	86	56	5.4 (7.9E-03)	308	169	4.5 (3.9E-03)	618	288	3.8 (2.6E-03)
6	0	0	n.a.	<b>41</b>	<b>29</b>	<b>5.8 (1.1E-02)</b>	161	97	5.0 (5.7E-03)	481	237	4.1 (3.1E-03)	944	362	3.2 (1.9E-03)
5	1	1	4.1 (n.a.)	<b>86</b>	<b>57</b>	<b>5.4 (7.7E-03)</b>	273	152	4.6 (4.2E-03)	731	312	3.5 (2.3E-03)	1376	426	2.6 (1.4E-03)
4	13	9	5.7 (2.9E-02)	158	98	5.1 (5.6E-03)	421	207	4.0 (3.4E-03)	1089	390	2.9 (1.7E-03)	1875	492	2.2 (1.0E-03)
3	30	20	5.5 (1.5E-02)	291	155	4.4 (4.2E-03)	693	291	3.5 (2.5E-03)	1616	465	2.4 (1.2E-03)	2470	572	1.9 (6.7E-04)
2	88	51	4.8 (8.6E-03)	591	255	3.6 (2.8E-03)	1198	412	2.8 (1.6E-03)	2331	571	2.0 (7.0E-04)	3245	634	1.6 (3.9E-04)
1	368	159	3.6 (4.0E-03)	1354	435	2.6 (1.4E-03)	2215	572	2.1 (7.3E-04)	3415	651	1.6 (3.2E-04)	4285	679	1.3 (1.5E-04)

**Table 7.** Enrichments achieved by Fragment Aminergic Consensus Score (FrACS) on the training set - F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval.

The performance of the consensus ranking protocol was assessed by comparing enrichments achieved with standalone structures and by using all nine structures together. The comparisons were made at the same top rank percentages of 1, 5, 10, 20 and 30%. **Table Appendix 4** summarizes the enrichments of the nine individual structures. The histamine H<sub>1</sub>R (3RZE)<sup>163</sup> and the muscarinic mAChR<sub>2</sub> (4MQT)<sup>164</sup> structures achieved the best enrichments for the training set highlighted in the **Table Appendix 4**. While two structures (3PBL<sup>165</sup> and 4LDE<sup>166</sup>) showed only moderate enrichments, seven structures demonstrated 2 to 5-fold enrichment factors. **Table 7** shows the corresponding enrichments by certain sum-vote cut-offs and Rank(x)% criteria for the consensus score. In most cases, we found that the consensus method yields better enrichments as compared to individual structures, in particular at FrACS  $\geq 5$  cut-off and Rank(5)% criterion. This result suggests that the combination of multiple receptor structures may provide promising ligands with higher diversity. The highlighted rows in **Table 7** show that within a range of about 20 to 90 selected compounds at a Rank(5)% criterion, about 10 to 60 actives are recovered yielding remarkable enrichments between 5.4 and 7.1.

## 2.2.2. Retrospective validation of FrACS

The same two screening datasets used in **Section 2.1.2., page 35**, targeting aminergic receptors were downloaded from the PubChem public repository<sup>152</sup>. The sets were docked to the ensemble of aminergic receptors, and FrACS consensus ranking scores were calculated for both validation sets (**Table 8**) (see procedures in **Section 3.1.2.6., page 88**).

5HT <sub>1</sub> R				TAAR <sub>1</sub>			
Rank(1)%				Rank(5)%			
F	S	A	EF (95% CI)	F	S	A	EF (95% CI)
9	<b>3</b>	<b>1</b>	<b>33.7 (1.78)</b>	9	7	0	0.0 (n.a.)
8	<b>12</b>	<b>1</b>	<b>8.4 (0.3)</b>	8	<b>32</b>	<b>2</b>	<b>6.3 (6.6E-02)</b>
7	<b>30</b>	<b>5</b>	<b>16.8 (0.2)</b>	7	<b>110</b>	<b>6</b>	<b>5.5 (3.2E-02)</b>
6	<b>57</b>	<b>6</b>	<b>10.6 (0.1)</b>	6	246	9	3.7 (1.8E-02)
5	<b>111</b>	<b>9</b>	<b>8.2 (7.5E-02)</b>	5	475	16	3.4 (1.2E-02)
4	198	12	6.1 (4.6E-02)	4	886	25	2.9 (7.7E-03)
3	358	14	4.0 (2.7E-02)	3	1723	43	2.5 (4.9E-03)
2	660	15	2.3 (1.5E-02)	2	3327	64	1.9 (2.9E-03)
1	1280	20	1.6 (7.3E-02)	1	6765	98	1.5 (1.4E-03)

**Table 8** - FrACS enrichments achieved on PubChem validation sets - F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval (see full content in **Table Appendix 6 and 7**)

The performance of the standalone structures for the 5HT<sub>1R</sub> screening set is shown in **Table Appendix 5**. In this case the 5HT<sub>2BR</sub> structure (4IB4) highlighted in the table yielded the highest enrichment factors (7.7 to 16.3) at Rank(5)%. The FrACS score achieved similar enrichments (8.2 to 16.8) at Rank(1)% and Rank(5)%, as shown in **Table 8**. In case of the trace-amine receptor HTS dataset the best enrichments (2.5 to 3.5) were achieved with the 5HT<sub>1BR</sub> structure (4IAQ) highlighted in **Table Appendix 5**. Conversely, the FrACS method, where an ensemble of structures was used, showed improved enrichments (**Table 8**) at the same RANK(x)% selections around Rank(5)%.

### 2.2.3. Combination of FrAGS with FrACS

The previous paragraphs showed the utility of the two complementary approaches in parallel, providing true positives with notable enrichments. Proposing the consecutive application of ligand- and structure-based methods we screened fragment-sized molecules from ChEMBL (version 16) database<sup>117</sup> using the ligand-based desirability score FrAGS followed by the structure-based docking protocol FrACS. We have calculated the FrAGS scores for a ChEMBL fragment set containing all fragments (327419 molecules, 2183 actives included). The dataset was scored and filtered by FrAGS  $\geq 5$  cut-off, with a 6.7-fold enrichment of known actives (detailed description in **Section 3.1.2.7., page 88**). The resulting pre-filtered set was docked to the ensemble of receptors, and was scored by FrACS. The resulting votes were summed up, and enrichments - shown in **Table 9** - were examined at different FrACS cut-offs (1-9).

		Rank(30)%	
F	S	A	EF (95% CI)
9	456	56	18.4 [2.5 (3.6E-03)]
8	1538	145	14.1 [1.9 (1.7E-03)]
7	3137	269	12.9 [1.7 (1.1E-03)]
6	5251	382	10.9 [1.5 (7.3E-04)]
5	7657	524	10.3 [1.4 (5.4E-04)]
4	10453	653	9.4 [1.3 (4.1E-04)]
3	13673	774	8.5 [1.2 (3.1E-04)]
2	17360	909	7.9

			[1.1 (2.2E-04)]
1	21899	1165	8.0 [1.1 (1.1E-04)]

**Table 9.** Sequential screening enrichments achieved by the consecutive application of FrAGS and FrACS on the ChEMBL fragment set. - FrAGS pre-filtering step EF = 6.7 (4.1E-04) (95% Confidence Interval); F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval; Best enrichments highlighted as bold entires. (see full content in **Table Appendix 8**)

In a realistic drug design scenario, sequential virtual screening may start with similarly sized datasets consisting of about  $\sim 10^5$  fragments as it was the case for our ChEMBL fragment training set (327.419 entries). The FrAGS score may serve as a pre-filtering step with low computational cost to select compounds with favourable physicochemical properties for class A aminergic GPCR affinity. The structure-based FrACS method can be applied as a second screening step to further enrich promising actives for experimental testing. Our studies on the ChEMBL dataset showed 8 to 18-fold total enrichments in case of selecting 4-500 fragments. Enrichments increased with increasing Rank(x)% criterion showing dependence on the size of the dataset. Furthermore, the pre-selected fragments already satisfy main features for aminergic promiscuity, thus more permissive Rank(x)% criterion should be applied. Our results showed the best enrichments of the known actives at Rank(30)% (**Table 9**).

The prospective validation was carried out on our in-house fragment library (Library assembly is described in **Section 3.1.2.4., page 88**) by first calculating the FrAGS desirability scores, and keeping fragments with a score higher than FrAGS  $\geq 5$ . The resulting 211 pre-filtered fragments are expected to be enriched in promiscuous aminergic fragments according to their physicochemical features. Filtered fragments were subjected to Ligand Preparation in the Schrödinger suite<sup>167</sup> with default settings, and the resulting 2126 isomers (tautomers and protomers) were docked into the nine aminergic receptor structures. The docked fragments were ranked by Glide Scores<sup>160</sup>, and 9 votes were assigned to each fragment at Rank(30)% value. As described above, the sum-vote cut-off may be defined according to the experimental screening capacity. We selected fragments at sum-vote cut-off  $\geq 6$ . The 36 virtual screening hits were subjected to antagonist activity measurements in a cell-based assay against 5-HT<sub>6</sub>R Receptor. We have identified 8 *in vitro* hits with remarkable inhibitory levels (biological activities are summarized in **Table 10**), two of them appeared to be PAINS frequent hitters<sup>168</sup> (phenolic Mannich-bases<sup>169</sup>). The remaining six fragment hits corresponding to a hit rate of 16.7 % (6 hits/36 screened) possess the pharmacophoric features of typical aminergic compounds and were not described previously as 5-HT<sub>6</sub>R receptor ligands.

Cmpd	Structure	FrACS	FrAGS	Inhibitory Level (%) at 5-HT6R	IC <sub>50</sub> (μM) at 5-HT6R	IC <sub>50</sub> (μM) at CB <sub>1</sub> R
33		8	5.82	115.0 ± 1.0	-	
34		8	5.69	101.1 ± 9.7	-	
35		8	5.39	87.7 ± 7.5	-	
36		7	5.71	102.8 ± 8.5	3.9 ± 0.5	
37		7	5.40	105.0 ± 9.2	3.9 ± 1.9	17.0 ± 9.1
38		6	5.81	89.3 ± 5.1	-	
39		6	5.72	101.1 ± 0.3	2.9 ± 0.3	28.3 ± 15.3
40		6	5.46	102.0 ± 7.9	0.1 ± 0.01	0.0 ± 22.2

**Table 10.** Antagonist efficacies (% of inhibition) and IC<sub>50</sub> values of FrACS hits against 5HT<sub>6</sub>R receptor. - 5HT<sub>6</sub>R serotonin receptor inhibition assay; Antagonist efficacy was measured at 50 μM final concentration; 100%: Same inhibition with the control antagonist, 0%: No inhibition (agonist only), Negative control: 1% DMSO, Positive control: 10 μM SB-271046; table sorted by FrACS scores.

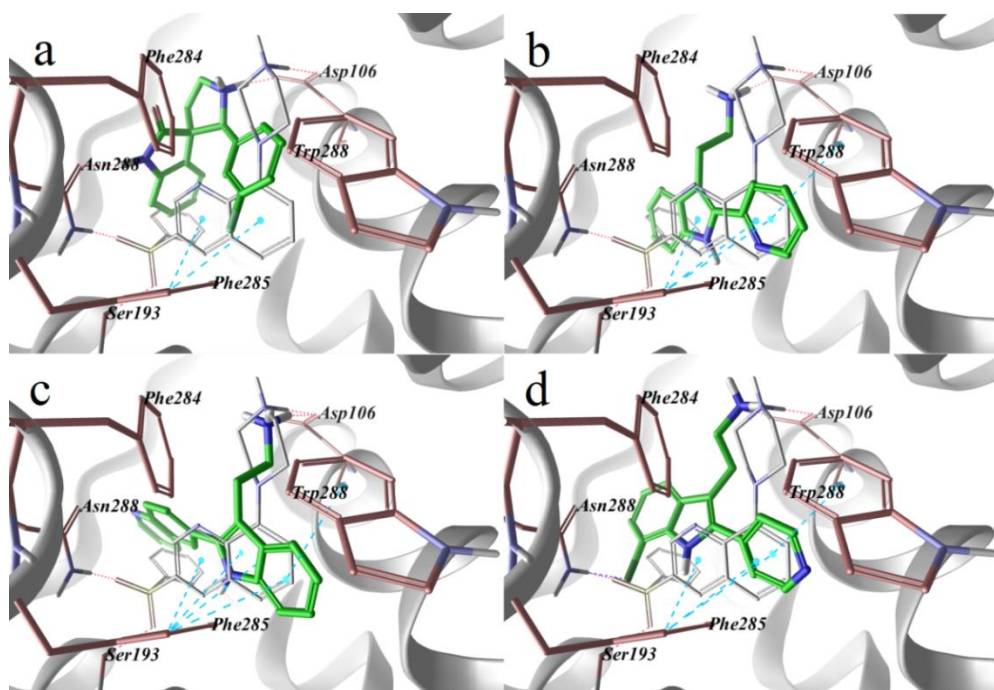
We selected **compounds 36, 37, 39 and 40** for IC<sub>50</sub> measurements (highlighted structures in blue in **Table 10**). We have found that our virtual screening fragment hits show low micromolar IC<sub>50</sub> values on 5-HT<sub>6</sub>R receptor.

We have accomplished a counter-screen for **compounds 37, 39 and 40** against the non-aminergic target, cannabinoid receptor subtype-1 (CB<sub>1</sub>R) in order to check specificity of compounds towards the aminergic subfamily. In fact, all of the hits were found to be inactive against CB<sub>1</sub>R (**Table 10**).

The prospective validation of our ligand and structure based methods showed that the similar characteristics of the aminergic binding pockets allows the identification of fragment hits active on an aminergic receptor not included in our predictive model. This finding underpins that FrACS based virtual screening either itself or in combination with FrAGS is an efficient tool for the design of focused libraries for class A aminergic GPCRs and can support the identification of fragment ligands for these receptors.

#### **2.2.4. Binding mode analysis of active fragments in 5-HT<sub>6</sub>R homology models**

We have examined the four best fragment hits of the prospective validation in 5-HT<sub>6</sub>R conformations obtained from the molecular dynamics simulation of its homology model by comparing their binding modes to a canonical 5-HT<sub>6</sub>R antagonist SB-742457 (**8** in **Table Appendix 2**)<sup>170,171</sup> (see details of docking procedure in **Section 3.1.2.8., page 89**). The most important interaction of the reference ligand is the H-bond formed between the protonated basic amine-moiety in the piperazine ring and the highly-conserved Asp106<sup>3.32</sup>. Further two H-bonds are formed between the sulfonyl-oxygens of the ligand and Asn288<sup>6.55</sup> and Ser193<sup>5.43</sup>. The quinoline ring occupies a hydrophobic cavity (first hydrophobic pocket), surrounded by three aromatic residues (Trp281<sup>6.48</sup>, Phe284<sup>6.51</sup>, Phe285<sup>6.52</sup>). The second hydrophobic cavity is formed by Val107<sup>3.33</sup>, Ala157<sup>4.56</sup>, Leu160<sup>4.59</sup>, Pro161<sup>4.60</sup>, Leu164<sup>4.63</sup>, etc. around the phenyl-moiety of SB-742457 (**8** in **Figure 5** and **Table Appendix 2**). Four representative docking poses of our primary hits are shown in **Figure 14**.



**Figure 14** - Representative docking poses of compounds **36** (a), **37** (b), **39** (c) and **40** (d).

Twenty-eight out of the thirty-six docking poses were found in an orientation that is closely analogous to the ones of SB-742457 (**8**) in all the nine MD frames. Characteristic interactions are discussed here using the example frame annotated as “1001” in reference<sup>171</sup>. The “spiro” **compound 36** (2'-(3-fluorophenyl)spiro[indoline-3,3'-pyrrolidin]-2-one) interacts through the protonated secondary amine nitrogen of the pyrrolidine-ring in all frames. The benzene-ring of the indoline-scaffold occupies the same hydrophobic pocket, like the benzene-sulfonyl part of the reference ligand. The fluorophenyl-part of **compound 36** is oriented towards the same hydrophobic pocket (Phe284<sup>6.51</sup>, Phe285<sup>6.52</sup>, and Trp281<sup>6.48</sup>) occupied by the quinoline-ring of SB-742457 (**8** in **Figure 5** and **Table Appendix 2**). Interestingly, an additional H-bond was formed between the backbone carbonyl oxygen of Asn288<sup>6.55</sup> and the amide-nitrogen of the indoline-2-one scaffold in 6 out of 9 poses. **Compound 37** (2-(2-(pyridin-2-yl)-1H-indol-3-yl)ethanamine) forms an H-bond with the conserved aspartate (Asp106<sup>3.32</sup>) in all poses, and  $\pi$ - $\pi$  stacking interactions are formed between the pyridine-ring and Phe285<sup>6.52</sup>, Trp281<sup>6.48</sup> hydrophobic residues, furthermore, the benzene part of the indole-ring occupies the same hydrophobic cavity as the benzene-sulfonyl part of SB-742457 (**8**). The pose of the hit compound (2-(2-(quinolin-6-yl)-1H-indol-3-yl)ethanamine) (**39**) is in good agreement with the reference ligand, by forming the desired H-bond with

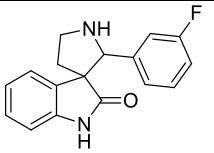
Asp106<sup>3,32</sup>, and both of the hydrophobic (first and second) subpockets are occupied by its quinoline and indole rings, respectively. In case of **compound 40** (2-(7-chloro-2-(pyridin-4-yl)-1H-indol-3-yl)ethanamine) an interesting halogen-bond is formed between the chlorine and the amino-moiety of Asn288<sup>6,55</sup> in six out of nine docking poses. The aromatic-stacking interactions and the formation of the H-bond with the aspartate are both well-represented.

## 2.3. Spiro-oxindoles, and -indolines as new 5-HT<sub>6</sub>R antagonists<sup>172</sup>

**Compound 36** (2'-(3-fluorophenyl)spiro[indoline-3,3'-pyrrolidin]-2-one) was identified as a fragment hit with micromolar inhibitory activity against 5-HT<sub>6</sub>R in the prospective validation of our combined FrAGS and FrACS approach. In this section we present a hit-to lead optimization performed by structure-activity relationship (SAR) analysis to explore the several derivatives of the spiro[pyrrolidine-3,3'-oxindole] scaffold. Throughout the process we introduced the phenylsulphonyl moiety and examined the corresponding reduced spiro[pyrrolidine-3,3'-indoline] scaffold.

### 2.3.1. Early SAR data on spiro[pyrrolidine-3,3'-oxindole] derivatives

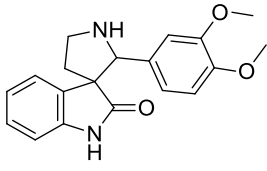
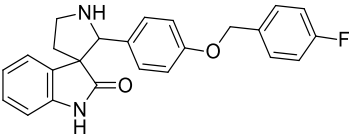
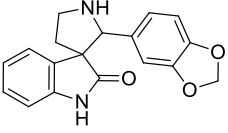
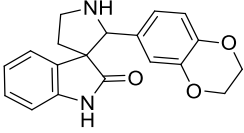
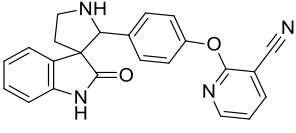
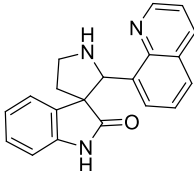
As an initial step, we have examined the affinities of **compound 36** across the closely related subtypes of serotonin receptors (5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, and 5-HT<sub>7</sub>R). The results of this GPCR-panel screen (**Table 11**) showed that a fairly non-selective fragment hit was in our hand. Thus, we aimed both to improve affinity and also selectivity of the initial fragment hit to reach the submicromolar affinity range with reasonable subtype-selectivity.

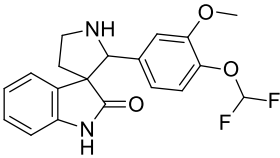
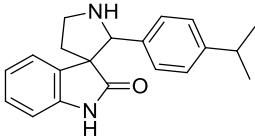
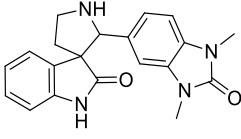
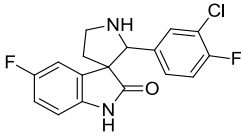
ID	Structure	5-HT <sub>1A</sub>	5-HT <sub>2A</sub>	5-HT <sub>6</sub>	5-HT <sub>7</sub>
36		16.91	5.05	6.75	8.48

**Table 11** – Serotonergic G-protein coupled receptor (GPCR) panel of **compound 36** as measured in binding assays of four serotonin receptors (measured values of K<sub>i</sub> (μM)); description of the assay in **Section 3.2.2., page 94**.

Next we have initiated a SAR-by-catalogue virtual screening. All purchasable screening compounds with the spiro[pyrrolidine-3,3'-oxindole] scaffold were collected from the Mcule database<sup>173,174</sup>. The collected set was docked (see details in **Section 3.1.3., page 89**) to the nine 5-HT<sub>6</sub>R MD frames<sup>171</sup>(see **Section 2.2.4., page 49**). Compounds (**41-50**) having the best

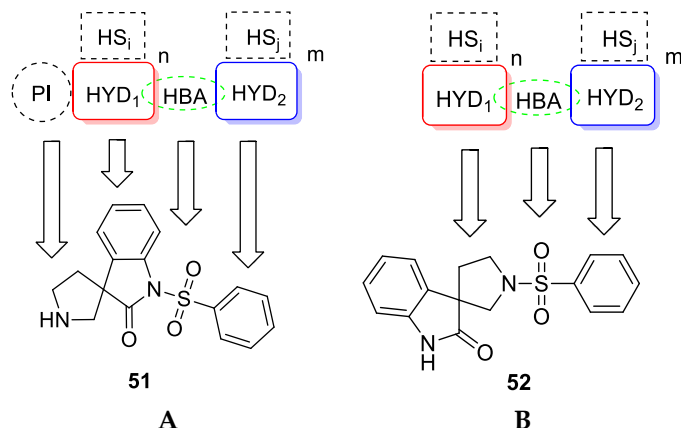
ranking in most of the 9 frames using the same consensus ranking as in **Section 2.2.1.2., page 42**<sup>154</sup> were purchased and tested in our serotonergic panel (**Table 12**) (see details of bioassay procedures in **Section 3.2.2., page 94**).

ID	Structure	5-HT <sub>1A</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>2A</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>6</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>7</sub> K <sub>i</sub> ( $\mu$ M)
41		7.07	5.45	<b>2.26</b>	2.69
42		1.74	2.28	<b>3.18</b>	1.09
43		9.84	2.69	<b>3.20</b>	8.82
44		12.08	4.33	<b>4.11</b>	12.74
45		49.11	8.68	<b>6.20</b>	58.96
46		45.67	1.12	<b>6.63</b>	10.45

47		7.14	5.77	7.73	7.49
48		7.76	4.54	8.62	8.98
49		40.50	23.70	9.49	46.72
50		not active	53.89	16.54	not active

**Table 12** - Serotonergic GPCR panel of derivatives substituted in the 2'-phenyl moiety as measured in binding assays of four serotonin receptors.

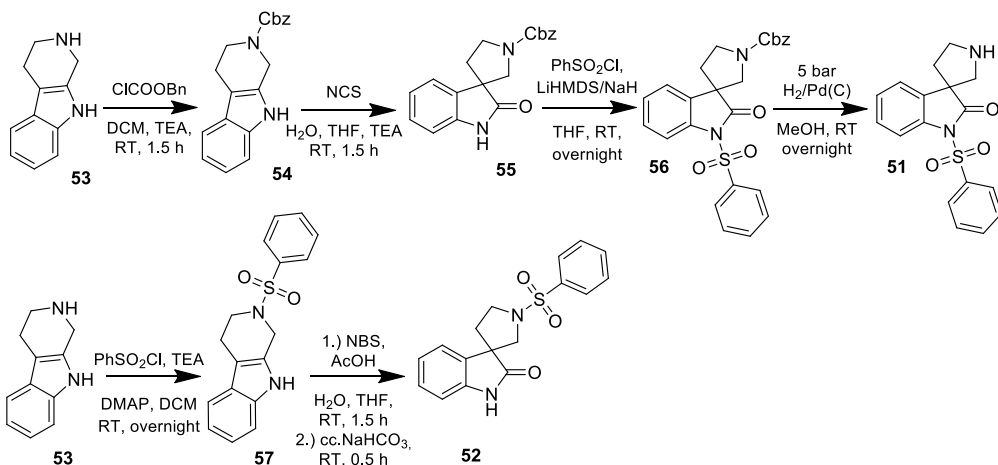
Compounds **41-49** were lacking substituents in the oxindole-core. Out of the collected set, three 3,4-dimethoxy- (**41**)/3,4-methylenedioxy- (**43**)/3,4-ethylenedioxy- (**44**) analogues showed minor improvements in affinity compared to **36** while 5-F analogue (**50**) was even weaker than that. These findings suggested to leave the oxindole core unsubstituted. Analysing the pharmacophore patterns (see **Figure 4** and **15**) of known antagonists<sup>64</sup>, however, suggested to introduce a phenylsulfonyl moiety either to the 1-nitrogen at the oxindole ring or to the 1'-nitrogen at the pyrrolidine ring in order to improve affinity and selectivity.



**Figure 15** – (A) 1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**51**) superimposed onto the classical pharmacophore; (B) 1'-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**52**) fitting to the pharmacophore pattern lacking a positive ionisable moiety.

### 2.3.2. Hit-to-lead optimization of spiro[pyrrolidine-3,3'-oxindoles]

The first synthetic route leading to the synthesis of **51** (Scheme 3) starts with the *N*-acylation of the tryptoline (**53**). The application of Cbz (carboxybenzyl) protecting group was necessary for two reasons. First, to achieve the sulfonylation at the oxindol-nitrogen, and second, the spiro-rearrangement reaction failed in case of unsubstituted pyrido-nitrogen. *N*-chlorosuccinimide<sup>175</sup> was used in the first experiments as halogenating reagent providing the Cbz-protected spiro[pyrrolidine-3,3'-oxindole]. Phenylsulfonylation was performed either by using lithium-hexamethyl disilazane<sup>176</sup> and sodium hydride<sup>177</sup> as deprotonating agents and acid scavengers. Finally, the deprotection of the Cbz-group by hydrogenation<sup>178</sup> resulted in the desired 1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**51**). Following an alternative way, sulfonylation of the more basic pyrido-nitrogen of tryptoline **53** might be performed as a first step, followed by the spiro-cyclization **52**, now by the use of *N*-bromosuccinimide<sup>120</sup>.



**Scheme 3** - Synthesis of 1- and 1'-phenylsulfonyl derivatives (**51**, **52**).

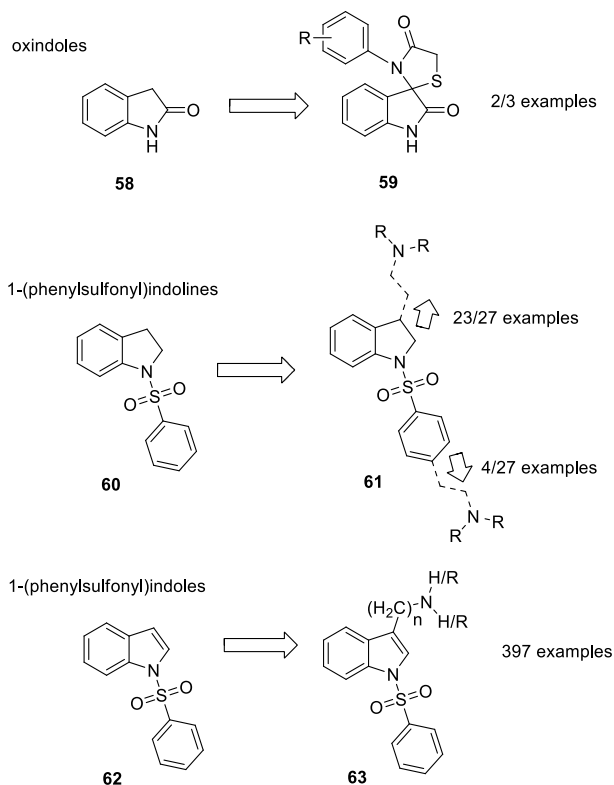
Compounds **51** and **52** contain phenylsulfonyl group instead of 2'-phenyl (compare with **41-50**). Although they still exhibited micromolar affinity towards 5-HT<sub>6</sub>R, they showed a promising trend of selectivity against the related subtypes (5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>7</sub>R).

ID	Structure	5-HT <sub>1A</sub> K <sub>i</sub> ( $\mu\text{M}$ )	5-HT <sub>2A</sub> K <sub>i</sub> ( $\mu\text{M}$ )	5-HT <sub>6</sub> K <sub>i</sub> ( $\mu\text{M}$ )	5-HT <sub>7</sub> K <sub>i</sub> ( $\mu\text{M}$ )
<b>51</b>		not active	not active	<b>12.30</b>	58.31
<b>52</b>		not measured	87.35	<b>6.86</b>	not active

**Table 13** - Serotonergic GPCR panel of the *N*-phenylsulfonylated derivatives at *N'* and *N*-positions (**51**, **52**) as measured in binding assays of four serotonin receptors.

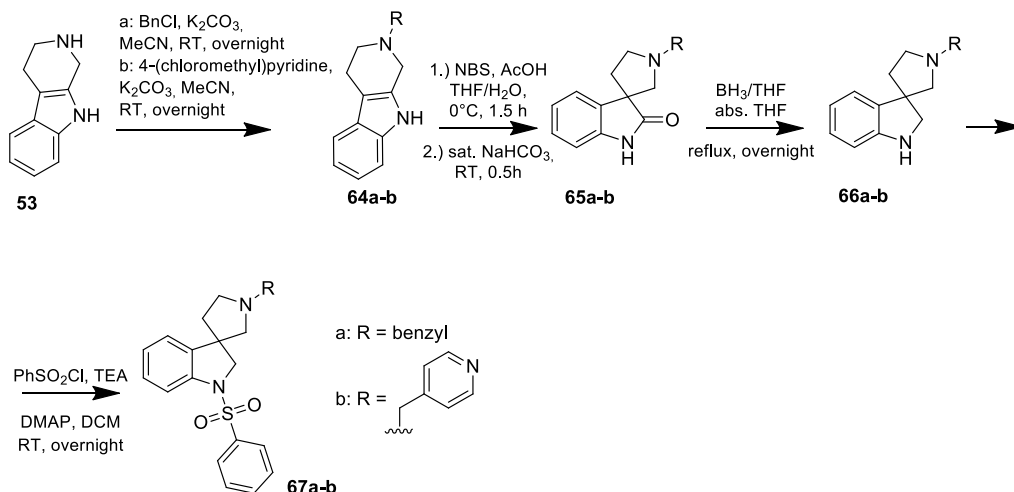
A feature not matching the pharmacophore pattern was the hydrogen bond acceptor oxo group in the oxindole core. For this purpose we analyzed the chemical space of 5-HT<sub>6</sub>R ligands with indole, indoline and oxindole substructures, having at least 0.1  $\mu$ M activity towards 5-HT<sub>6</sub>R from the ChEMBL database.

Substructure filtering by the query of **58** (**Figure 16**) revealed that altogether three oxindole-based 5-HT<sub>6</sub>R ligands are reported in the literature, all belonging to the 3'-phenylspiro[indoline-3,2'-thiazolidine]-2,4'-dione scaffold **59** that was presented by Hostetler *et al*<sup>179</sup>. When searching for the 1-phenylsulfonyl indoline substructure (**60**), a generalized pattern (**61**) was identified, representing 27 active ligands. Out of the 27 active molecules 23 were equipped with the classical tryptamine-like amino-ethylene moiety, and 4 examples contained the secondary amine moiety (positive ionizable) linked to the phenyl-side of the sulfonyl structure. The most abundant group identified contains the 1-(phenylsulfonyl)indole (**62**) substructure with 397 examples (**63**). Five compounds had a short one-methylene long linker between the amine (two primary amines, three tertiary amines) and the indole core, while the majority (51%) of the ligands possessed an ethylene linker (201 examples out of 397), out of which 123 (31%) were tertiary amines, 51 (13%) were secondary amines and 19 (5%) were primary amines. Furthermore 54 (14%) analogues contained a three-methylene long linker between the C-3 and the amino-group. Out of these 54 analogues 21 (5%) were tertiary amines and 33 (8%) were secondary amines and no primary amines were found. No ligands were found with a positive ionizable group farther than 3 atoms from the indole core.



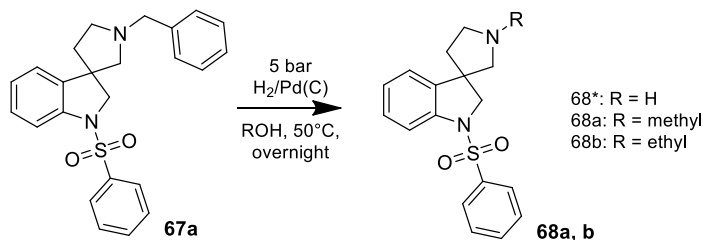
**Figure 16.** Substructure analysis of ChEMBL 5-HT<sub>A</sub>R active compounds.

As a conclusion of the analysis we designed compounds **67a-b** (Scheme 4), being substituted in the 1'-pyrrolidine-nitrogen with either basic-nitrogen containing, or lipophilic groups. First we benzylated<sup>180</sup> the tryptoline (**53**) to afford **64a** that was further converted to the corresponding spiro-derivative<sup>181</sup> **65a**. The reduction of the amide oxo-group leading to the indoline scaffold (**66a**) was performed by borane-tetrahydrofuran complex in refluxing absolutized tetrahydrofuran<sup>182</sup>. We observed that the reduction step has to precede the sulfonylation to avoid the formation of side products when treating the phenylsulfonamide-derivative with the boronic reagent. The phenyl-sulfonylation was performed either by using LiHMDS<sup>176</sup> or triethylamine/dimethylaminopyridine<sup>183</sup> reagents to yield **67a**. The pyridine-4-ylmethyl derivative (**67b**) was synthesized following the same procedure, however, as a first step, the tryptoline (**53**) was alkylated with 4-chloromethylpyridine<sup>184,185</sup>.



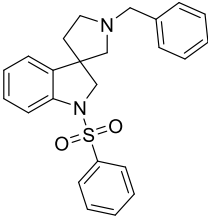
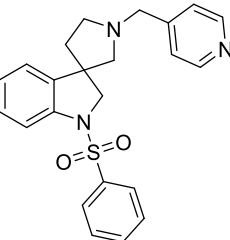
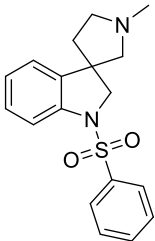
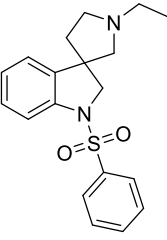
**Scheme 4.** Synthesis of *N'*-substituted and *N*-phenylsulfonylated indolines.

To examine the effect of omitting the oxo-group of **51** we planned to synthesize the indoline analogue **68\***, however the debenzoylation step of **67a** led to unexpected products (**Scheme 5**, **68a** and **68b**). When methanol was applied as solvent for the debenzoylation by catalytic hydrogenation, a methylated product **68a** has formed and the use of ethanol led to an ethyl-alkylated derivative **68b**.



**Scheme 5.** Synthesis of *N'*-methyl and *N'*-ethyl derivatives.

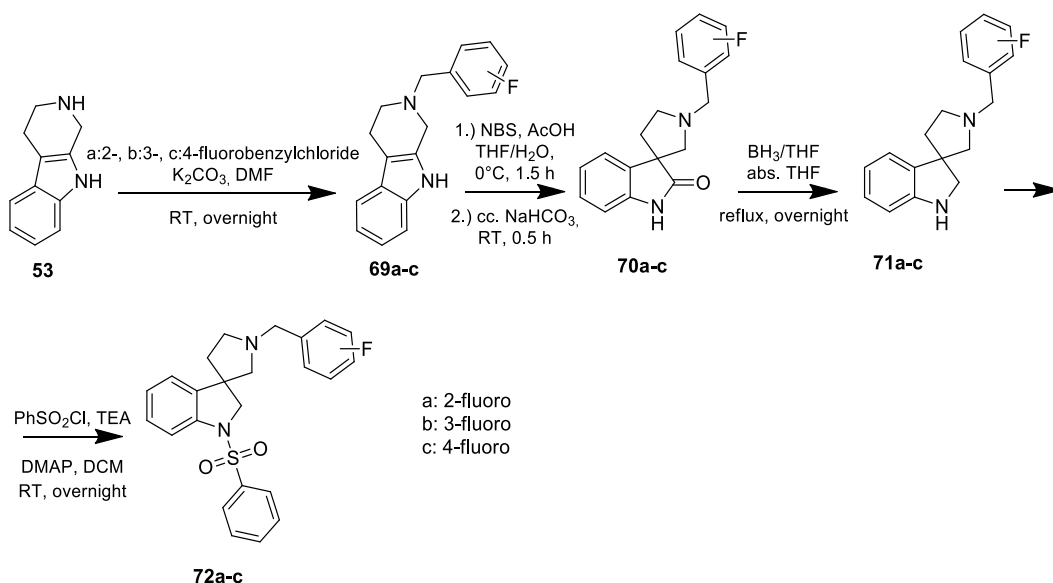
The *N'*-alkylated products (**68a-b**) showed low micromolar affinities towards 5-HT<sub>6</sub>R, and good selectivities against 5-HT<sub>7</sub>R, however no selectivity against 5-HT<sub>2A</sub>R. The *N'*-benzylated analogue (**67a**) reached the submicromolar affinity at 5-HT<sub>6</sub>R with more reasonable selectivities across the related subtypes (**Table 14**).

ID	Structure	5-HT <sub>1A</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>2A</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>6</sub> K <sub>i</sub> ( $\mu$ M)	5-HT <sub>7</sub> K <sub>i</sub> ( $\mu$ M)
67a		not measured	4.76	0.76	55.76
67b		not active	26.85	6.60	not active
68a		not measured	4.52	1.85	40.17
68b		not measured	3.82	2.42	not active

**Table 14.** Serotonergic GPCR panel of *N*-phenylsulfonyl indolines with different *N'*-substituents as measured in binding assays of four serotonin receptors.

Therefore we selected **67a** for further optimization. As a next step, we examined the substituent vectors at both the benzyl and the phenylsulfonyl rings by walking fluorine substituents around. This methodology, often referred to as fluoro-scan has been used effectively for the identification of sites tolerant to functionalization. In particular, fluoro-scan explores the impact of enhanced lipophilicity, H-bonding and/or filling a small pocket<sup>186</sup>. In case of the *N'*-benzylic substitution pattern, 2-, 3- and 4-fluorobenzylchlorides were used as

alkylating agents in order to synthesize the corresponding fluorinated indolines (Scheme 6, 72a-c) following the same synthetic steps as for 67a in Scheme 4.



Scheme 6 - Fluoro-scan of the *N'*-benzyl direction.

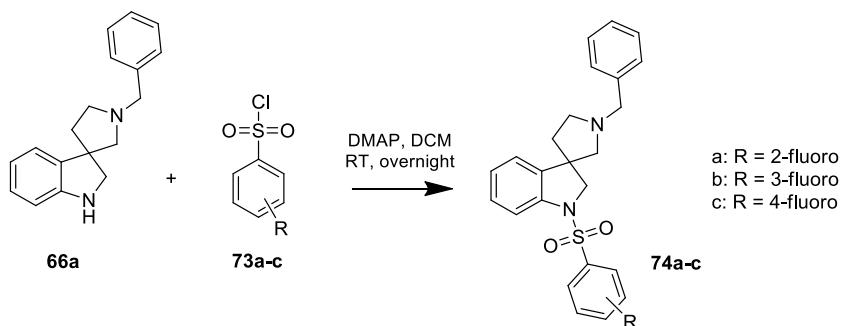
Substitution of the benzyl ring with fluorine in all the three (ortho-, meta-, para-, 72a-c respectively) positions caused a minor decrease in affinity still at the low micromolar level, and retaining selectivity against 5-HT<sub>1A</sub>R and 5-HT<sub>7</sub>R. We concluded, that growing towards this direction is not beneficial (Table 15).

ID	Structure	5-HT <sub>1A</sub> K <sub>i</sub> (μM)	5-HT <sub>6</sub> K <sub>i</sub> (μM)	5-HT <sub>7</sub> K <sub>i</sub> (μM)
72a		not active	4.53	78.93

72b		not active	1.97	31.49
72c		not active	1.77	15.74

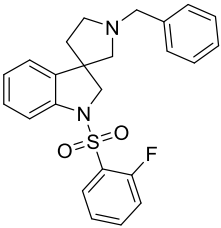
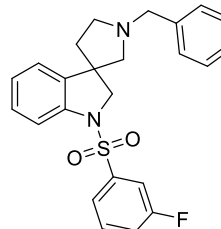
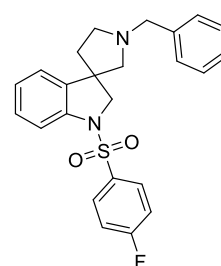
**Table 15** - Serotonergic GPCR panel of the fluorobenzyl derivatives as measured in binding assays of three serotonin receptors.

The fluorophenylsulfonyl derivatives with ortho-, meta-, para-substitutions (**74a–c**) were synthesized from compound **66a** (**Scheme 7**), previously prepared as an intermediate for the synthesis of **67a**.



**Scheme 7** - Fluoro-scan of the *N*-phenylsulfonyl direction.

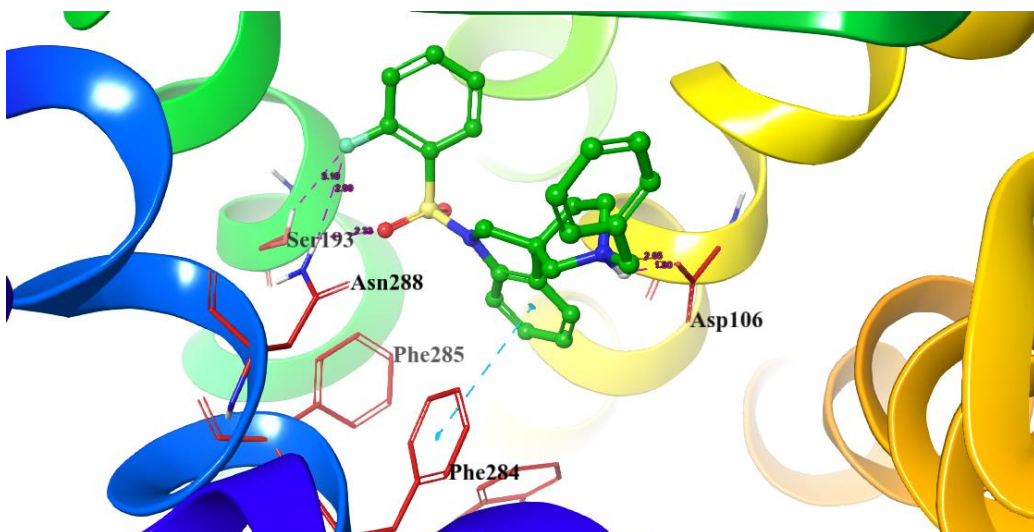
The alternative growing direction towards the phenylsulfonyl ring (compounds **74a–c**) showed improvement in the affinities (see **Table 16**), compared to the results of **67a**, also retaining the good selectivity against the 5-HT<sub>1A</sub>R and 5-HT<sub>7</sub>R subtypes (200 nM affinity towards 5-HT<sub>6</sub>R, and around 200-fold selectivity against 5-HT<sub>7</sub>R). These data suggest position-2 of the phenylsulfonyl group of **74a** can be beneficially substituted and therefore it is worth to explore this vector further in the lead optimization phase.

ID	Structure	5-HT <sub>1A</sub> K <sub>i</sub> (μM)	5-HT <sub>6</sub> K <sub>i</sub> (μM)	5-HT <sub>7</sub> K <sub>i</sub> (μM)
74a		not active	0.19	42.57
74b		not active	0.71	76.00
74c		not active	2.45	56.05

**Table 16** - Serotonergic GPCR panel of the fluorophenylsulfonyl derivatives as measured in binding assays of three serotonin receptors.

### 2.3.3. Binding mode analysis of the Optimized 5-HT<sub>6</sub>R Ligand (74a)

The improved affinity of **74a** was interpreted by docking the compound to the 5-HT<sub>6</sub>R receptor model reported earlier<sup>171</sup> and used in **Section 2.2.4., page 49**. Docking of **74a** by Schrödinger–Glide either with, or without applying any constraint (requiring hydrogen bonding with both Asn288<sup>6.55</sup> and/or Ser193<sup>5.43</sup>) resulted in unchanged, and consistent docking poses inside the orthosteric binding pocket (**Figure 17**).



**Figure 17.** Binding mode of compound **74a** in the orthosteric pocket of 5-HT<sub>6</sub>R

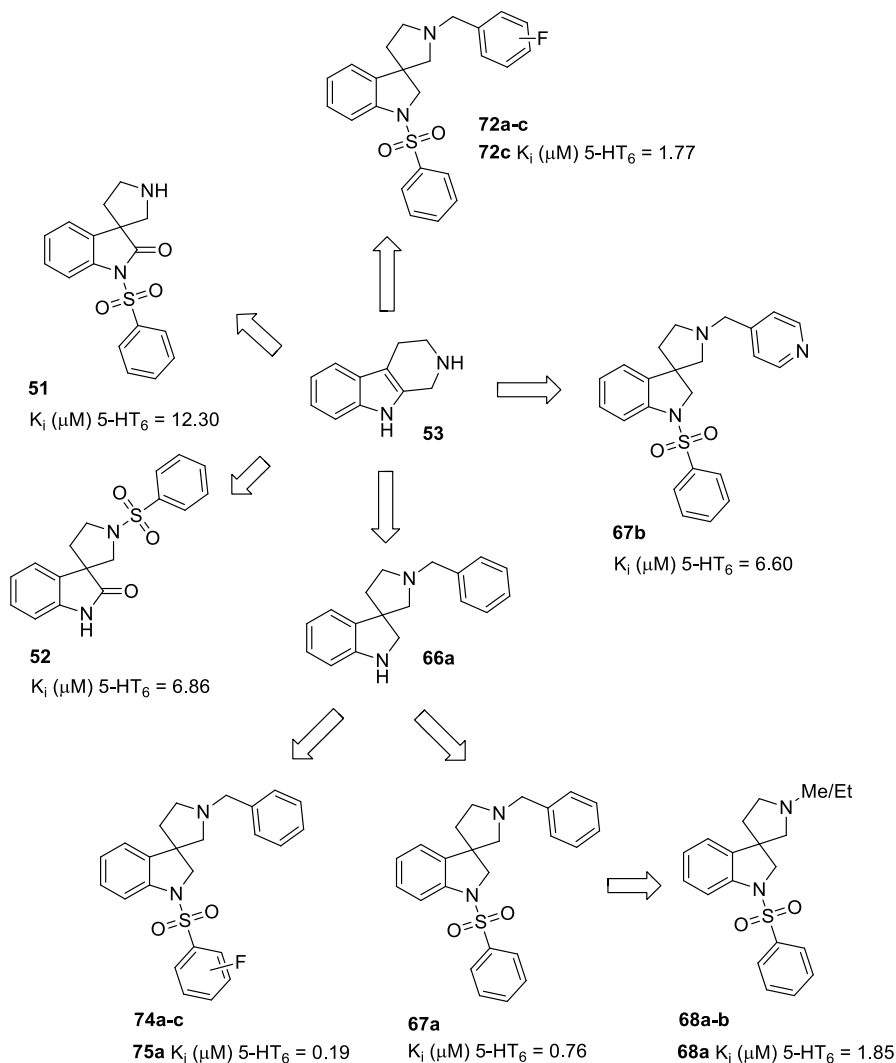
The indoline core (HYD1 in **Figure 4**) occupies the primary hydrophobic cavity of the binding pocket (formed by Trp281<sup>6.48</sup>, Phe284<sup>6.51</sup>, Phe285<sup>6.52</sup>), while the protonated quaternary pyrrolidine-nitrogen (PI in **Figure 4**) is oriented towards the conserved Asp106<sup>3.32</sup>, forming a well-established hydrogen bond. The phenyl-sulfonyl moiety (HYD2 in **Figure 4**) of **74a** is sitting at a secondary hydrophobic cleft defined by Val107<sup>3.33</sup>, Ala157<sup>4.56</sup>, Leu160<sup>4.59</sup>, Pro160<sup>4.60</sup> and Leu164<sup>4.63</sup>, positioning the sulfonyl-linker as hydrogen-bond acceptor (HBA in **Figure 4**) against Ser193<sup>5.43</sup> and Asn288<sup>6.55</sup>. Interestingly, the fluorine atom as hydrogen bond acceptor is offering a possible polar interaction (**Figure 4**) with the Asn288<sup>6.55</sup> residue, underlining the preference of the ortho-substitution at the phenyl-sulfonyl ring.

### 2.3.4. Conclusion

The structure-activity relationship around the original spiropyrrolidinyl-oxindole hit (**36**) was first investigated using the “SAR-by-catalog” approach that resulted in some moderate improvement in affinity and unacceptable selectivity. Therefore, we decided to explore the novel 5-HT<sub>6</sub>R chemotype by a more conventional medicinal chemistry strategy. A synthetic tree (summarized in **Scheme 8**) of oxindoles and indolines was elaborated starting from the core tryptoline intermediate (**53**).

The removal of the bulky lipophilic group at the 2'-position of the pyrrolidine ring and the insertion of the classical phenylsulfonyl moiety resulted in notable selectivity across the serotonergic GPCR panel of 4 receptors. Analysis of known indole, indoline and oxindole-based 5-HT<sub>6</sub>R inhibitors in

ChEMBL suggested to incorporate the phenylsulfonyl group. We therefore synthesized different 1-(phenylsulfonyl) indolines substituted at the 1'-pyrrolidine nitrogen with either lipophilic, or basic nitrogen containing groups. As a result, 1'-benzyl-1-(phenylsulfonyl) spiro[indoline-3,3'-pyrrolidine] (**67a**) and its corresponding 2-, 3-substituted analogues **74a-b** were identified as promising 5-HT<sub>6</sub>R ligands.

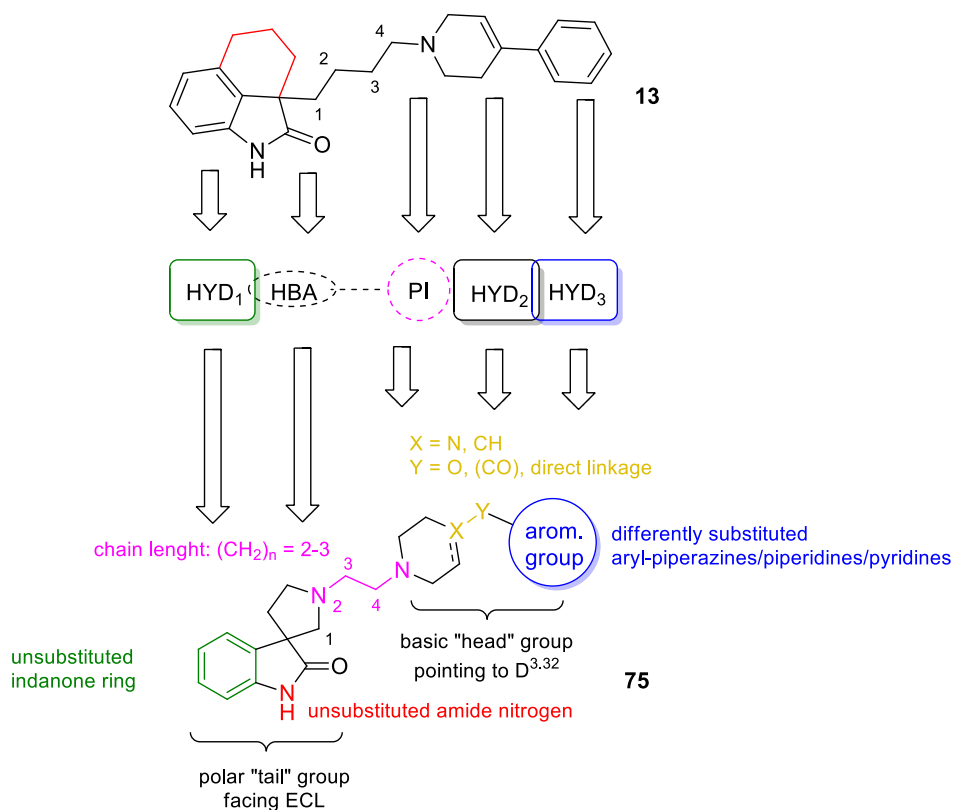


**Scheme 8** - Spiro[pyrrolidine-3,3'-oxindoles] and spiro[pyrrolidine-3,3'-indolines] as 5HT<sub>6</sub>R ligands.

## 2.4. Spiro-oxindoles as new 5-HT<sub>7</sub>R antagonist

### 2.4.1. Identifying novel 5-HT<sub>7</sub>R ligands with spiro-oxindole scaffold

As discussed in chapter 1.3.2.2. (Figure 7, 18 and Table Appendix 3), tetrahydrobenzindoles (e.g. DR-4004 **13**<sup>187,188</sup>) were validated as potent 5-HT<sub>7</sub>R scaffolds. As an attempt to synthesize new, selective and less lipophilic compounds we designed (arylpiperazinylethyl)spiro[indoline-3,3'-pyrrolidin]-2-ones (**75**) as potential 5-HT<sub>7</sub>R ligands.



**Figure 18** – Design concept for the SAR analysis of (arylpiperazinylethyl)spiro[indoline-3,3'-pyrrolidin]-2-ones (**75**).

Designing selective 5-HT<sub>7</sub>R compounds against other serotonin receptors (5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R) we used the pharmacophore models of López *et al.*<sup>98,100</sup> and Medina *et al.*<sup>189</sup> These models suggest that it is beneficial to decrease the distance between PI (positive ionizable) and HBA (H-bond acceptor), to facilitate interactions to Ser<sup>6.55</sup> (Ala<sup>6.55</sup> at 5-HT<sub>1A</sub>R), and to introduce polar substituents at HYD<sub>1</sub>-HYD<sub>2</sub> to contact Arg<sup>7.36</sup> (that is absent in 5-HT<sub>1A</sub>R).

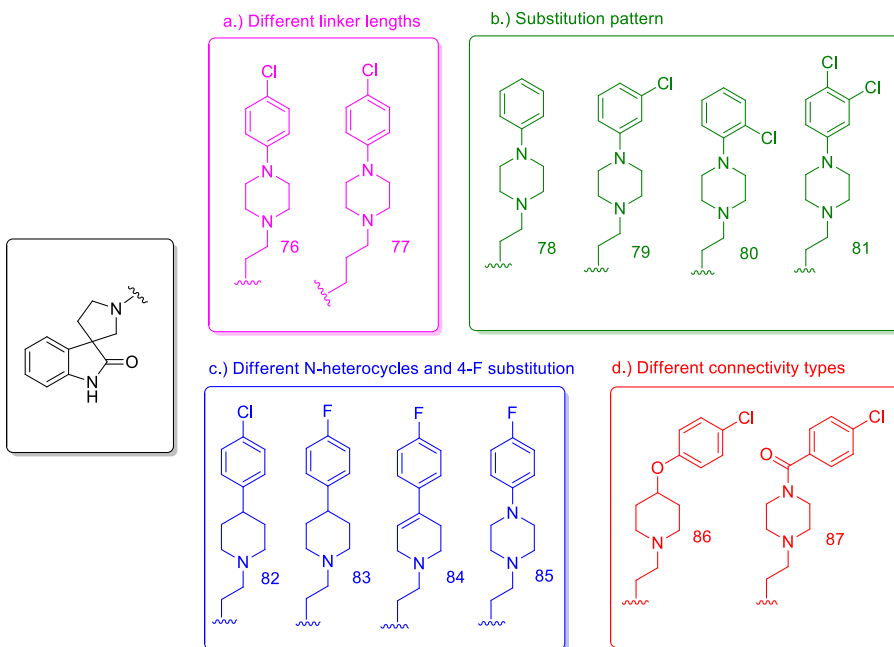
Following these guidelines we set our objective for exploring selectivity drivers around the spiro[pyrrolidine-3,3'-oxindoles] core.

The 5-HT<sub>7</sub>R pharmacophore (**Figure in 1.3.2.2.**) also suggested that we need a four-atom linker from the spiro carbon C-3 to the pyrrolidine-nitrogen and therefore connected this with two methylenes (**Figure 18**) to the arylpiperazine moiety (**75**). Thus in line with other studies<sup>188,190-194</sup> we prepared compounds with two atom spacer (**Figure 19**, compound **76** (CH<sub>2</sub>)<sub>n</sub> n=2), although n=3 (**77**) was also synthesized in order to investigate its effect on flexibility and HBA-PI distances.

Following the findings of SAR data around the oxindole,<sup>195</sup> benzimidazolone<sup>194</sup> and quinazolinone cores,<sup>196</sup> the substitution of the aromatic ring either did not substantially affect or even decreased the 5-HT<sub>7</sub>R affinity. Thus we kept our spiro[pyrrolidine-3,3'-oxindole] core unsubstituted (**Figure 18** in green, **Figure 19**).

Halo-scan, however, was planned to explore the substituent effects around the phenyl ring of the arylpiperazine moiety comparing the profile of the unsubstituted (**78**) and the meta-Cl (**79**), ortho-Cl (**80**), para-Cl (**76**), and 3,4-dichloro-derivatives (**81**) (see **Figure 19**)

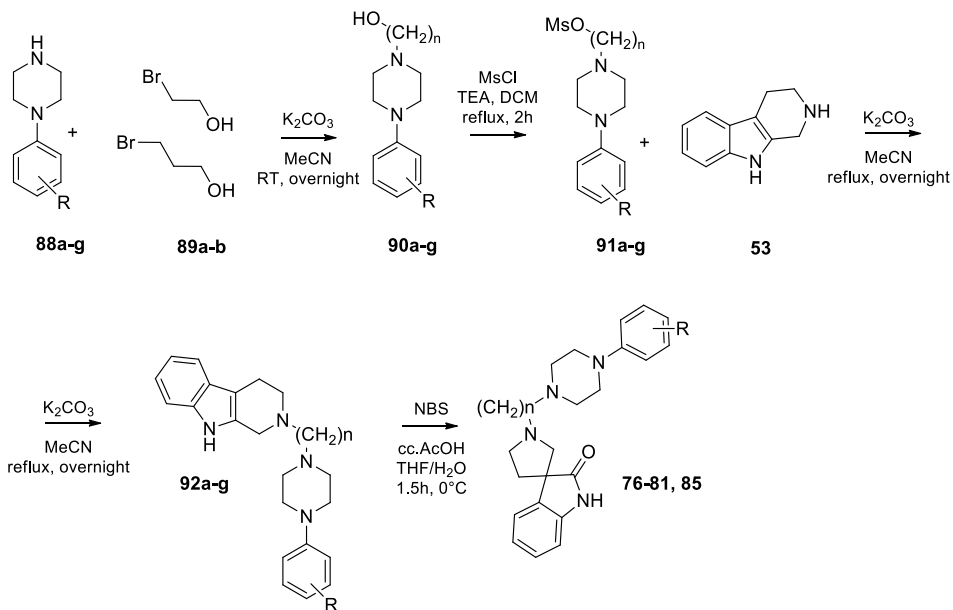
Furthermore, we plan to explore the impact of the HYD<sub>1</sub>/HYD<sub>2</sub> features on the 5-HT<sub>7</sub>R affinity and selectivity we replaced the canonical arylpiperazine by phenyl-, and phenoxy-piperidines (**82**, **83**, **86**), 5,6-dihydropyridine (**84**), and benzoyl-piperazine (**87**) moieties (see **Figure 19**).



**Figure 19** – Designed spiro[indoline-3,3'-pyrrolidin]-2-ones for SAR analysis. a.) different linker lengths (76-77), b.) substitution pattern (78-81), c.) different N-heterocycles and 4-F substitution (82-85), d.) different connectivity types (86-87).

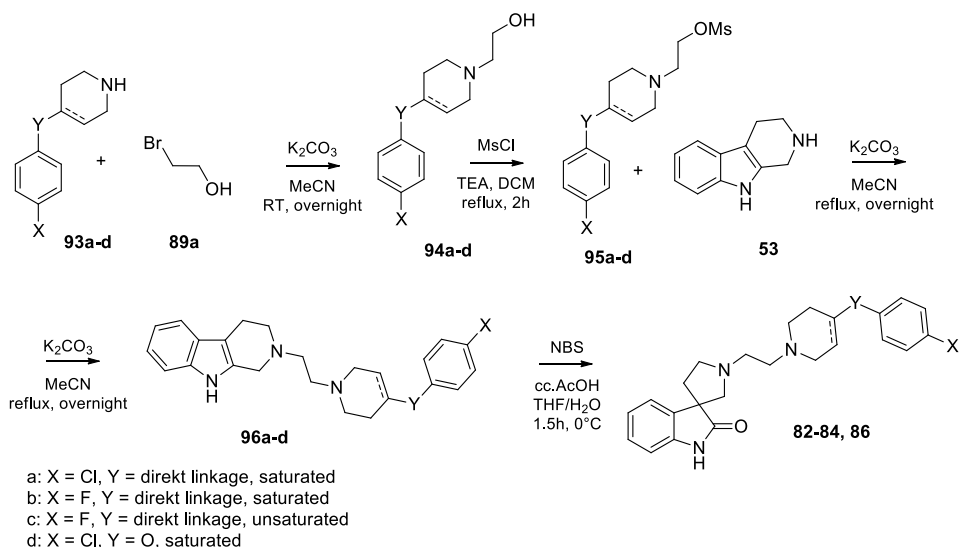
## 2.4.2. Optimization of spiro[indoline-3,3'-pyrrolidin]-2-ones as 5-HT<sub>7</sub>R ligands

The synthesis of the spiro-oxindole derivatives was based on the strategies presented in **Sections 1.4., page 25** and **2.3.2., page 54 (Schemes 3-4)**. Arylpiperazine derivatives (76-81, 85) were synthesized from the corresponding secondary amines (88a-g). First, cyclic secondary amines were alkylated by 2-bromoethanol (89a) (in case of 4-chlorophenylpiperazine (88a) either by 3-bromopropanol (89b)) followed by the mesylation of the appropriate alcohols using mesyl chloride. A protecting group on the basic nitrogen of the tryptoline was necessary to avoid decomposition during the spiro-rearrangement reaction (**Scheme 1** in **Section 1.4., page 25**). N-alkylation by the corresponding mesylates led to intermediates **92a-g (Scheme 9)** that served the desired protection of the tetrahydro- $\beta$ -carboline (**53**) nitrogen for the final spiro-cyclization step to afford derivatives **76-81** and **85**.



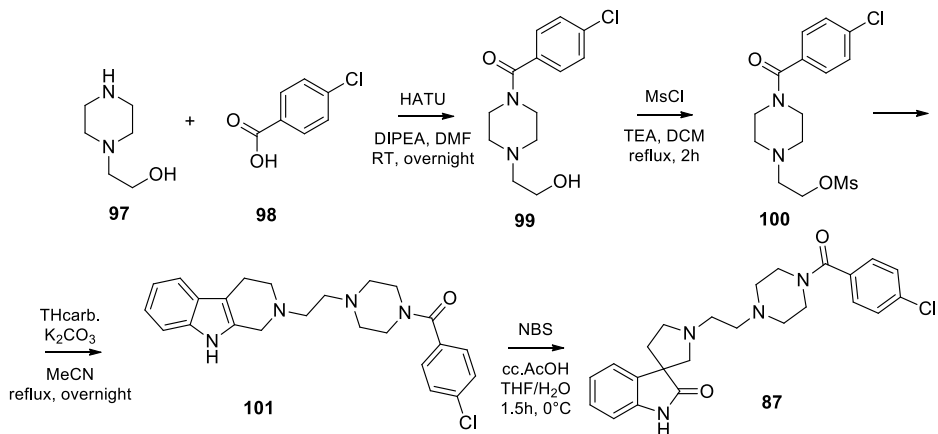
**Scheme 9** – Synthesis strategy leading to the aryl-piperazine derivatives (**76-81, 85**)

The synthesis of the aryl-piperidine (**82-83**), the 5,6-dihydro-pyridine (**84**), and the phenoxy (**86**) analogues (**Scheme 10**) was also started from the corresponding secondary amines (**93a-d**), followed by the alkylation with 2-bromoethanol (**89a**), to afford the piperidiny-1-/5,6-dihydropyridyl/phenoxy-piperidiny-1-alcohols (**94a-d**). The crude mesylated derivatives (**95a-d**) were then used for alkylation of the tetrahydro-β-carboline (**53**) to get **96a-d** that was spirocyclicized to **82-84** and **86**, respectively.



**Scheme 10** – Synthetic route for aryl-piperidine and tetrahydro-pyridine derivatives (85-87)

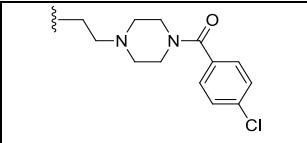
Alternatively 2-(piperazin-1-yl)ethanol (**97**) was used as starting material for the synthesis of the benzoylpiperazinyl analogue **87** (**Scheme 12**). The acylation<sup>197</sup> of **97** with the 4-chlorobenzoic acid (**98**) gave the corresponding amide (**99**), which was treated with mesyl-chloride to yield the mesylate (**100**). After the alkylation of tetrahydro-β-carboline (**53**) with **100**, the spiro-cyclization of intermediate **101** afforded **87**.



**Scheme 12** – Synthetic route to the amide-analogue (**87**)

The synthesized compounds were investigated in competition binding assays against 5-HT<sub>7</sub>R and other related serotonin receptor subtypes including 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R and 5-HT<sub>6</sub>R (**Table 17**).

ID	Structure	5-HT <sub>1A</sub>	5-HT <sub>2A</sub>	5-HT <sub>6</sub>	5-HT <sub>7</sub>
		K <sub>i</sub> (μM)	K <sub>i</sub> (μM)	K <sub>i</sub> (μM)	K <sub>i</sub> (μM)
76		3.755	0.068	4.737	0.029
		129.5	2.3	163.3	
77		2.891	0.715	5.559	0.067
		43.1	10.7	83.0	
78		6.005	0.496	6.556	0.100
		60.1	5.0	65.6	
79		2.687	0.351	2.005	0.055
		48.9	6.4	36.5	
80		1.096	0.452	2.843	0.245
		4.5	1.8	11.6	
81		4.054	0.270	1.131	0.038
		106.7	7.1	29.8	
82		14.900	0.404	4.577	0.284
		52.5	1.4	16.1	
83		52.010	0.259	5.285	0.275
		189.1	0.9	19.2	
84		3.949	0.075	6.267	0.043
		91.8	1.7	145.7	
85		14.160	0.059	4.744	0.021
		674.3	2.8	225.9	
86		13.790	0.343	1.693	0.121
		114.0	2.8	14.0	

87		48.240	17.330	7.746	0.529
		91.2	32.8	14.6	

**Table 17** - Serotonergic GPCR panel of the spiro-oxindole derivatives as measured in binding assays of four serotonin receptors (measured values of  $K_i$  ( $\mu\text{M}$ )). 5-HT<sub>7</sub>R selectivities are shown in italics.

The unsubstituted phenylpiperazine derivative (**78**) showed reasonably high affinity towards 5-HT<sub>7</sub>R, however its selectivity was moderate. Halo-scan around the phenyl ring revealed that the 5-HT<sub>7</sub>R affinity and selectivities against 5-HT<sub>1A</sub> and 5-HT<sub>6</sub>R are increasing from ortho (**80**) – meta (**79**) – para (**76**) direction. In fact, the para-Cl derivate (**76**) showed low nanomolar affinity for the target and more than hundred-fold selectivity against two of the three other serotonin receptors. Selectivity against these receptors was further improved for the para-F analogue (**85**). All of these compounds have submicromolar 5-HT<sub>2A</sub> affinity that might be interesting due to potentially comorbid indications<sup>194</sup>. As compared to compound **76**, the longer ([CH<sub>2</sub>]<sub>3</sub>) alkyl chain in **77** increased the distance between the PI and HBA features, accounting for an overall loss of affinity for the 5-HT<sub>7</sub>R. Selectivity against 5-HT<sub>1A</sub>R decreased, however, the selectivity against the 5-HT<sub>2A</sub>R subtype has improved. **78** showed lower selectivity (65.6-fold) towards 5-HT<sub>6</sub>R than **76** (163.3-fold).

Starting from the most active 4-fluorine derivative (**85**) the 5,6-dihydropyridine derivative (**84**) showed decreased selectivity (91.8-fold) against 5-HT<sub>1A</sub>R. Replacing the piperazine ring by piperidine (**83**) gave 10-times lower affinity that was further confirmed by the corresponding 4-chloro analogue (**82**). Selectivities against 5-HT<sub>1A</sub>R and 5-HT<sub>2A</sub>R did not change significantly, but the 5-HT<sub>6</sub>R selectivity decreased ten-times. Similar to **82** and **83**, introduction of the phenoxy-piperidine moiety (**86**) improved the affinity towards 5-HT<sub>6</sub>R. Finally, the benzoyl-piperazine derivative (**87**) showed reduced affinity to 5-HT<sub>7</sub>R, however, it has the highest selectivity against 5-HT<sub>2A</sub>R (32.8-fold).

In summary, preliminary SAR data demonstrates that spiro[pyrrolidine-3,3'-oxindoles] can be optimized to potent 5-HT<sub>7</sub>R ligands. We confirmed that the 2-methylene linker ensures the ideal distance between HYD<sub>1</sub> and HYD<sub>2</sub> and therefore it is beneficial for the 5-HT<sub>7</sub>R affinity. Para substitution at the aryl-piperazine moiety is advantageous both for affinity and selectivity against 5-HT<sub>1A</sub>R and 5-HT<sub>6</sub>R. Actually, the 4-fluoro analogue (**85**) showed the best affinity and most remarkable selectivity against 5-HT<sub>1A</sub>R and 5-HT<sub>6</sub>R. Selectivity against 5-HT<sub>2A</sub>R might be improved by replacing the phenyl substituent of the piperazine by a benzoyl group. These results show the potential of this novel

chemotype and validate its further optimization for more detailed *in vivo* characterization in disease models.

## **2.5. Consensus-scoring based docking for the design of 5-HT<sub>2B</sub>R selective antagonists<sup>198</sup>**

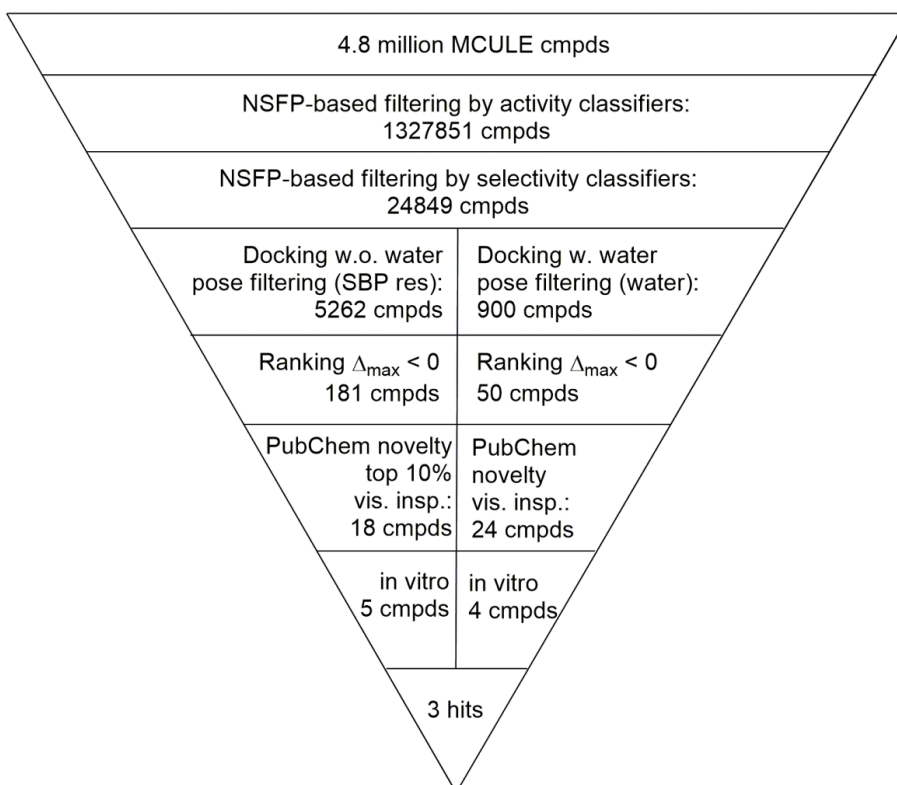
A thorough analysis of the crystal structures of aminergic GPCR proteins revealed that most of the receptors have a secondary binding pocket (SBP) formed at the extracellular part of the protein by the participation of the extracellular loop 2 (ECL2). This site contains a significant proportion of non-conserved amino acids across certain aminergic GPCR's that provides an opportunity to obtain subtype selectivity. Michino and colleagues<sup>55</sup> collected all such binding cavities with at least three residues interacting with ligands outside of the canonical orthosteric binding pocket (OBP). Their analysis identified residues involved in interactions with ligands constituting the SBP in 30 out of 36 X-ray structures of aminergic receptors. These cavities consisting of non-conserved residues open the possibilities for the design of subtype-selective ligands by maintaining crucial ligand-contacts in the OBP by an initial fragment, and optimizing it by fragment-linking with subtype-specific SBP-fragments. Based on this approach, a two-step ligand docking strategy was developed for the design of dopamine receptor subtype selective bitopic ligands<sup>199</sup>. In this study we combined Structure Connectivity Fingerprint (SCFP) based machine-learning tool with SBP-targeted docking.

### **2.5.1. Concept of SCFP-based machine-learning method and combination with an SBP-targeted consensus ranking approach**

Understanding the key drivers of subtype selectivity (chapter 1.3.3.2.) we aim to identify novel selective 5-HT<sub>2B</sub>R ligands using a combination of ligand- and structure-based methods. At the first stage of screening, we applied machine learning tools trained on the molecular structure of known 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R ligands (developed in collaboration with the Department of Medicinal Chemistry (Polish Academy of Sciences))<sup>200,201</sup>. Starting from the main principle of fragment-based drug design (FBDD) these compounds were represented by Neighbouring Substructures Fingerprint (NSFP) that opens the possibilities of quickly creating fragment libraries with desired target-specific, class-specific or even family-specific properties. In our present study we combined both methodologies, i.e. the fragment based NSFP fingerprint with fragment docking. Applying this approach for 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R serotonin receptors we were able to identify new and selective 5-HT<sub>2B</sub>R ligands providing potential chemical starting points for further optimization.

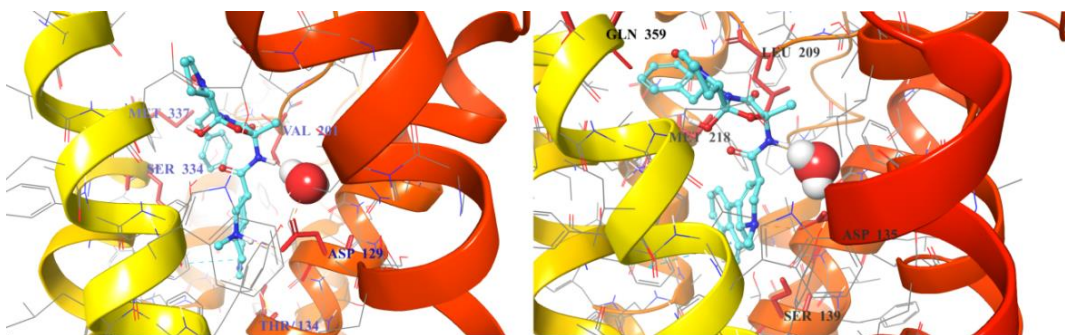
In this study we built a NSFP fingerprint-based machine learning model using *in vitro* activity data available for human 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R receptors

in ChEMBL<sup>202</sup> (The collection of the training set is described in **Section 3.1.4.1, page 89**). Compounds were divided into actives and inactives using binding affinity thresholds. In this way four sets were acquired: active for 5-HT<sub>1B</sub> (1B<sub>active</sub>), active for 5-HT<sub>2B</sub> (2B<sub>active</sub>), inactive for 5-HT<sub>1B</sub> (1B<sub>inactive</sub>) and inactive for 5-HT<sub>2B</sub> (2B<sub>inactive</sub>). 5-HT<sub>1B</sub> actives and 5-HT<sub>2B</sub> inactives, and 5-HT<sub>2B</sub> actives and 5-HT<sub>1B</sub> inactives were used to collect 1B<sub>selective</sub> and 2B<sub>selective</sub> subsets. The nonselective subset contains compounds active on both 5-HT<sub>1B</sub> and 5-HT<sub>2B</sub> receptors. Altogether, 7 training sets were used for building machine learning models (1-2. actives, 3-4. inactives, 5-6. selectives for both 1B and 2B receptor subtypes, and 7. nonselectives). The trained model (see **Section 3.1.4.1, page 89**) was used prospectively for the screening of the commercial database of MCule<sup>173,174</sup> database (method description in **Section 3.1.4.2, page 90**). The in-stock MCule database containing 4.8M molecules was filtered prospectively in an activity-classification (resulting in 1327851 putative 5-HT<sub>2B</sub>R active compounds) and consecutively in a selectivity-classification step resulting in 24849 putative 5-HT<sub>2B</sub>R selective compounds (**Figure 20**).



**Figure 20.** Flowchart of the virtual screening cascade used for the identification of 5-HT<sub>2B</sub> selective ligands

In the next sequential filtering step, the pre-filtered 5-HT<sub>2B</sub> selective compound set of 24849 compounds was subjected to two complementary docking workflows (i) considering non-conserved amino acid pairs and ranking-based consensus scoring, and (ii) accounting for water molecules (**Figure 20**). The set of 24849 putative 5-HT<sub>2B</sub> selective compounds was docked into four crystal structures available to date (PDB ID: 4IAQ, 4IAR<sup>203</sup> for h5-HT<sub>1B</sub>R, and 4IB4<sup>204</sup>, 4NC3<sup>205</sup> for h5-HT<sub>2B</sub>R). Our docking constraints involved forming a hydrogen bond to Asp129/132<sup>3,32</sup> (in 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R respectively), that anchors the charged amines of aminergic ligands<sup>206</sup>, and any interactions with the designated non-conserved amino acid pairs present in the SBPs of the two receptors based on the work of Michino and colleagues<sup>55</sup> (see **Figure 21 – A, B** and **Table 22**). Each pose-filtered ligand was ranked at the four receptors based on their Glide-scores. The consensus scoring<sup>154</sup> resulted in the selection of 181 compounds showing preference to 5-HT<sub>2B</sub>R based on their rankings. The set of 181 ligands was sorted by their  $\Delta_{\max}$  values = [(RANK at 4IB4 receptor) + (RANK at 4NC3 receptor)] – [(RANK at 4IAQ receptor) + (RANK at 4IAR receptor)]. The top 10% (18 ligands) were subjected to a novelty check against PubChem<sup>152</sup>, and were selected for visual inspection, analyzing their binding modes. This nominated five compounds (**102, 104-107** marked with “a” in **Table 18**) with the largest difference of ranks obtained on 5-HT<sub>2B</sub> and 5-HT<sub>1B</sub> structures for in vitro tests.

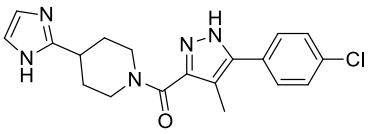
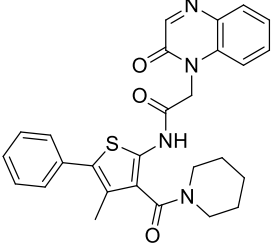
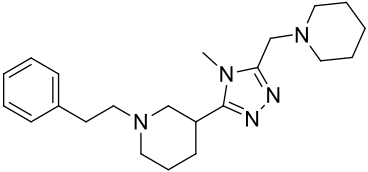
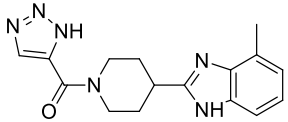
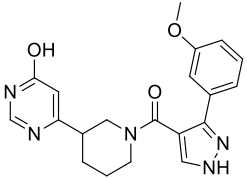
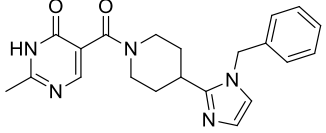


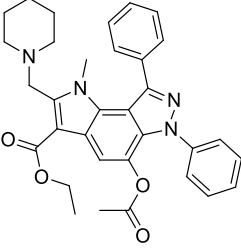
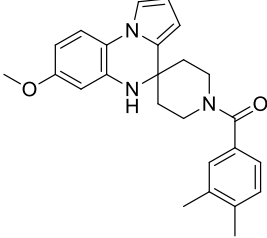
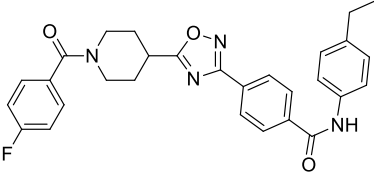
**Figure 21.** Residues (red) used in pose filtering for 5-HT<sub>1B</sub>R (PDB ID: 4IAQ) (panel A) and 5-HT<sub>2B</sub>R X-Ray structures (PDB ID: 4IB4) (panel B); Water molecules forming interactions with the Asp<sup>3,32</sup> shown as sphere representation.

As a complementary approach, we have filtered a second pool of compounds, considering the structural waters found in the X-Ray structures of the receptors (PDB ID: 4IAQ<sup>203</sup> for h5-HT<sub>1B</sub>R, and 4IB4<sup>204</sup> for h5-HT<sub>2B</sub>R). In both structures, one water molecule anchoring the oxo-group of the ergoline ligands' amide-moiety with the Asp<sub>129/132</sub><sup>3,32</sup> (in 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R respectively) in TM3 was considered. Whereas the waters settle the actual position and conformation of the co-crystallized ligand, they might also govern other ligands to reach the SBP (see **Figure 18**). Interactions to these water molecules or directly to the protein were used for pose filtering yielding 900 ligands with satisfactory binding modes. The 900 pose-filtered compounds were ranked, and  $\Delta_{\max} = (\text{RANK at 4IB4 receptor}) - (\text{RANK at 4IAQ receptor})$  values were calculated. The ligands were sorted by  $\Delta_{\max}$  values, and 50 structurally diverse compounds with high rank differences (i.e.,  $\Delta_{\max} \ll 0$ ) were selected for novelty screening. We identified 24 compounds with no known similar and active analogues in PubChem that were subjected to visual inspection selecting four hits (**103**, **108-110** marked with "b" in **Table 18**) with feasible binding mode (compounds marked with "b" in **Table 18**).

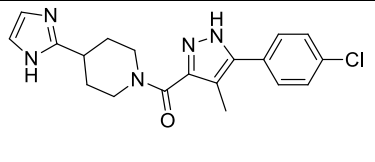
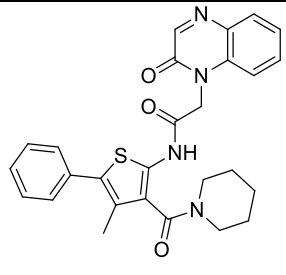
### 2.5.2. Selection of virtual screening hits for GPCR-panel screening

Altogether nine compounds - predicted to be 2B selective - were selected for competition binding assay for 5-HT<sub>1B</sub>R, and 5-HT<sub>2B</sub>R (**Table 18**). Three hit compounds were identified (compounds **102-104**), showing preference towards the desired 5-HT<sub>2B</sub>R target over 5-HT<sub>1B</sub>R off-target. Moreover, selectivity of two hit 5-HT<sub>2B</sub>R ligands **102** and **103** over four serotonin receptors (5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R and 5-HT<sub>7</sub>R) was confirmed (**Table 19**).

Structure	ID	5HT <sub>1B</sub> [%] <sup>1</sup>	5HT <sub>2B</sub> [%] <sup>1</sup>	5HT <sub>1B</sub> [nM] <sup>2</sup>	5HT <sub>2B</sub> [nM] <sup>2</sup>
	102 a	-	-	2934.7±321.3	0.3±0.07
	103 b	-	-	2099±622.2	235.1±15.9
	104 a	7±4	-	inactive at 10 μM	2612.9±422.1
	105 a	12±4	40±3	-	-
	106 a	7±1	36±4	-	-
	107 a	19±3	34±4	-	-

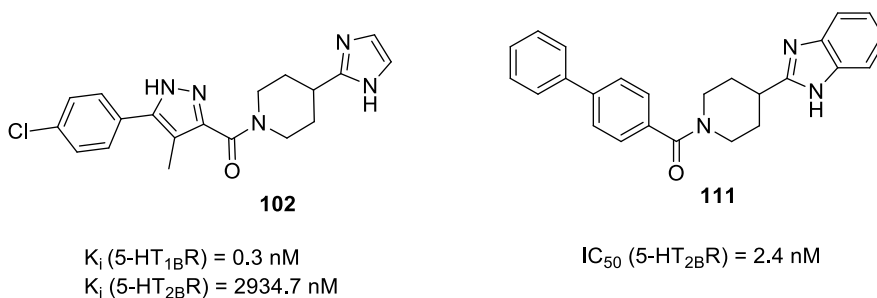
	<b>108</b> <sup>b</sup>	26±4	34±4	-	-
	<b>109</b> <sup>b</sup>	28±3	31±4	-	-
	<b>110</b> <sup>b</sup>	1±1	24±4	-	-

**Table 18.** Structures and measured *in vitro* assay data of top-ranked compounds (sorted by percentage of inhibition and  $K_i$  [nM] in h5-HT<sub>2B</sub>R competition binding assay). Values represent the mean±SD of three independent assays with duplicate measurements. <sup>1</sup>Percentage of inhibition measured at 10  $\mu$ M concentration; <sup>2</sup> $K_i$  measured (reported in nM units) if percentage of inhibition exceeded 50%; <sup>a</sup>Selection not accounting waters, <sup>b</sup>Selection accounting waters.

Compound	ID	5-HT <sub>1A</sub> K <sub>i</sub> (μM)	5-HT <sub>2A</sub> K <sub>i</sub> (μM)	5-HT <sub>6</sub> K <sub>i</sub> (μM)	5-HT <sub>7</sub> K <sub>i</sub> (μM)
	102	> 20	> 20	13.2±1.5	> 20
	103	> 20	> 20	7.5±0.9	> 20

**Table 19.** 5-HT panel screening of the best hits. Values represent the mean±SD of three independent assays with duplicate measurements.

Our sequential screening protocol was able to identify potent and selective compounds with a 33% hit rate, including a highly active (subnanomolar activity) compound with almost 10<sup>4</sup> selectivity factor. The 5-HT<sub>2B</sub> vs. 5-HT<sub>1B</sub> selectivity of the closest literature analogue (compound **111** shown in **Figure 19**) was not confirmed in ref.<sup>207</sup>. Lacking the relevant 5-HT<sub>1B</sub> affinity, **111** was not part of the ChEMBL training set that underlines the efficiency of the NSFP based classification models.

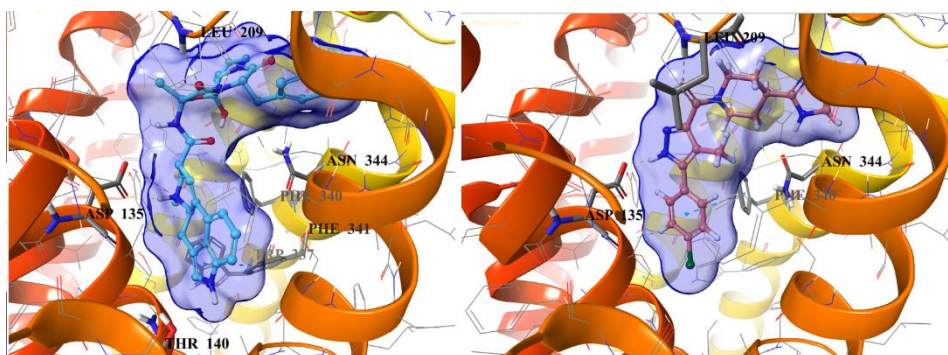


**Figure 19.** Structure of hit compound (**105**) and its' closely related analogue (**114**)

### 2.5.3. Binding Mode Analysis of the selective 5-HT<sub>2B</sub>R Ligand

The extended binding pocket (secondary binding pocket) of 5-HT<sub>2B</sub>R, originally occupied by the peptide portion of the ergoline ligand is defined by several residues of ECL2, TM5, TM7 and TM6. The 4-chlorophenyl tail of the





**Figure 21.** Binding site surfaces (blue spheres) of ergotamine (panel A) and **105** (panel B) in the 5-HT<sub>2B</sub>R crystal structure. Important ligand-contacting residues are shown in thick tube representation.

In summary, we aimed identifying selective 5-HT<sub>2B</sub>R ligands by virtual screening. Given the sequential and structural similarities of serotonergic receptors we hypothesized that receptor selectivity is basically driven from the secondary binding pocket (SBP). Consequently we developed a screening strategy combining both ligand-based and structure-based approaches. For the first level of our hierarchical approach we used Neighbouring Substructures Fingerprints (NSFP) and developed machine learning tools select potential 5HT<sub>2B</sub>R selective molecules from a large database of druglike compounds. Next this subset was subjected to docking calculations and we identified compounds showing different interactions with the secondary binding pockets of 5-HT<sub>2B</sub>R and 5-HT<sub>1B</sub>R. Careful analysis of the binding mode allowed us to select nine compounds for biological testing. Out of these, three compounds showed significant 5-HT<sub>2B</sub>R affinity, one in the low micromolar (**104**), one in the submicromolar (**103**) and one in the subnanomolar range (**102**). Compounds with submicromolar 5-HT<sub>2B</sub>R affinity were further profiled against a set of serotonin receptors including 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R and 5-HT<sub>7</sub>R. The best compound (**102**) showed subnanomolar affinity and ten thousand-fold selectivity for 5-HT<sub>2B</sub>R that nominates it for *in vivo* testing.

## Chapter 3 – Computational Methods and Experimental Section

### 3.1. Computational Methods and Procedures

#### 3.1.1. Computational procedures for physicochemical-property based scoring method (FrAGS)

##### 3.1.1.1. Compilation of training set

The assembly of the training set was initiated by downloading the public database of GPCR-SARfari<sup>208</sup> as part of the ChEMBL database<sup>209,210</sup>. The database was processed further by a database workflow-manager program Knime<sup>211</sup>. GPCR SARfari (version 3.00, June 2012) contained 947914 entries of chemical structures with related *in vitro* activity data for different GPCR-targets. The database was processed following these steps: only binding-related data were kept, while data of functional assays and those related to ADME properties were discarded. Binding affinities defined as IC<sub>50</sub>, logIC<sub>50</sub>, K<sub>i</sub> and logK<sub>i</sub> were converted uniformly to pActivity and were treated on an equal footing with pK<sub>i</sub> values. When binding affinities for different species of the certain receptor were available (i.e. human, mouse, rat, etc.), only the entry corresponding to the highest activity was used in the further steps. Counter ions of salts were removed, and only one desalted entry was kept with the corresponding highest pK<sub>i</sub> value. This process resulted in 166699 structurally unique entries (removing duplicates), with activity data for several GPCR targets. A fragment-based interpretation of the ligand-efficiency metrics<sup>3</sup>, the size-independent ligand efficiency (SILE)<sup>212</sup> was used to classify compounds as active or inactive on a target. SILE is an empirically derived measure suitable for the comparison of compounds with different sizes (advantageous for fragment-sized molecules):

$$SILE = pK_i / N_{heavy}^{0.3}$$

The cut-off for sorting of GPCR-active fragments was determined by analysing the relationship between pK<sub>i</sub> and SILE in the  $8 \leq N_{heavy} \leq 22$  range (see **Table 19**).

N <sub>heavy</sub>	pK <sub>i</sub> at different SILE values				
	1.85	1.9	1.95	2	2.05
8	3.5	3.5	<b>3.6</b>	3.7	3.8
9	3.6	3.7	<b>3.8</b>	3.9	4
10	3.7	3.8	<b>3.9</b>	4	4.1
11	3.8	3.9	<b>4</b>	4.1	4.2
12	3.9	4.0	<b>4.1</b>	4.2	4.3
13	4.0	4.1	<b>4.2</b>	4.3	4.4
14	4.1	4.2	<b>4.3</b>	4.4	4.5
15	4.2	4.3	<b>4.4</b>	4.5	4.6
16	4.3	4.4	<b>4.5</b>	4.6	4.7
17	4.3	4.4	<b>4.6</b>	4.7	4.8
18	4.4	4.5	<b>4.6</b>	4.8	4.9
19	4.5	4.6	<b>4.7</b>	4.8	5
20	4.5	4.7	<b>4.8</b>	4.9	5
21	4.6	4.7	<b>4.9</b>	5	5.1
22	4.7	4.8	<b>4.9</b>	5.1	5.2

**Table 20.** – Relationship between the activity (pK<sub>i</sub>) and SILE in the  $8 \leq N_{\text{heavy}} \leq 22$  range

Actives were collected by SILE  $\geq 1.95$  criterion that corresponds to K<sub>i</sub> values suitable for fragment-screening activity (ranging between 100 $\mu$ M and 10 $\mu$ M – ranging between 3.6 and 4.9). After removing entries with a SILE lower than 1.95, a remaining data-set of 144759 entries was used for the identification of GPCR-active fragments. Unlike the general approach applying molecular weight as size-determining definition, our measure was the number of heavy atoms<sup>213</sup>. Molecular-weight based filters penalize certain heteroatoms, such as: halogens (e.g. iodine) and sulphur atoms (including the markedly important sulfonyl-type functionalities, such as sulphonamides), due to their proportionately higher atom-weight. Since one heavy atom stands for an average of a 13.3 Da molecular weight<sup>3</sup>, the commonly used Rule of Three (RO3)<sup>2</sup> cut-off MW  $\leq 300$  Da ( $13.3 \times 22 \approx 300$  Da) can be converted to N<sub>heavy</sub>  $\leq 22$ . Keeping molecules with at least 8, and 22 heavy atoms at most, resulted in 10477 unique fragments with their corresponding binding affinity data for a range of GPCR targets. Compounds classified as “GPCR-like fragments” (2370), were defined as having SILE of at least 1.95 on at least four different GPCR-targets. We have analysed the activity data from the side of receptors. We concluded that the active set contained activity data related to 139 different receptors (and subtypes), out of which 87 receptors were aminergic GPCRs (serotonine,  $\beta$ -adrenerg, dopamine, histamine, trace-amine, octopamine, muscarinic-acetylcholine receptors) and the remaining 52 were not part of the aminergic subclass, but members of the class-A GPCRs

(e.g. opioid-, adenosine-, sphingosine-1-phosphate-, thromboxane-, melatonin-, orphan-, and angiotensin-receptors). The vast majority of the fragments (2183 out of 2370, 92%) were characterized exclusively with aminergic GPCR activity data, used as the training set.

For the determination of the typical aminergic GPCR-fragment characteristics, we created a reference set derived from the ChEMBL as a counter-part during the analysis. First, the compounds of the active GPCR set were removed from the 1292344 entries of ChEMBL (version 16)<sup>209,210</sup>. Salt-forms were disentangled from the core ligand structures<sup>214</sup>, and duplicates were removed, resulting in structurally unique entries. After filtering by our size cut-off:  $8 \leq N_{\text{heavy}} \leq 22$ , 309962 fragments remained, out of which 5000 fragments were picked out randomly to generate a set proportional in size with the active set of 2183 compounds.

### 3.1.1.2. Compilation of ChEMBL retrospective validation set

The inactive set was created similarly to the reference set generated for deriving the FrAGS. The entire ChEMBL database<sup>209,210</sup> of about 1.3 million entries was filtered by the previously created active fragments. Counter ions of salts were removed; finally 96539 unique fragments were kept. This set was merged with the 2183 active fragments to give a validating set with 2.2 % actives. Relevant properties were calculated in Knime<sup>211</sup> for the compiled 98722 fragments, and FrAGS were calculated.

### 3.1.1.3. Compilation of PubChem retrospective validation set

Second validation was carried by retrieving aminergic GPCR targeted HTS data from PubChem<sup>152</sup>. High-throughput screenings for allosteric-binders, or targeting  $\beta$ -arrestin pathways, or used as counter-screening were sorted out. Only confirmed active fragments with better than 10  $\mu\text{M}$  activity were taken into consideration. The following aminergic GPCR targets were represented in HTS campaigns: 5HT<sub>1A</sub>R with 7 confirmed actives, 5HT<sub>1E</sub>R with 51 confirmed actives, trace-amine receptor (TAAR<sub>1</sub>) with 375 confirmed agonists and with 36 confirmed antagonists. Molecules satisfying fragment-size criterion (heavy atom count between 8 and 22) were kept providing 200 fragments with confirmed active data on three aminergic-GPCRs (5HT<sub>1A</sub>R, 5HT<sub>1E</sub>R, TAAR<sub>1</sub>). 31 confirmed 5HT<sub>1R</sub>-active fragments and 169 TAAR<sub>1</sub>-active fragments were added to a set of inactive fragments extracted from the corresponding primary screening sets by filtering all inactive compounds by the number of heavy atoms ( $8 \leq N_{\text{heavy}} \leq 22$ ). After removing duplicates 296510 fragments remained out of which 3100 and 16900 randomly sampled molecules (1:100 proportion of actives versus inactives) from the 5HT<sub>1R</sub> and TAAR<sub>1</sub> screens, respectively, were selected for validation.

### 3.1.2. Computational procedures for the consensus-scoring based docking method (FrACS)

#### 3.1.2.1. Collection of protein structures

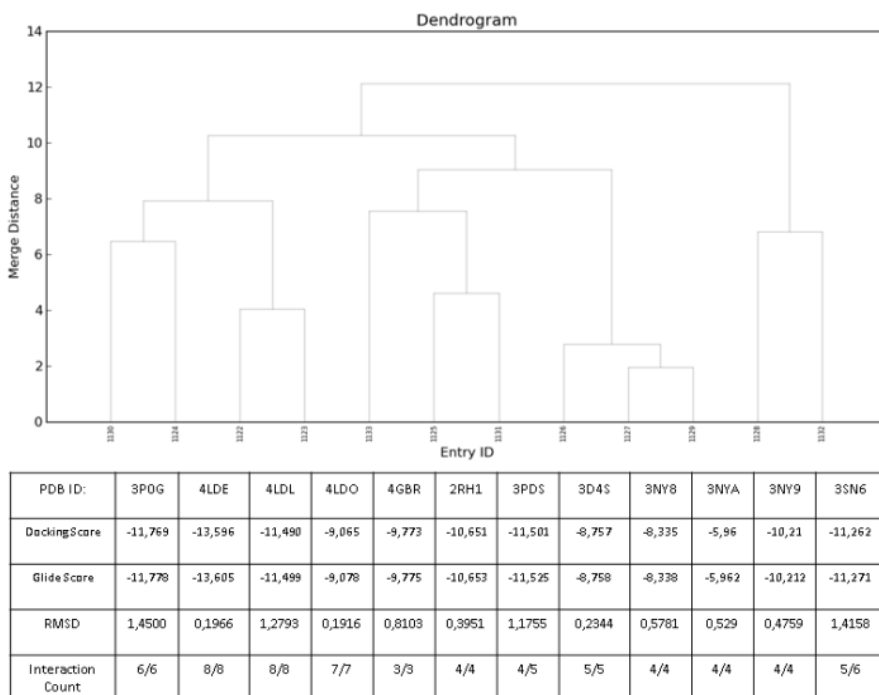
Three structures of the human M2 Muscarinic Acetylcholine receptor (PDB IDs: 3UON, 4MQS, 4MQT)<sup>164,215,216</sup> structures of the human  $\beta$ 2 adrenoceptors (2RH1, 3D4S, 3NY8, 3NY9, 3NYA, 3P0G, 3PDS, 4GBR, 4LDE, 4LDL, 4LDO, 3P0G, 2R4R, 2R4S, 3KJ6, 4QKX)<sup>57,166,217–223</sup>, 2-2 structures of human 5-HT<sub>1B</sub>R (4IAQ, 4IAR)<sup>224</sup> and 5-HT<sub>2B</sub>R receptors (4IB4, 4NC3)<sup>225,226</sup> and the X-ray structures of human dopamine DRD<sub>3</sub> receptor (3PBL)<sup>165</sup> and human histamine H<sub>1</sub>R receptor (3RZE)<sup>163</sup> were collected from the Protein Data Bank (PDB)<sup>227</sup>. Two adrenergic receptor apo structures (2R4R and 2R4S), and one structure (4QKX) containing a covalent agonist ligand were excluded from our analysis. The remaining structures were first prepared for docking by keeping the monomeric receptor, removing crystallization appendices (fatty acids, diethylene glycol, triethyleneglycol, butanediol, palmitic acid, tris buffer, phosphate ions, sodium ions, chloride ions, sulphate ions, glucose, maltose, cholesterol and co-crystallized antibodies) and using Schrödinger's Protein Preparation Wizard<sup>228</sup> with default settings, including the assignment of bond orders, adding missing sidechains, and missing hydrogens, creating disulfide bonds, removing crystalline waters, and applying protomer-optimization, and restrained minimization.

#### 3.1.2.2. Selection of structures for the docking ensemble

The ligands were prepared for docking by generating possible predicted ionization states (at a pH range of  $7.0 \pm 2.0$ ), tautomers, and stereoisomers using Schrödinger's Ligand Preparation Wizard<sup>167</sup> with default settings. Grid generation and self-dockings were performed with Grid Generation application and Glide Docking SP (standard precision)<sup>160</sup>, respectively. RMSD values were calculated by superimposing the docked ligands using Schrödinger's Atom Specification Language. Additionally the 2D interaction maps of the binding sites were downloaded from PDB<sup>227</sup>, and the number of recovered secondary interactions was considered in the selection process. Conformational clustering of receptors with multiple experimental structures (muscarinic M<sub>2</sub>R, adrenergic  $\beta$ <sub>2</sub>AR, 5HT<sub>1B</sub>R and 5HT<sub>2B</sub>R) was performed by the average-type linkage method available in Schrödinger's Clustering of Conformers script<sup>229</sup> using the RMSD-matrix calculated on atomic RMSD of the 5 Å proximity of the ligand.

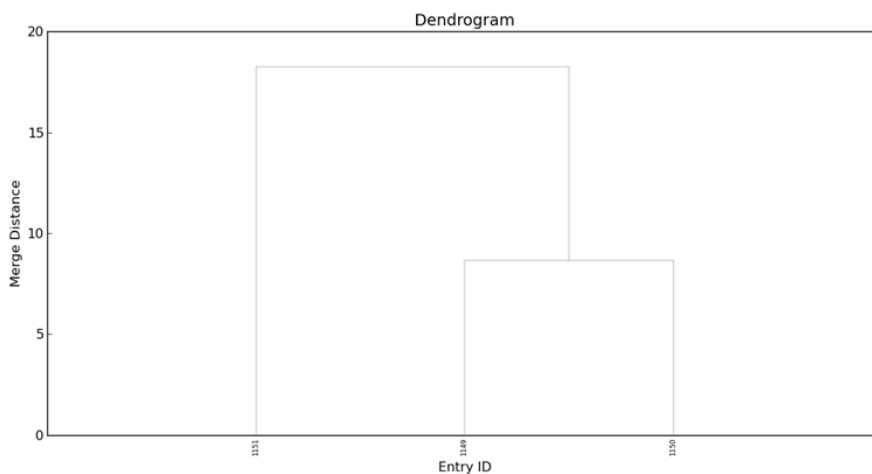
Conformational clustering of the 12  $\beta$ <sub>2</sub>-adrenergic receptor structures (**Figure 22**) resulted in three clusters at 10.27 Å merge distance. One representative structure was selected from each cluster based on their self-docking Glide scores and ligand RMSD values. The  $\beta$ <sub>2</sub> adrenergic structure

(4LDE)<sup>166</sup> with agonist ligand P0G was selected from the first cluster showing a remarkable self-docking result (RMSD = 0.2 Å, Glide score: -13.605, all 8 original binding site interactions recovered). In case of the second cluster 2RH1<sup>217</sup> was selected, containing S-Carazolol agonist as ligand. Despite the moderate docking score, 2RH1 structure was favored over 3PDS<sup>219</sup> due to its lower ligand RMSD value and better recovery of crucial interactions. The third selected structure was 3NY9<sup>218</sup> co-crystallized with an inverse agonist that showed the best RMSD value (0.5 Å) in self docking.



**Figure 22**– Conformational clustering of  $\beta_2$  adrenergic structures

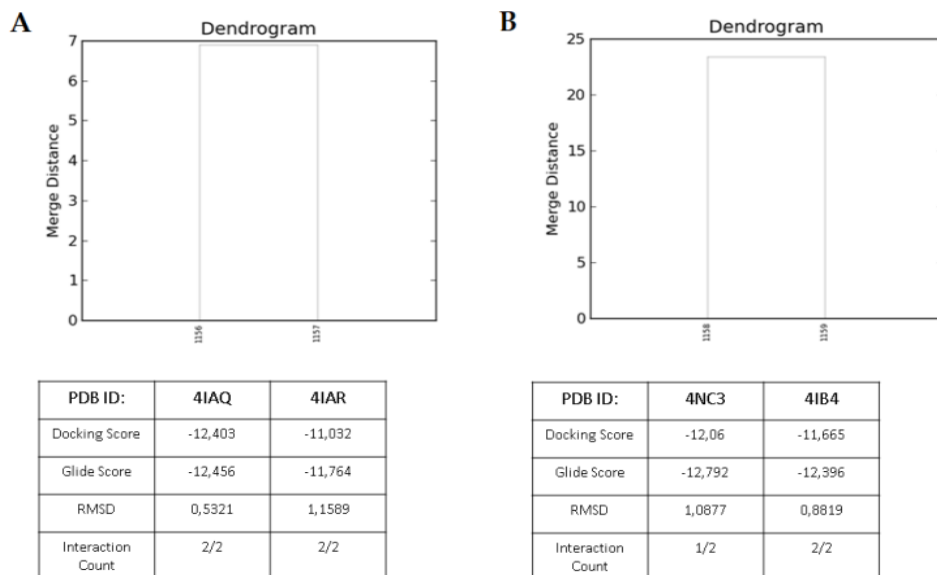
The three available structures (**Figure 23**) of the M2-muscarinic receptor (3UON<sup>215</sup>, 4MQS, 4MQT<sup>164</sup>) showed conformational difference at 18.2 Å merge distance, resulting in two main clusters, from among which 3UON and 4MQT structures were selected for docking studies. 4MQT was favoured over 4MQS due to its docking RMSD (0.2 Å).



PDB ID:	3UON	4MQS	4MQT
Docking Score	-11,343	-9,172	-8,993
Glide Score	-11,345	-9,172	-8,993
RMSD	1,0716	0,3601	0,2025
Interaction Count	3/3	1/1	1/1

**Figure 23**– Conformational Clustering of M2 muscarinic receptor structures

Out of the two 5-HT<sub>1B</sub>R and two 5-HT<sub>2B</sub>R structures 1-1 representatives (4IAQ<sup>224</sup> and 4IB4<sup>225</sup>, respectively) were chosen based on their RMSD values (**Figure 24**). For the human DRD<sub>3</sub> (3PBL<sup>165</sup>) and H1 (3RZE<sup>163</sup>) receptors the single available X-ray structures were selected.



**Figure 24**– Conformational clustering of 5-HT<sub>1B</sub>R (A) and 5-HT<sub>2B</sub>R (B) structures

### 3.1.2.3. Collection of the training set for the optimization of the consensus scoring

The same set of 2,183 GPCR-fragments collected in **Section 2.1.1., page 29** was used in the docking study. A diverse subset (692 fragments in total, clustered by Tanimoto distance of 0.3) was extracted to represent distinct molecular recognition types in the training set. The diversity selection was realized in Knime.com AG's Konstanz Information Miner (Knime)<sup>211</sup>, using Tanimoto-based distance<sup>230</sup> matrix calculator (based on ChemAxon's Chemical Hashed Fingerprint<sup>146</sup>; binary type; length: 1024bit), Hierarchical Clustering (DistMatrix; linkage type: complete linkage), and Hierarchical Cluster Assigner (distance threshold: 0.3). The inactive training set was extracted from the open-source database ChEMBL (version 16)<sup>209,210</sup>, out of which the same 5,000 fragments were filtered as in our previous study<sup>143</sup>. The training set of 692 active and 5,000 inactive fragments were prepared for docking with the same ligand preparation method used for the co-crystallized ligands resulting in 2,381 active and 10,227 inactive generated structures (12,608 entries in total). The training sets were subjected to standard precision docking into the nine previously selected structures using Schrödinger's Virtual Screening Workflow script<sup>160</sup>.

For the validation of using FrAGS and FrACS in consecutive application (**Section 2.2.3., page 46**), a training set was prepared as following: Salt-form duplicates were first removed from ChEMBL, and after fragment-size filtering

327,419 molecules remained containing the 2,183 active training set molecules (intersection with GPCR SARfari<sup>231</sup>).

#### **3.1.2.4. In-house fragment library**

We have compiled an in-house fragment screening library of 1183 compounds in the Medicinal Chemistry Research Group for the purpose of our studies. Compounds with the following characteristics were collected:

- number of heavy atoms between 8 and 22
- $\text{clogP} \leq 3$
- H-bond donors  $\leq 3$
- H-bond acceptors  $\leq 3$
- number of rotatable bonds  $\leq 3$
- $\text{TPSA} \leq 60 \text{ \AA}$

All compounds were collected in 2 mg solid deposit, 20 mM stock solutions and 2 mM diluted solutions in dimethyl-sulfoxide (DMSO).

#### **3.1.2.5. Molecular docking procedures for the consensus-scoring method**

The ligands were prepared for docking by generating possible predicted ionization states (at a pH range of  $7.0 \pm 2.0$ ), tautomers, and stereoisomers using Schrödinger's Ligand Preparation Wizard<sup>167</sup> with default settings. Grid generation and dockings into the ensemble of nine receptor structures were performed with Grid Generation application and the Virtual Screening Workflow script/Glide Docking SP (standard precision)<sup>160</sup>.

#### **3.1.2.6. Procedures of PubChem validation of the consensus scoring method**

The 5HT<sub>1R</sub> (3,131 fragments) and TAAR<sub>1</sub> (17,069 fragments) sets were prepared for docking with the ligand preparation method described above resulting in 37,167 generated structures. The screening datasets were subjected to docking into the nine previously selected structures using Schrödinger's Virtual Screening Workflow script<sup>160</sup>.

#### **3.1.2.7. Procedures of consecutive screening by combining FrAGS with FrACS**

The dataset was sorted by descending desirability scores, and filtered by a FrAGS  $\geq 5$  cut-off. The resulting 28,816 fragments are predicted to be promiscuous aminergic compounds based on their physicochemical properties. FrAGS filtering enriched the actives by 6.7-fold by retrieving 1,285 out of 2,183 actives. The resulting 28,816 fragments were then subjected to the default

Ligand Preparation method resulting in 87,176 molecules that were docked into the nine aminergic structures. In case of multiple generated states for a single compound, the Glide Scores of the best states were translated into rankings, and votes were assigned according to the Rank(x)% criteria.

### 3.1.2.8. Docking of fragment hits to the 5-HT<sub>6</sub>R homology model

In a recent study<sup>171</sup> Vass and colleagues have analysed a 5-HT<sub>6</sub>R homology model (built using h5-HT<sub>2B</sub>R in complex with ergotamine as template (PDB ID: 4IB4) by membrane-embedded molecular dynamics simulations. A chemometric analysis of enrichments from retrospective virtual screening studies resulted in nine MD-frames containing SB-742457 (8 in **Figure 3** and **Table Appendix 2**) selective 5-HT<sub>6</sub>R antagonist as ligand. We projected our primary hits (**compounds 36, 37, 39, 40** to docking studies using the nine 5-HT<sub>6</sub>R model structures. The experiments were performed with prepared ligands using Glide Single Precision<sup>160</sup> docking. Poses corresponding to the best Glide scores were compared to the orientation and interactions of SB-742457 (8) as reference.

### 3.1.3. Computational procedures for the study of spiro-oxindoles and indolines as 5-HT<sub>6</sub>R antagonists

A substructure search on the spiro[pyrrolidine-3,3'-oxindole] scaffold was performed on the downloaded Molecule purchasable database<sup>173,174</sup> (containing 5 million compounds), using Instant JChem<sup>146</sup> resulting in 887 derivatives. The molecules were prepared by Schrödinger's LigPrep<sup>167</sup> creating possible conformers, tautomers and protonation states by default settings. The ligands were then projected to single precision (SP) molecular docking analysis on a nine membered ensemble of molecular dynamics frames of the same 5-HT<sub>6</sub>R homology model<sup>171</sup> used in **Section 2.2.4., page 49**. The docking poses were filtered in a post-processing step keeping only binding modes forming hydrogen bonds towards Asp106<sup>3.32</sup>, Asn288<sup>6.55</sup> and optionally Ser193<sup>5.43</sup>, occupying a primary hydrophobic cleft defined by Trp281<sup>6.48</sup>, Phe284<sup>6.51</sup> and Phe285<sup>6.52</sup>. A secondary hydrophobic subpocket was defined by the pose filtering criteria as following: Val107<sup>3.33</sup>, Ala157<sup>4.56</sup>, Leu160<sup>4.59</sup>, Pro161<sup>4.60</sup>, Leu164<sup>4.63</sup>. The top ten pose-filtered ligands (compounds **41-50**) were selected for exploratory SAR-by-catalogue screening.

### 3.1.4. Computational procedures for the consensus-scoring based docking method targeting 5-HT<sub>2B</sub>R

#### 3.1.4.1. Compilation of the training set, NSFP-based machine learning model building

The active/inactive sets were selected based on binding affinity recalculated from various units reported in the ChEMBL to K<sub>i</sub> (actives: K<sub>i</sub> ≤ 500 nM,

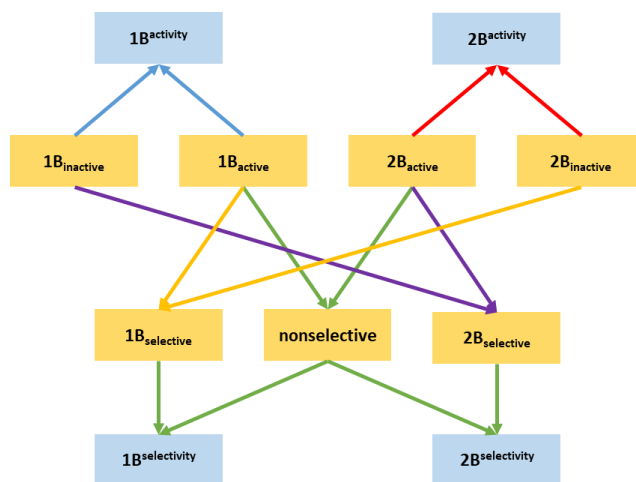
inactives:  $1000 \text{ nM} \leq K_i$ ) (see **Table 21**). Since our design concept was based on the role of the secondary binding pocket in selectivity, only compounds with 22 or more heavy atoms were considered, as they are more likely to bind to both orthosteric binding pocket (OBP) and SBP of 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R.

Receptor	Number of actives <sup>1</sup>	Number of inactives <sup>2</sup>	Number of selectives <sup>3</sup>
5-HT <sub>1B</sub>	858 (1011)	339 (477)	86
5-HT <sub>2B</sub>	478 (718)	259 (351)	33

**Table 21.** Active, inactive and selective 5-HT<sub>1B</sub> and 5-HT<sub>2B</sub> ligands retrieved from the ChEMBL with at least 22 heavy atoms. <sup>1</sup>  $K_i \leq 500 \text{ nM}$ ; <sup>2</sup>  $K_i \geq 1000 \text{ nM}$ ; <sup>3</sup> Classified as 5-HT<sub>2B</sub> selective for being active at 5-HT<sub>2B</sub>R, and inactive at 5-HT<sub>1B</sub>R; numbers in parenthesis are the total number of all actives without size-filtering

### 3.1.4.2. Building NSFP-based machine learning model and screening of the MCule database

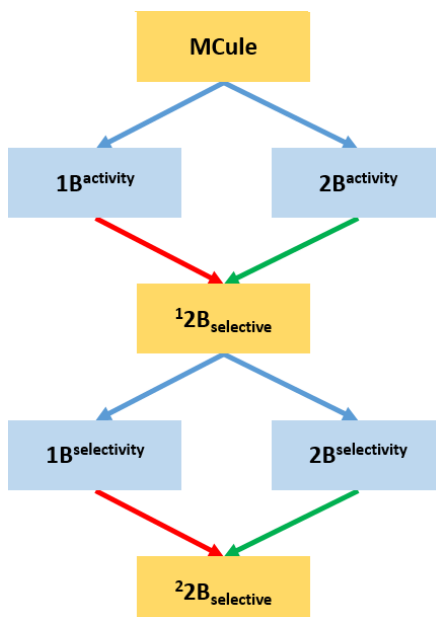
NSFP fingerprints were calculated for all compound sets using Klekota-Roth fingerprint (KRFP) substructure keys. A series of machine learning classifiers were created, aimed to properly discriminate compounds. Activity classifier for 2B (2B<sup>activity</sup>) was built using 2B<sup>active</sup> and 2B<sup>inactive</sup> compounds. Selectivity classifier (2B<sup>selectivity</sup>) for 2B was built using 2B<sup>selective</sup> and nonselective compounds in order to predict putative selectivity of screened compounds. The same procedure was applied to 1B with its selective sets, resulting in 1B<sup>selectivity</sup> and 2B<sup>selectivity</sup> classifiers. Altogether we developed two activity (for 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R) and two selectivity (for 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R) classifiers using known active, inactive, selective and non-selective 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R ligands. Final models were selected from the total of 4x117 classifiers based on the highest acquired Matthews correlation coefficient (MCC) values. These models were used for filtering the in-stock MCule database. The summary of the model development is shown in **Figure 25**.



**Figure 25.** Depiction of NSFP representations of compound sets (orange) and machine learning models (blue) constructed.

First, NSFP fingerprints have also been calculated for the entire MCule database<sup>173,174</sup> of commercially available compounds. Next, each compound was classified using the four machine learning models. If a compound's classification by  $2B^{\text{activity}}$  was positive (compound classified as 2B active) and classification by  $1B^{\text{activity}}$  was negative (compound classified as 1B inactive), the compound was treated as putative 2B selective compound ( ${}^12B^{\text{selective}}$  – First classification step marked in superscript). If the opposite was true (compound classified as 1B active and 2B inactive), the compound was regarded as putative 1B selective compound ( ${}^11B^{\text{selective}}$ ). Compounds achieving any other combination of classification results were disregarded from further research.

The second round of classification consisted of validating the compounds against the  $1B^{\text{selectivity}}$  and  $2B^{\text{selectivity}}$  selectivity classifiers. If an  ${}^12B^{\text{selective}}$  compound was classified as positive by the  $2B^{\text{selectivity}}$  and negative by the  $1B^{\text{selectivity}}$ , the compound was regarded as  ${}^22B^{\text{selective}}$  (Second classification step marked in superscript). In any other case, the compound was discarded. The summary of the prospective classification is presented in **Figure 26**.



**Figure 26.** Classification procedure for putative 2B selective compounds. Green arrows represent positive classification and red arrows represent negative classification.

### 3.1.4.3. Ligand and protein preparations for 5-HT<sub>2B</sub>R dockings

Ligands were prepared for docking using Schrodinger's LigPrep algorithm<sup>167</sup>, generating possible conformers and tautomers with the pH set to  $7.4 \pm 0.0$ , and minimizing their structures. Crystal structures of 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R were extracted from the PDB database<sup>232</sup> (PDB ID: 4IAQ and 4IAR<sup>203</sup> for 1B, and 4IB4<sup>204</sup> and 4NC3<sup>205</sup> for 2B). Proteins were prepared using PrepWiz software from Schrodinger<sup>233</sup> and transformed into a receptor grid using Schrodinger software package<sup>228</sup>, with the docking centred to the Asp<sup>3.32</sup> residue. Single precision dockings were launched with Schrodinger's Glide algorithm<sup>160</sup>. The putative 2B-selective sets were docked to crystal structures of both 5HT<sub>1B</sub> and 5HT<sub>2B</sub> receptors, and the number of reported poses was set to 10. As a post-processing step, poses were filtered for interactions with the SBP characteristic amino acids (see **Table 22**).

	Interactions	1B crystals	2B crystals
OBP		D129 <sup>3.32</sup>	D135 <sup>3.32</sup>
		T134 <sup>3.37</sup>	S139 <sup>3.36</sup>
		S212 <sup>5.42</sup>	
SBP			Q359 <sup>7.32</sup>
		M337 <sup>6.58</sup>	M218 <sup>5.39</sup>
		V201 <sup>ECL2.52</sup>	L209 <sup>ECL2.52</sup>

**Table 22.** All of the interactions considered during the pose filtering phase. The blue cells represent hydrogen bonds, and the red cells represent distance criterion ( $\leq 5.0$  Å between the compound and any atom of a residue). OBP: orthosteric binding pocket; SBP: secondary binding pocket.

Scoring of reported poses was performed using custom scoring function [39], taking into consideration the rankings of compounds based on the assigned Glide docking score. It compares the numerical sum of ranks of all poses acquired by a compound in both crystals of the target protein with those found in both crystals of the antitarget:

$$\Delta_{\max} = [\sum (\text{rank}_{\text{target}}) - \sum (\text{rank}_{\text{antitarget}})]$$

Since better rankings have lower numerical values (a pose with rank 1 is better than a pose with rank 5), if the compound's  $\Delta_{\max} < 0$ , it was considered target-selective. For further research the compounds with the lowest  $\Delta_{\max}$  were considered, as the difference in ranks between target and antitarget were the highest.

## 3.2. Biological assay descriptions

### 3.2.1. Cell-Based assays for 5-HT<sub>6</sub>R and CB<sub>1</sub>R

CHO-K1 cells expressing stably the mitochondrially targeted aequorin (luminescent indicator) and G $\alpha$ 16 were used for the assay<sup>42</sup>. The principle of the assay is that aequorin (derived from *Aequorea victoria*) complex emits blue light while binding Ca<sup>2+</sup> ions. The aequorin and G $\alpha$ 16 expressing CHO-K1 cells were transiently transfected with plasmid harbouring the gene expressing human 5-HT<sub>6</sub> or CB<sub>1</sub> respectively, using Roche X-treme GENE HP DNA Transfection Reagent. Cells were grown for 48 hours after transient transfection. One day before the assay the culture medium was changed to antibiotics free medium and the cells were grown for an additional 6 hours, then the cells were detached by gentle flushing with PBS/0.5 mM EDTA, recovered by centrifugation and resuspended at 1x10<sup>6</sup> cells/ml density in assay medium (DMEM/HAM's F12 with HEPES, without phenol red (Sigma, D6434), + 0.1% BSA + 2mM glutamine) in a Falcon tube. Coelenterazine h (LifeTechnologies C6780) was added at a final concentration of 5  $\mu$ M and the cells were incubated overnight at room temperature using constant shaking protected from light. Before the measurement the cells were diluted 2 fold in assay medium and incubated for 60min.

The measurement was executed in Optiplate TM-96 plates (PerkinElmer 6005290) and the luminescent emission was detected by an AppliscanTM (Thermo) plate-reader. Cell suspension (45 $\mu$ l)/well + 5  $\mu$ l antagonist (diluted in assay medium) was preincubated in the plate; the reaction was initiated with the addition of agonist EMD 386088 (50 $\mu$ l) at a concentration of 50nM. In case of CB<sub>1</sub> measurement agonist CP55940 (50 $\mu$ l) was added at a concentration of 25nM.

The time lapse curves of receptor activation signals were recorded well by well. After detection of the base line (8 sec) the agonist was injected in one well and the change of intracellular Ca<sup>2+</sup> level released due to receptor activation was monitored for 45 seconds. Peak luminescence signal was used for evaluation of the measurements. For assay development and validation SB-271046<sup>47</sup> was used as antagonist control (Supporting Information, Figures S4 and S5) for 5-HT<sub>6</sub>

assay, and AM251<sup>48</sup> was used as antagonist control (Supporting Information Figures S6 and S7) respectively. All compounds were screened at a fixed concentration of 50  $\mu$ M. The hit compounds with better than 100 % inhibitory levels were selected for IC<sub>50</sub> measurements.

### **3.2.2. Radioligand binding assays for 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R, 5-HT<sub>7</sub>R**

#### **3.2.2.1. Cell culture and preparation of cell membranes**

HEK293 cells with stable expression of human serotonin 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>, 5-HT<sub>6</sub> or 5-HT<sub>7</sub> receptor (all prepared with the use of Lipofectamine 2000) were maintained at 37 °C in a humidified atmosphere with 5% CO<sub>2</sub> and were grown in Dulbecco's Modified Eagle Medium containing 10% dialysed foetal bovine serum and 500 mg/ml G418 sulphate. For membrane preparations, cells were subcultured in 10 cm diameter dishes, grown to 90% confluence, washed twice with prewarmed to 37 °C phosphate buffered saline (PBS) and were pelleted by centrifugation (200 g) in PBS containing 0.1 mM EDTA and 1 mM dithiothreitol. Prior to membrane preparations pellets were stored at -80 °C.

#### **3.2.2.2. Radioligand binding assays**

Cell pellets were thawed and homogenized in 20 volumes of assay buffer using an Ultra Turrax tissue homogenizer and centrifuged twice at 35 000 g for 20 min at 4 °C, with incubation for 15 min at 37 °C in between. The composition of the assay buffers was as follows: for 5-HT<sub>1A</sub>R: 50 mM Tris-HCl, 0.1 mM EDTA, 4 mM MgCl<sub>2</sub>, 10  $\mu$ M pargyline and 0.1% ascorbate; for 5-HT<sub>2A</sub>R: 50 mM Tris-HCl, 0.1 mM EDTA, 4 mM MgCl<sub>2</sub> and 0.1% ascorbate; for 5-HT<sub>6</sub>R: 50 mM Tris-HCl, 0.5 mM EDTA and 4 mM MgCl<sub>2</sub>, for 5-HT<sub>7</sub>R: 50 mM Tris-HCl, 4 mM MgCl<sub>2</sub>, 10  $\mu$ M pargyline and 0.1% ascorbate. All assays were incubated in a total volume of 200  $\mu$ l in 96-well microtiter plates for 1 h at 37 °C, except for 5-HT<sub>1A</sub>R and 5-HT<sub>2A</sub>R that were incubated at room temperature for 1 h and 1.5 h, respectively. The process of equilibration was terminated by rapid filtration through Unifilter plates with a 96-well cell harvester and radioactivity retained on the filters was quantified on a Microbeta plate reader (PerkinElmer).

For displacement studies the assay samples contained as radioligands: 1.5 nM [<sup>3</sup>H]-8-OH-DPAT (135.2 Ci/mmol, K<sub>D</sub> = 1.5 nM) for 5-HT<sub>1A</sub>R; 2 nM [<sup>3</sup>H]-Ketanserin (53.4 Ci/mmol, K<sub>D</sub> = 2 nM) for 5-HT<sub>2A</sub>R; 2 nM [<sup>3</sup>H]-LSD (83.6 Ci/mmol, K<sub>D</sub> = 2 nM) for 5-HT<sub>6</sub>R or 0.6 nM [<sup>3</sup>H]-5-CT (39.2 Ci/mmol, K<sub>D</sub> = 0.6 nM) for 5-HT<sub>7</sub>R. Non-specific binding was defined with 10  $\mu$ M of 5-HT in 5-HT<sub>1A</sub>R and 5-HT<sub>7</sub>R binding experiments, whereas 10  $\mu$ M of chlorpromazine or 10  $\mu$ M of methiothepine were used in 5-HT<sub>2A</sub>R and 5-HT<sub>6</sub>R assays, respectively. Each compound was tested in triplicate at 7 concentrations (10<sup>-10</sup>-10<sup>-4</sup> M).

Reference compounds for 5-HT<sub>1A</sub>R: buspirone  $K_i = 32.2 \text{ nM} \pm 2.9$ ; for 5-HT<sub>2A</sub>R: olanzapine  $K_i = 5.6 \text{ nM} \pm 1.1$ ; for 5-HT<sub>6</sub>R: olanzapine  $K_i = 8.8 \text{ nM} \pm 1.3$ ; and for 5-HT<sub>7</sub>R: 5-CT  $K_i = 0.8 \text{ nM} \pm 0.2$ . The inhibition constants ( $K_i$ ) were calculated from the Cheng-Prusoff equation<sup>234</sup>. Results were expressed as means of at least two separate experiments.

### 3.2.3. Radioligand binding assays for 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R

All tested compounds were purchased via MCule, Inc. Compounds were tested for their ability to displace radioligand from the membrane of 5-HT<sub>1B</sub>R and 5-HT<sub>2B</sub>R expressing cells. First test was performed with 10  $\mu\text{M}$  concentration of a compound, and the inhibition percentage was calculated. If the value was above 50%, a full affinity screening was performed, and the data was gathered as  $K_i$  with units reported as nM.

#### 3.2.3.1. Competition binding in human 5-HT<sub>1B</sub> receptor

Serotonin 5-HT<sub>1B</sub>R competition binding experiments were carried out in a polypropylene 96-well plate. In each well was incubated 5  $\mu\text{g}$  of membranes from HeLa-5-HT<sub>1B</sub> cell line prepared in our laboratory (Lot: A001/14-11-2011, protein concentration = 3179  $\mu\text{g}/\text{ml}$ ), 1.5 nM [<sup>3</sup>H]-GR125743 (83.9 Ci/mmol, 0.1 mCi/ml, Perkin Elmer NET1172100UC,  $K_D = 0.74 \text{ nM}$ ) and compounds studied and standard. Non-specific binding was determined in the presence of GR55562 10  $\mu\text{M}$  (TOCRIS 1054) and total binding was determined in the absence of any unlabeled compound. The reaction mixture ( $V_t$ : 250  $\mu\text{l}$ /well) was incubated at 25°C for 90 min, 200  $\mu\text{L}$  was transferred to GF/C 96-well plate (Millipore, Madrid, Spain) pretreated with 0.5% of PEI and treated with binding buffer (Tris-HCl 50mM, EDTA 1mM, MgCl<sub>2</sub> 10mM, pH=7.4), then it was filtered and washed four times with 250  $\mu\text{l}$  wash buffer (Tris-HCl 50mM, pH=7.4), before measuring in a microplate beta scintillation counter (Microbeta Trilux, PerkinElmer, Madrid, Spain). 5-Carboxytryptamine was included as a reference compound in all the assays. Compounds were first tested at 10 micromolar and those compounds showing a percentage of displacement of specific binding higher than 50% were classified as active compounds and  $K_i$  values were determined by means of concentration-response curves

#### 3.2.3.2. Competition binding in human 5-HT<sub>2B</sub> receptor

Serotonin 5-HT<sub>2B</sub>R competition binding experiments were carried out in a polypropylene 96-well plate. In each well was incubated 5  $\mu\text{g}$  of membranes from CHO-5-HT<sub>2B</sub> cell line prepared in our laboratory (Lot: A003/27-03-2012, protein concentration=3431  $\mu\text{g}/\text{ml}$ ), 1 nM [<sup>3</sup>H]-LSD (82.4 Ci/mmol, 1 mCi/ml, PerkinElmer NET638250UC,  $K_D = 0.57 \text{ nM}$ ) and compounds studied and standard. Non-specific binding was determined in the presence of 5-HT 50  $\mu\text{M}$  (Sigma H9523) and total binding was determined in the absence of any

unlabelled compound. The reaction mixture (Vt: 250  $\mu$ l/well) was incubated at 37°C for 30 min, 200  $\mu$ L was transferred to GF/C 96-well plate (Millipore, Madrid, Spain) pretreated with 0.5% of PEI and treated with binding buffer (Tris-HCl 50mM, Ascorbic acid 0.1%, CaCl<sub>2</sub> 4mM, pH=7.4), after was filtered and washed four times with 250  $\mu$ l wash buffer (Tris-HCl 50mM, pH=7.4), before measuring in a microplate beta scintillation counter (Microbeta Trilux, PerkinElmer, Madrid, Spain). Methysergide was included as a reference compound in all the assays. Compounds were first tested at 10 micromolar concentration and those compounds showing a percentage of displacement of specific binding higher than 50% were classified as active compounds and K<sub>i</sub> values were determined by means of concentration-response curves.

### 3.3. Synthetic procedures

All chemical reagents used were purchased from commercial chemical suppliers. The NMR experiments were performed at 500 MHz (<sup>1</sup>H) on a Varian VNMR SYSTEM spectrometer. Chemical shifts are referenced to the residual solvent signals, 2.50 ppm for <sup>1</sup>H in DMSO-d<sub>6</sub> and 7.28 ppm for <sup>1</sup>H in CDCl<sub>3</sub>. The MS measurements were performed on Shimadzu LCMS2020 LC/MS system. Flash chromatography was performed using Teledyne ISCO CombiFlash Lumen+ Rf. Purifications by preparative-HPLC were performed with Hanbon NS4205 Binary high pressure semi-preparative HPLC. Thin-layer chromatography was performed on TLC Silica Gel 60 F<sup>254</sup>. High resolution mass spectrometric measurements were performed using a Q-TOF Premier mass spectrometer (Waters Corporation, Milford, MA, USA) in positive electrospray ionization mode. Screening compounds were purchased from Mcule Inc.

Synthetic procedures for compounds reported in K3 publication (*Molecules* **2017**, *22*, 2221) are referred with the corresponding page and compound number.

#### 3.3.1. Synthesis of spiro-oxindoles and indolines as new 5-HT<sub>6</sub>R antagonists

Benzyl 2-oxospiro[indoline-3,3'-pyrrolidine]-10-carboxylate (**55**)

See compound **27** in K3 publication on page 14.

Benzyl 2-oxo-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine]-10-carboxylate (**56**)

See compound **28** in K3 publication on page 14.

1-(Phenylsulfonyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**51**)

See compound **17** in K3 publication on page 15.

1'-(Phenylsulfonyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**52**)

See compound **18** in K3 publication on page 15.

1'-Benzyl-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**67a**)

See compound **39a** in K3 publication on page 15.

1'-(Pyridin-4-ylmethyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**65b**)

See compound **37b** in K3 publication on page 16.

1-(Phenylsulfonyl)-10-(pyridin-4-ylmethyl)spiro[indoline-3,3'-pyrrolidine] (**67b**)

See compound **39b** in K3 publication on page 16.

1'-Methyl-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**68a**)

See compound **40a** in K3 publication on page 17.

1'-Ethyl-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**68b**)

See compound **40b** in K3 publication on page 17.

1'-(2-Fluorobenzyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**70a**)

See compound **42a** in K3 publication on page 18.

1'-(2-Fluorobenzyl)-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**72a**)

See compound **44a** in K3 publication on page 18.

1'-(3-Fluorobenzyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**70b**)

See compound **42b** in K3 publication on page 19.

1'-(3-Fluorobenzyl)-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**72b**)

See compound **44b** in K3 publication on page 19.

1'-(4-Fluorobenzyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**70c**)

See compound **42c** in K3 publication on page 20.

1'-(4-Fluorobenzyl)-1-(phenylsulfonyl)spiro[indoline-3,3'-pyrrolidine] (**72c**)

See compound **44c** in K3 publication on page 20.

1'-Benzyl-1-((2-fluorophenyl)sulfonyl)spiro[indoline-3,3'-pyrrolidine] (**74a**)

See compound **46a** in K3 publication on page 20.

1'-Benzyl-1-((3-fluorophenyl)sulfonyl)spiro[indoline-3,3'-pyrrolidine] (**74b**)

See compound **46b** in K3 publication on page 21.

1'-Benzyl-1-((4-fluorophenyl)sulfonyl)spiro[indoline-3,3'-pyrrolidine] (**74c**)

See compound **46c** in K3 publication on page 21.

### 3.3.2. Synthesis of spiro-oxindoles as new 5-HT<sub>7</sub>R antagonists

#### *Synthesis of the piperazine analogues (76-81, 85)*

1'-(2-(4-(4-chlorophenyl)piperazin-1-yl)ethyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**76**)

To the solution of 4-chlorophenylpiperazine (**88a**) (39 mg, 0.2 mmol) in acetonitrile (50 mL) was added potassium carbonate (55 mg, 0.4 mmol) and 2-bromoethanol (**89a**) (0.028 mL, 0.4 mmol), and the mixture was stirred overnight at room temperature. The reaction mixture was quenched with 20 ml brine, and was extracted three times with ethyl-acetate (5 mL x 3). The collected organic phases were dried over sodium-sulphate, and evaporated to give the crude 3-(4-(4-chlorophenyl)piperazin-1-yl)propan-1-ol (**90a**). The crude alcohol was dissolved in dry dichloromethane (10 mL), and trimethylamine (0.032 mL, 0.2293 mmol) was added, followed by the addition of mesyl-chloride (0.015 mL, 0.1911 mmol) at room temperature, and the reaction mixture was then refluxed for 2h. The mixture was quenched by brine (20 ml), extracted three times with dichloromethane (5 mL x 3), and washed with distilled water (10 mL). The collected organic phases were dried over sodium-sulphate, and evaporated to give the crude mesylate 3-(4-(4-chlorophenyl)piperazin-1-yl)propyl methanesulfonate (**91a**). Next the crude mesylate (**91a**) was dissolved in acetonitrile (10 mL), then tetrahydro- $\beta$ -carboline (**53**) (19 mg, 0.1098 mmol) and potassium carbonate (46 mg, 0.3294 mmol) were added. The suspension was refluxed overnight. The reaction mixture was quenched with brine (15 mL), and was extracted three times with ethyl-acetate (5 mL x 3). The collected organic phases were dried over sodium-sulphate, and evaporated to give the crude alkylated intermediate 2-(3-(4-(4-chlorophenyl)piperazin-1-yl)propyl)-2,3,4,9-tetrahydro-1H-pyrido[3,4-b]indole (**92a**). Intermediate **92a** was dissolved in [tetrahydrofuran]:[water]:[glacial acetic acid] = 3:3:1 (3-3-1 mL), and *N*-bromosuccinimide (11 mg, 0.0583 mmol) was added in small portions at 0°C. The reaction mixture was stirred for 1.5 hour at 0°C. Afterwards the reaction mixture was quenched at 0 °C with 15 mL concentrated sodium-carbonate solution dropwise (until gas-evolution ceased) and was stirred for 0.5 h, allowing to warm up to room temperature. The mixture was extracted three times using ethyl acetate (3 x 5 mL). The combined organic phases were washed with distilled water (10 mL). The collected organic phases were dried over sodium-sulfate, filtered and evaporated under reduced pressure. The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 3.1 mg. <sup>1</sup>H-NMR (500 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  ppm 10.30 (s, 1H, NH), 7.31 (d, *J* = 6.9 Hz, 1H, ArH-4), 7.20 (d, *J* = 8.6 Hz, 2H, ArH-3'',5''), 7.13 (m, 1H, ArH-6), 6.95-6.91 (m, 3H, ArH-5,2'',6''), 6.81 (d, *J* = 7.7 Hz, 1H, ArH-7), 3.15-3.12

(m, 1H, pyrrolidine H-2'), 3.10 (m, 4H, piperazine), 2.93 (m, 1H, pyrrolidine H-2'), 2.83-2.81 (m, 1H, pyrrolidine H-5'), 2.67-2.60 (m, 3H, ethylene-linker CH<sub>2</sub>, pyrrolidine H-5'), 2.60-2.58 (m, 1H, ethylene-linker CH<sub>2</sub>), 2.55 (m, 4H, piperazine), 2.51 (m, 1H, ethylene-linker CH<sub>2</sub>), 2.17-2.12 (m, 1H, pyrrolidine H-4'), 1.89-1.83 (m, 1H, pyrrolidine H-4'). HR-MS (ESI+): *m/z* [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>28</sub>N<sub>4</sub>OCl: 411.1952; found: 411.1960.

1'-[3-[4-(4-chlorophenyl)piperazin-1-yl]propyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**77**)

Compound **77** was prepared using the same procedure as used in case of **76**, starting from 4-chlorophenylpiperazine (**88a**) (196 mg, 1 mmol), and using 3-bromopropanol (**89b**) (0.181 mL, 2 mmol) for the first alkylation step (instead of 2-bromoethanol **89a**). Using the following quantities of reagents: potassium carbonate for the alkylation with 3-bromopropanol (276 mg, 2 mmol), triethylamine (0.145 mL, 1.044 mmol), mesyl chloride (0.0675 mL, 0.87 mmol), tetrahydro-β-carboline (**53**) (112 mg, 0.6525 mmol), potassium carbonate for alkylation with the mesylate (108 mg, 0.783 mmol), *N*-bromosuccinimide (11 mg, 0.06128 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 5.4 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>) δ ppm 8.42 (s, 1H, NH), 7.40 (d, *J* = 7.4 Hz, 1H, ArH-4), 7.25-7.21 (m, 3H, ArH-6, ArH-3'',5''), 7.07 (t, *J* = 7.2 Hz, 1H, ArH-5), 6.98 (d, *J* = 8.9 Hz, 2H, ArH-2'',6''), 6.90 (d, *J* = 7.6 Hz, 1H, ArH-7), 3.49-3.42 (m, 1H, pyrrolidine H-2'), 3.41-3.35 (m, 4H, piperazine), 3.34-3.26 (m, 2H, pyrrolidine H-2'), 3.17-3.15 (m, 1H, pyrrolidine H-5'), 3.12-3.08 (m, 2H, propylene-linker CH<sub>2</sub>), 3.06-3.03 (m, 4H, piperazine), 2.97-2.95 (m, 2H, propylene-linker CH<sub>2</sub>), 2.41-2.36 (m, 1H, pyrrolidine H-4'), 2.30-2.24 (m, 1H, pyrrolidine H-4'), 2.06 (m, 1H, propylene-linker CH<sub>2</sub>), 1.95-1.90 (m, 1H, propylene-linker CH<sub>2</sub>). HR-MS (ESI+): *m/z* [M + H]<sup>+</sup> calcd. for C<sub>24</sub>H<sub>30</sub>N<sub>4</sub>OCl: 425.2115; found: 425.2108.

1'-[2-(4-phenylpiperazin-1-yl)ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**78**)

Derivative **78** was prepared using the same procedure as used in case of **76**, starting from 1-phenylpiperazine (**88b**) (185 mg, 1.1420 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.162 mL, 2.2840 mmol), potassium carbonate for alkylation with 2-bromoethanol (315 mg, 2.284 mmol), triethylamine (0.158 mL, 1.1418 mmol), mesyl chloride (0.074 mL, 0.9515 mmol), tetrahydro-β-carboline (**53**) (129 mg, 0.7525 mmol), potassium carbonate for alkylation with mesylate (312 mg, 2.2575 mmol), *N*-bromosuccinimide (141 mg, 0.7916 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 4.8 mg. <sup>1</sup>H-NMR (500 MHz,

methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.42 (s, 1H, NH), 7.39 (d, *J* = 7.5 Hz, 1H, ArH-4), 7.26-7.23 (m, 3H, ArH-6, ArH-2'',6''), 7.06 (t, *J* = 7.5 Hz, 1H, ArH-5), 6.99 (t, *J* = 8.1 Hz, 2H, ArH-3'',5''), 6.91-6.84 (m, 2H, ArH-7, ArH-4''), 3.34-3.25 (m, 8H, piperazine), 3.13-3.10 (m, 2H, pyrrolidine H-2'), 3.07 (t, *J* = 4.8 Hz, ethylene-linker CH<sub>2</sub>), 3.02-2.99 (m, 2H, pyrrolidine H-5'), 2.93 (t, *J* = 4.8 Hz, ethylene-linker CH<sub>2</sub>), 2.41-2.35 (m, 1H, pyrrolidine H-4'), 2.23-2.17 (m, 1H, pyrrolidine H-4'). HR-MS (ESI<sup>+</sup>): *m/z* [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>29</sub>N<sub>4</sub>O: 377.2341; found: 377.2341.

1'-[2-[4-(3-chlorophenyl)piperazin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**79**)

**79** was prepared using the same procedure as used in case of **76**, starting from 3-chlorophenylpiperazine (**88c**) (28 mg, 0.1421 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.020 mL, 0.2842 mmol), potassium carbonate for the alkylation with 2-bromoethanol (39 mg, 0.2842 mmol), triethylamine (0.035 mL, 0.25 mmol), mesyl chloride (0.016 mL, 0.2083 mmol), tetrahydro- $\beta$ -carboline (**53**) (22 mg, 0.1258 mmol), potassium carbonate for the alkylation with the mesylate (52 mg, 0.3774 mmol), *N*-bromosuccinimide (20 mg, 0.1124 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 1.1 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.46 (s, 1H, NH), 7.27 (d, *J* = 7.9 Hz, 1H, ArH-7), 7.18 (t, *J* = 7.3 Hz, 1H, ArH-5''), 7.13 (t, *J* = 7.9 Hz, 1H, ArH-6), 7.06 (t, *J* = 7.8 Hz, 1H, ArH-5), 6.93-6.87 (m, 2H, ArH-2'',4'') 6.79-6.77 (m, 2H, ArH-6'',7), 3.24-3.19 (m, 8H, piperazine), 2.95-2.87 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.76-2.67 (m, 6H, pyrrolidine H-5' + ethylene-linker CH<sub>2</sub>), 2.35-2.28 (m, 1H, pyrrolidine H-4'), 2.10-2.03 (m, 1H, pyrrolidine H-4'). HR-MS (ESI<sup>+</sup>): *m/z* [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>28</sub>N<sub>4</sub>OCl: 411.1952; found: 411.1947.

1'-[2-[4-(2-chlorophenyl)piperazin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**80**)

**80** was prepared using the same procedure as used in case of **76**, starting from 2-chlorophenylpiperazine hydrochloride (**88d**) (47 mg, 0.2 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.142 mL, 0.4 mmol), potassium carbonate for the alkylation with 2-bromoethanol (55 mg, 0.4 mmol), triethylamine (0.018 mL, 0.1299 mmol), mesyl chloride (0.008 mL, 0.1083 mmol), tetrahydro- $\beta$ -carboline (**53**) (10 mg, 0.05643 mmol), potassium-carbonate for the alkylation with the mesylate (23 mg, 0.1693 mmol), *N*-bromosuccinimide (9 mg, 0.0508 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 3.0 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.48 (s, 1H, NH), 7.39 (m, 2H, ArH-4, ArH-3''), 7.28 (t, *J* = 7.8 Hz, 1H, ArH-4''), 7.22 (t, *J* = 7.8 Hz, 1H, ArH-6), 7.18 (d, *J* = 7.9 Hz, 1H, ArH-6''),

7.07-7.02 (m, 2H, ArH-5, ArH-5''), 6.91 (d,  $J = 7.3$  Hz, 1H, ArH-7), 3.26-3.17 (m, 8H, piperazine), 3.10-3.05 (m, 6H, pyrrolidine H-2', H-5' + ethylene-linker CH<sub>2</sub>), 2.98 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.40-2.34 (m, 1H, pyrrolidine H-4'), 2.20-2.16 (m, 1H, pyrrolidine H-4'). HR-MS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>28</sub>N<sub>4</sub>OCl: 411.1952; found: 411.1951.

1'-{2-[4-(3,4-dichlorophenyl)piperazin-1-yl]ethyl}-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**81**)

**81** was prepared using the same procedure as used in case of **79**, starting from 3,4-dichlorophenylpiperazine (**88e**) (46 mg, 0.2 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.028 mL, 0.4 mmol), potassium carbonate for the alkylation with 2-bromoethanol (55 mg, 0.4 mmol), triethylamine (0.020 mL, 0.1445 mmol), mesyl chloride (0.009 mL, 0.1204 mmol), tetrahydro- $\beta$ -carboline (**53**) (20 mg, 0.1180 mmol), potassium carbonate used for the alkylation with the mesylate (49 mg, 0.3540 mmol), *N*-bromosuccinimide (21 mg, 0.1168 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 1.2 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.46 (s, 1H, NH), 7.41 (d,  $J = 6.9$  Hz, 1H, ArH-4), 7.36 (d,  $J = 7.3$  Hz, 1H, ArH-5''), 7.31 (t,  $J = 9.3$  Hz, 1H, ArH-6''), 7.26-7.20 (m, 1H, ArH-6), 7.08-7.05 (m, 2H, ArH-5, ArH-2''), 6.91 (d,  $J = 8.3$  Hz, 1H, ArH-7), 3.81-3.71 (m, 2H, pyrrolidine H-2'), 3.65-3.63 (m, 1H, pyrrolidine H-5'), 3.56-3.54 (m, 1H, piperazine), 3.21 (m, 2H, piperazine), 3.16 (m, 3H, pyrrolidine H-5', ethylene-linker CH<sub>2</sub>), 3.03 (m, 1H, piperazine), 2.84 (m, 3H, ethylene-linker CH<sub>2</sub>, piperazine), 2.75-2.69 (m, 2H, piperazine), 2.39-2.33 (m, 2H, pyrrolidine H-4', piperazine), 2.29-2.24 (m, 1H, pyrrolidine H-4'). HR-MS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>27</sub>N<sub>4</sub>OCl<sub>2</sub>: 445.1562; found: 445.1559.

1'-{2-[4-(4-fluorophenyl)piperazin-1-yl]ethyl}-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**85**)

**85** was prepared using the same procedure as used in case of **76**, starting from 4-fluorophenylpiperazine (**88f**) (63 mg, 0.3512 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.050 mL, 0.7024 mmol), potassium carbonate used for the alkylation with 2-bromoethanol (97 mg, 0.7024 mmol), trimethylamine (0.057 mL, 0.4111 mmol), mesyl chloride (0.027 mL, 0.3426 mmol), tetrahydro- $\beta$ -carboline (**53**) (38 mg, 0.2235 mmol), potassium carbonate used for the alkylation with the mesylate (93 mg, 0.6705 mmol), *N*-bromosuccinimide (21 mg, 0.1189 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 2.8 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.49 (s, 1H, NH), 7.41 (d,  $J = 7.1$  Hz, ArH-4), 7.22 (t,  $J = 7.9$  Hz, ArH-6), 7.05 (t,  $J = 7.9$  Hz, 1H, ArH-5), 6.99 (d,  $J = 6.6$  Hz, 4H, ArH-

2'',3''5'',6''), 6.91 (d,  $J = 7.9$  Hz, 1H, ArH-7), 3.24-3.22 (m, 4H, piperazine), 3.16-3.08 (m, 4H, pyrrolidine H-2', H-4'), 3.04-3.02 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.95 (m, 4H, piperazine), 2.90-2.88 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.38-2.33 (m, 1H, pyrrolidine H-4'), 2.19-2.16 (m, 1H, pyrrolidine H-4'). HR-MS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>23</sub>H<sub>28</sub>N<sub>4</sub>O: 395.2247; found: 395.2249.

*Synthesis of the piperidine-analogues, and of the 5,6-dihydropyridine-analogue (82-84)*

1'-[2-[4-(4-chlorophenyl)piperidin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**82**)

**82** was prepared using the same procedure as used in case of **76**, starting from 4-(4-chlorophenyl)piperidine (**93a**) (49 mg, 0.25 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.043 mL, 0.5 mmol), potassium carbonate used for the alkylation with 2-bromoethanol (69 mg, 0.5 mmol), triethylamine (0.027 mL, 0.1958 mmol), mesyl chloride (0.013 mL, 0.1632 mmol), tetrahydro- $\beta$ -carboline (**53**) (24 mg, 0.1384 mmol), potassium carbonate used for the alkylation with the mesylate (57 mg, 0.4153 mmol), *N*-bromosuccinimide (26 mg, 0.1470 mmol). The crude product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 1.3 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.47 (s, 1H, NH), 7.37 (d,  $J = 7.39$  Hz, 1H, ArH-4), 7.31-7.26 (m, 3H, ArH-6, ArH-3'',5''), 7.05 (t,  $J = 7.3$  Hz, 1H, ArH-5), 6.97 (d,  $J = 8.4$  Hz, 2H, ArH-2'',6''), 6.91 (d,  $J = 7.5$  Hz, 1H, ArH-7), 3.82-3.80 (m, 1H, pyrrolidine H-2'), 3.76-3.73 (m, 1H, pyrrolidine H-2'), 3.67-3.63 (m, 2H, ethylene CH<sub>2</sub>), 3.19-3.15 (m, 2H, pyrrolidine H-5'), 3.11-3.05 (m, 2H, piperidine CH<sub>2</sub>), 2.98-2.95 (m, 1H, piperidine CH), 2.91-2.87 (m, 2H, piperidine CH<sub>2</sub>), 2.69-2.65 (m, 1H, pyrrolidine H-4'), 2.57-2.47 (m, 2H, ethylene CH<sub>2</sub>), 2.38-2.33 (m, 1H, pyrrolidine H-4'), 2.00-1.82 (m, 4H, piperidine 2 CH<sub>2</sub>). HR-MS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>24</sub>H<sub>29</sub>N<sub>3</sub>OCl: 410.1999; found: 410.2003.

1'-[2-[4-(4-fluorophenyl)piperidin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**83**)

**83** was prepared using the same procedure as used in case of **76**, starting from 4-(4-fluorophenyl)piperidine (**93b**) (86 mg, 0.4 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.057 mL, 0.8 mmol), potassium carbonate used for alkylation with 2-bromoethanol (110 mg, 0.8 mmol), trimethylamine (0.052 mL, 0.377 mmol), mesyl chloride (0.024 mL, 0.314 mmol), tetrahydro- $\beta$ -carboline (**53**) (41 mg, 0.239 mmol), potassium carbonate used for alkylation with the mesylate (99 mg, 0.718 mmol), *N*-bromosuccinimide (42 mg, 0.236 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 1.0 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.47 (s, 1H, NH), 7.36 (d,  $J = 7.6$  Hz, 1H, ArH-4), 7.26 (t,  $J =$

7.6 Hz, 1H, ArH-6), 7.08-7.03 (m, 3H, ArH-5, ArH-3'',5''), 6.94 (d,  $J = 8.4$  Hz, 2H, ArH-2'',6''), 6.91 (d,  $J = 7.5$  Hz, 1H, ArH-7), 3.87-3.83 (m, 1H, pyrrolidine H-2'), 3.76-3.72 (m, 1H, pyrrolidine H-2'), 3.65-3.57 (m, 2H, ethylene CH<sub>2</sub>), 3.20-3.13 (m, 2H, pyrrolidine H-5'), 3.09-3.01 (m, 2H, piperidine CH<sub>2</sub>), 2.99-2.95 (m, 2H, piperidine 2 CH<sub>2</sub>), 2.89-2.85 (m, 1H, piperidine CH), 2.75-2.67 (m, 1H, pyrrolidine H-4'), 2.62-2.55 (m, 2H, ethylene CH<sub>2</sub>), 2.43-2.33 (m, 1H, pyrrolidine H-4'), 2.04-1.85 (m, 4H, piperidine 2 CH<sub>2</sub>). HR-MS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>24</sub>H<sub>29</sub>N<sub>3</sub>O: 394.2295; found: 394.2290.

1'-(2-(4-(4-fluorophenyl)-5,6-dihydropyridin-1(2H)-yl)ethyl)spiro[indoline-3,3'-pyrrolidin]-2-one (**84**)

**84** was prepared using the same procedure as used in case of **76**, starting from 4-(4-fluorophenyl)-5,6-dihydropyridine hydrochloride (**93c**) (85mg, 0.4 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.057 mL, 0.8 mmol), potassium carbonate used for the alkylation with 2-bromoethanol (110 mg, 0.8 mmol), trimethylamine (0.061 mL, 0.4423 mmol), mesyl chloride (0.029 mL, 0.3686 mmol), tetrahydro- $\beta$ -carboline (**53**) (40 mg, 0.2316 mmol), potassium carbonate used for the alkylation with the mesylate (96 mg, 0.6948 mmol), *N*-bromosuccinimide (36 mg, 0.2014 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 1.0 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>)  $\delta$  ppm 8.42 (s, 1H, NH), 7.49-7.46 (m, 2H, ArH-2'',6''), 7.40 (d,  $J = 7.6$  Hz, 1H, ArH-4), 7.20 (t,  $J = 8.0$  Hz, 1H, ArH-6), 7.09-7.03 (m, 3H, ArH-3'',5'', ArH-5), 6.90 (d,  $J = 7.3$  Hz, 1H, ArH-7), 6.10 (m, 1H, 5,6-dihydropyridine unsaturated), 3.59 (m, 2H, ethylene-linker CH<sub>2</sub>), 3.20-3.17 (m, 2H, pyrrolidine H-2'), 3.10-3.02 (m, 5H, 5,6-dihydropyridine saturated, pyrrolidine H-5'), 2.99-2.93 (m, 3H, 5,6-dihydropyridine saturated, pyrrolidine H-5'), 2.73 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.36-2.31 (m, 1H, pyrrolidine H-4'), 2.15-2.08 (m, 1H, pyrrolidine H-4'). HRMS (ESI+):  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>24</sub>H<sub>27</sub>N<sub>3</sub>O: 392.2138; found: 392.2138.

*Synthesis of phenoxy-, and amide derivatives (86-87)*

1'-[2-[4-(4-chlorophenoxy)piperidin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**86**)

**86** was prepared using the same procedure as used in case of **76**, starting from 4-(4-chlorophenoxy)piperidine (**93d**) (64 mg, 0.3 mmol). Using the following quantities of reagents: 2-bromoethanol (**89a**) (0.043 mL, 0.6 mmol), potassium carbonate used for the alkylation with 2-bromoethanol (83 mg, 0.6 mmol) trimethylamine (0.034 mL, 0.2447 mmol), mesyl chloride (0.016 mL, 0.2039 mmol), tetrahydro- $\beta$ -carboline (**53**) (26 mg, 0.1532 mmol), potassium carbonate used for the alkylation with the mesylate (63 mg, 0.4596 mmol), *N*-

bromosuccinimide (26 mg, 0.1439 mmol). The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 2.4 mg. <sup>1</sup>H-NMR (500 MHz, methanol-*d*<sub>4</sub>) δ ppm 8.50 (s, 1H, NH), 7.28 (m, 4H, ArH-6, 4, 3'', 5''), 7.07-7.03 (m, 1H, ArH-5), 6.99 (d, *J* = 8.7 Hz, 1H, ArH-7), 6.95-6.90 (m, 2H, ArH-2'', 6''), 4.01-3.97 (m, 1H, piperidine CH), 3.44-3.38 (m, 4H, piperidine), 3.26-3.21 (m, 2H, piperidine), 3.16-3.14 (m, 4H, piperidine, ethylene-linker CH<sub>2</sub>), 3.08-2.98 (m, 2H, ethylene-linker CH<sub>2</sub>), 2.93 (m, 1H, pyrrolidine H-5'), 2.21-2.14 (m, 2H, pyrrolidine H-4', H-5'), 2.09-2.03 (m, 2H, pyrrolidine H-4'). HR-MS (ESI+): *m/z* [M + H]<sup>+</sup> calcd. for C<sub>24</sub>H<sub>29</sub>N<sub>3</sub>O<sub>2</sub>Cl: 426.1948; found: 426.1948.

1'-[2-[4-(4-chlorobenzoyl)piperazin-1-yl]ethyl]-1,2-dihydrospiro[indole-3,3'-pyrrolidine]-2-one (**87**)

To the solution of 4-chlorobenzoic acid (**98**) (63 mg, 0.4 mmol) in 10 mL DMF was added 2-(piperazin-1-yl)ethanol (**97**) (63 mg, 0.48 mmol), HATU (167 mg, 0.44 mmol), and DIPEA (0.084 mL, 0.48 mmol). The suspension was stirred overnight at room temperature, and it was quenched by 15 mL brine, and extracted three times with ethyl-acetate (3 × 5 mL). The combined organic phases were dried over sodium-sulfate, and evaporated under reduced pressure to give the crude (4-chlorophenyl)(4-(2-hydroxyethyl)piperazin-1-yl)methanone (**99**). The crude **99** was taken up in dichloromethane (20 mL), and triethylamine (0.062 mL, 0.4511 mmol), and mesyl chloride (0.029 mL, 0.3759 mmol) were added to it at room temperature. The reaction mixture was stirred at reflux for 2h and then quenched by brine. The mixture was extracted with dichloromethane (3 × 5 mL), and the collected organic phases were dried over sodium-sulfate, and evaporated under reduced pressure to give the crude mesylate (**100**). Next the crude mesylate (**100**) was dissolved in 15 mL acetonitrile, and tetrahydro-β-carboline **53** (35 mg, 0.2015 mmol) and potassium carbonate (83 mg, 0.6045 mmol) were added. The suspension was refluxed overnight. The reaction mixture was quenched with brine (15 mL), and was extracted three times with ethyl-acetate (5 mL × 3). The collected organic phases were dried over sodium-sulphate, and evaporated to give the crude alkylated intermediate (4-chlorophenyl)(4-(2-(3,4-dihydro-1H-pyrido[3,4-*b*]indol-2(9*H*)-yl)ethyl)piperazin-1-yl)methanone (**101**). Intermediate **101** was dissolved in [tetrahydrofuran]:[water]:[glacial acetic acid] = 3:3:1 (3-3-1 mL), and *N*-bromosuccinimide (35 mg, 0.196 mmol) was added in small portions at 0°C. The reaction mixture was stirred for 1.5 hour at 0°C. Afterwards the reaction mixture was quenched at 0 °C with 15 mL concentrated sodium-carbonate solution dropwise (until gas-evolution ceased) and was stirred for 0.5 h, allowing to warm up to room temperature. The mixture was extracted three times using ethyl acetate (3 × 5 mL). The combined organic phases were washed with distilled water (10 mL). The collected organic phases were dried over sodium-

sulfate, filtered and evaporated under reduced pressure. The crude spiro-cyclized product was purified by preparative-HPLC (eluent A: 0.1% formic acid in water; eluent B: 0.1% formic acid in acetonitrile; gradient: 0% to 100% B) Yield: 3.7 mg.  $^1\text{H-NMR}$  (500 MHz, methanol- $d_4$ )  $\delta$  ppm 8.48 (s, 1H, NH), 7.49-7.48 (m, 2H, ArH-2'',6''), 7.45-7.39 (m, 3H, ArH-4, ArH-3'',5''), 7.28 (t,  $J = 7.2$  Hz, 1H, ArH-6), 7.10 (t,  $J = 7.8$  Hz, 1H, ArH-5), 6.95 (d,  $J = 7.6$  Hz, 1H, ArH-7), 3.88-3.85 (m, 1H, pyrrolidine H-2'), 3.77-3.73 (m, 2H, pyrrolidine H-2', H-5'), 3.69-3.67 (m, 2H, piperazine), 3.64-3.60 (m, 2H, piperazine), 3.56-3.53 (m, 2H, piperazine), 3.44-3.39 (m, 3H, ethylene-linker  $\text{CH}_2$ , pyrrolidine H-5'), 2.85 (t,  $J = 6.0$  Hz, 2H, ethylene-linker  $\text{CH}_2$ ), 2.48-2.44 (m, 1H, pyrrolidine H-4'), 2.43-2.38 (m, 1H, pyrrolidine H-4'). HRMS (ESI+):  $m/z$   $[\text{M} + \text{H}]^+$  calcd. for  $\text{C}_{24}\text{H}_{28}\text{N}_4\text{O}_2\text{Cl}$ : 439.1901; found: 439.1895.

## Chapter 4 – Conclusions

### 4.1. Conclusions and thesis points

In my PhD work I aimed the development of ligand-, and structure-based virtual screening methodologies for the design of aminergic GPCR ligands. The prospective screening of our in-house fragment library by our physicochemical property-based and consensus scoring based methods led to the identification of new chemotypes such as spiro[indoline-3,3'-pyrrolidines], and spiro[pyrrolidine-3,3'-oxindoles] as 5-HT<sub>6</sub>R and 5-HT<sub>7</sub>R ligands, respectively. The usefulness of the consensus scoring methodology was also demonstrated by the design of 5-HT<sub>2B</sub>R antagonists.

1. A physicochemical property-based desirability scoring method (FrAGS – Fragment Aminergic GPCR Score) was developed to identify aminergic Class A GPCR ligands. This method might serve as a pre-filtering step with low computational cost for the design of fragment-libraries and for the compilation of ligand sets for structure-based virtual screenings. The performance of FrAGS was validated in retrospective and prospective applications. [K1, ref <sup>143</sup>]

2. A structure-based virtual screening method (FrACS – Fragment Aminergic Consensus Score) was developed for the design of aminergic GPCR focused fragment libraries, using an ensemble of available aminergic X-ray structures to date. The consecutive application of the FrAGS and FrACS methods was validated in a prospective screening of our in-house fragment library. As a result, four low micromolar to nanomolar 5-HT<sub>6</sub>R hits were identified including spiro[pyrrolidine-3,3'-oxindoles] as a novel starting point for 5-HT<sub>6</sub>R ligands. [K2, ref <sup>154</sup>]

3. The pharmacophore supported optimization of the spiro[pyrrolidine-3,3'-oxindoles] led to the identification of aryl-sulfonylated spiro[indoline-3,3'-pyrrolidines] with improved 5-HT<sub>6</sub>R affinities. A structure-activity relationship analysis of 11 SAR-by-catalogue and 12 synthesized compounds was performed around the scaffold of spiro[indoline-3,3'-pyrrolidines] resulting in new 5-HT<sub>6</sub>R antagonists with submicromolar affinities and 100-to-200-fold selectivities against 5-HT<sub>7</sub>R. [K3, K4 ref <sup>172, 64</sup>]

4. Aryl-piperazine type analogues of spiro[pyrrolidine-3,3'-oxindoles] were identified as potent 5-HT<sub>7</sub>R ligands. A pharmacophore supported optimization led to the synthesis and identification of 12 novel 5-HT<sub>7</sub>R ligands with double-digit nanomolar potency and high selectivity (100-to-200-fold) against 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R, 5-HT<sub>7</sub>R.

5. Structure-based consensus scoring was applied in combination with a ligand-based machine learning pre-filter for the design of 5-HT<sub>2B</sub>R selective antagonists that provided two highly potent and selective 5-HT<sub>2B</sub>R ligands. [K5, ref <sup>198</sup>]

## 4.2. Összefoglalás és tézispontok

Munkám során aminerg-GPCR fókuszált fragmenskönyvtárak tervezésére alkalmas ligandum-, és szerkezet-alapú virtuális szűrési módszerek kidolgozását és alkalmazását tűztem ki célul. A sikeres módszerfejlesztést követően végzett prospektív szűrővizsgálatok segítségével új szerotonin receptor (5-HT<sub>6</sub>R, 5-HT<sub>7</sub>R) kemotípusokat (spiro[pyrrolidine-3,3'-oxindolok], spiro[indoline-3,3'-pyrrolidinek]) azonosítottunk, valamint új, szelektív 5-HT<sub>2B</sub>R ligandumokat terveztünk.

1. Kidolgoztam egy fizikokémiai tulajdonságokra épülő kívánatossági-függvény alapú módszert (FrAGS – Fragmens Aminerg GPCR pontozás) aminerg GPCR ligandumok azonosítására. Ez az alacsony számításigényű eljárás előszűrési módszerként szolgálhat fragmenskönyvtárak tervezéséhez és szerkezet-alapú virtuális szűrések ligandum készletének összeállításához. A módszer teljesítőképességét retrospektív és prospektív alkalmazásokkal támasztottuk alá. [K1, ref <sup>143</sup>]

2. Kidolgoztam egy szerkezet-alapú konszenzus pontozási módszert (FrACS – Fragmens Aminerg Konszenzus Pontozás) a rendelkezésre álló aminerg GPCR röntgenszerkezetek felhasználásával, mely alkalmas aminerg-GPCR fókuszált fragmenskönyvtárak tervezésére. A FrAGS és FrACS módszerek lépésenkénti kombinált alkalmazását prospektív szűréssel validáltam. Ennek eredményeképpen összesen négy 5-HT<sub>6</sub>R receptoron hatékony fragmens találatot azonosítottam. A spiro[pyrrolidine-3,3'-oxindol] szerkezetet tartalmazó vegyületek ígéretes kiindulópontok lehetnek új 5-HT<sub>6</sub>R ligandumok fejlesztéséhez. [K2, ref <sup>154</sup>]

3. Az előző tézispontban azonosított spiro-oxindolokat 5-HT<sub>6</sub>R farmakofórmodell alapján optimaltunk, így jutottunk a megfelelő aril-szulfonil spiro[indoline-3,3'-pyrrolidin] származékokhoz. Szerkezet-hatás összefüggés elemzést (SAR) végeztem a spiro[indoline-3,3'-pyrrolidinek] alapváza körül, melynek során új szubmikromólos affinitású, 100-200-szoros szelektivitású 5-HT<sub>6</sub>R antagonistákat azonosítottam. Összesen 11 vegyületet az úgynevezett katalógus alapú SAR segítségével vizsgáltam, további 12 származékot egyedileg szintetizáltam, és szerotonin receptor panelen (5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R, 5-HT<sub>7</sub>R) jellemeztünk. [K3, K4, ref <sup>172, 64</sup>]

4. A spiro[pyrrolidine-3,3'-oxindolok] aril-piperazin származékait potenciális 5-HT<sub>7</sub>R ligandumokként azonosítottam. Farmakofór-modell alapján szerkezet-hatás összefüggés elemzést (SAR) végeztem, melynek során 12 származékot szintetizáltam és jellemeztem 5-HT<sub>1A</sub>R, 5-HT<sub>2A</sub>R, 5-HT<sub>6</sub>R, 5-HT<sub>7</sub>R tesztekben. Ennek eredményeképpen nagy szelektivitású (100-200-szoros) és affinitású (nanomoláris) vegyületeket azonosítottam.

5. A második tézispontban kidolgozott szerkezet-alapú konszenzus pontozási módszert egy ligandum-alapú számítógépes tanulási módszerrel kombinálva 5-HT<sub>2B</sub>R szelektív antagonisták tervezésében alkalmaztam, melynek során két nagy affinitású és szelektív 5-HT<sub>2B</sub>R ligandumot azonosítottam. [K5, ref <sup>198</sup>]

*Publications related to the thesis*

**K1.** Kelemen Á. A.; Ferenczy Gy. G.; Keserű Gy. M. A desirability function-based scoring scheme for selecting fragment-like class A aminergic GPCR ligands. *Journal of Computer-Aided Molecular Design* **2015**, 29(1), 59-66. [IF: 3.199 (2015)]

**K2.** Kelemen Á. A.; Kiss R.; Ferenczy Gy. G.; Kovács L.; Flachner B.; Lőrincz Zs.; Keserű Gy. M. Structure-Based Consensus Scoring Scheme for Selecting Class A Aminergic GPCR Fragments. *Journal of Chemical Information and Modeling* **2016**, 56(2), 412–422. [IF: 3.76 (2016)]

**K3.** Kelemen Á. A.; Satała G.; Bojarski A. J.; Keserű Gy. M. Spiro[pyrrolidine-3,3'-oxindoles] and Their Indoline Analogues as New 5-HT<sub>6</sub> Receptor Chemotypes. *Molecules* **2017**, 22(12), 2221. [IF: 2.861 (2016)]

**K4.** Kelemen Á. A.; Mordalski S.; Bojarski A. J.; Keserű Gy. M. Computational Modeling of Drugs for Alzheimer's Disease: Design of Serotonin 5-HT<sub>6</sub> Antagonists. *Neuromethods (Springer Protocols)* **2017**, 132, 419-461. ISBN: 978-1-4939-7404-7

**K5.** Rataj K.; Kelemen Á. A.; Brea J.; Loza M. I.; Bojarski A. J.; Keserű Gy. M. Fingerprint-Based Machine Learning Approach to Identify Potent and Selective 5-HT<sub>2B</sub>R Ligands. *Molecules* **2018**, 23, 1137. [IF: 2.861 (2016)]

## References

1. Rees, D. C., Congreve, M., Murray, C. W. & Carr, R. Fragment-based lead discovery. *Nat Rev Drug Discov* **3**, 660–672 (2004).
2. Congreve, M., Carr, R., Murray, C. & Jhoti, H. A ‘Rule of Three’ for fragment-based lead discovery? *Drug Discov Today* **8**, 876–877 (2003).
3. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **9**, 430–431 (2004).
4. Erlanson, D. A., McDowell, R. S. & O’Brien, T. Fragment-based drug discovery. *Journal of Medicinal Chemistry* **47**, 3463–3482 (2004).
5. Zeng, S., Baillargeat, D., Ho, H.-P. & Yong, K.-T. Nanomaterials enhanced surface plasmon resonance for biological and chemical sensing applications. *Chem Soc Rev* **43**, 3426 (2014).
6. Navratilova, I. & Hopkins, A. L. Fragment Screening by Surface Plasmon Resonance. *ACS Med Chem Lett* **1**, 44–48 (2010).
7. Neumann, T., Junker, H.-D., Schmidt, K. & Sekul, R. SPR-based fragment screening: advantages and applications. *Curr Top Med Chem* **7**, 1630–42 (2007).
8. Ladbury, J. E., Klebe, G. & Freire, E. Adding calorimetric data to decision making in lead discovery: a hot tip. *Nat Rev Drug Discov* **9**, 23–27 (2010).
9. Ciulli, A. Biophysical screening for the discovery of small-molecule ligands. *Methods Mol Biol* **1008**, 357–88 (2013).
10. Vedadi, M. *et al.* Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* **103**, 15835–40 (2006).
11. Davies, T. G. & Tickle, I. J. in *Topics in Current Chemistry* 33–59 (Springer, Berlin, Heidelberg, 2011).
12. Blundell, T. L., Jhoti, H. & Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat Rev Drug Discov* **1**, 45–54 (2002).
13. Chilingaryan, Z., Yin, Z. & Oakley, A. J. Fragment-Based Screening by Protein Crystallography: Successes and Pitfalls. *Int J Mol Sci* **13**, 12857–12879 (2012).
14. Lepre, C. A. Practical Aspects of NMR-Based Fragment Screening. *Methods Enzymol* **493**, 219–239 (2011).
15. Li, Y. & Kang, C. Solution NMR Spectroscopy in Target-Based Drug Discovery. *Molecules* **22**, 1399 (2017).
16. Arroyo, X., Goldflam, M., Feliz, M., Belda, I. & Giralt, E. Computer-Aided Design of Fragment Mixtures for NMR-Based Screening. *PLoS One* **8**, e58571 (2013).
17. Pedro, L. & Quinn, R. Native Mass Spectrometry in Fragment-Based Drug Discovery. *Molecules* **21**, 984 (2016).
18. Vivat Hannah, V., Atmanene, C., Zeyer, D., Van Dorsselaer, A. &

- Sanglier-Cianférani, S. Native MS: an 'ESI, way to support structure- and fragment-based drug discovery. *Future Med Chem* **2**, 35–50 (2010).
19. Zanella, F., Lorens, J. B. & Link, W. High content screening: Seeing is believing. *Trends Biotechnol* **28**, 237–245 (2010).
  20. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531–1534 (1996).
  21. Erlanson, D. A. Practical Fragments - Fragments in the clinic: 2016 edition. Available at: <http://practicalfragments.blogspot.hu/2016/07/fragments-in-clinic-2016-edition.html>.
  22. Singh, M., Tam, B. & Akabayov, B. NMR-Fragment Based Virtual Screening: A Brief Overview. *Molecules* **23**, 233 (2018).
  23. Lamoree, B. & Hubbard, R. E. Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem* **61**, 453–464 (2017).
  24. Keseru, G. M. *et al.* Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *Journal of Medicinal Chemistry* **59**, 8189–8206 (2016).
  25. Lau, W. F. *et al.* Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J Comput Aided Mol Des* **25**, 621–636 (2011).
  26. Souers, A. J. *et al.* ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* **19**, 202–208 (2013).
  27. Maly, D. J., Choong, I. C. & Ellman, J. A. Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc Natl Acad Sci U S A* **97**, 2419–24 (2000).
  28. Rao, J., Lahiri, J., Weis, R. M. & Whitesides, G. M. Design, synthesis, and characterization of a high-affinity trivalent system derived from vancomycin and L-Lys-D-Ala-D-Ala. *J Am Chem Soc* **122**, 2698–2710 (2000).
  29. Rao, J. & Whitesides, G. M. Tight binding of a dimeric derivative of vancomycin with dimeric L-Lys-D-Ala-D-Ala. *J Am Chem Soc* **119**, 10286–10290 (1997).
  30. Pang, Y. P., Quiram, P., Jelacic, T., Hong, F. & Brimijoin, S. Highly potent, selective, and low cost bis-tetrahydroaminacrine inhibitors of acetylcholinesterase. Steps toward novel drugs for treating Alzheimer's disease. *J Biol Chem* **271**, 23646–9 (1996).
  31. Burgess, L. E. *et al.* Potent selective nonpeptidic inhibitors of human lung tryptase. *Proc Natl Acad Sci U S A* **96**, 8348–52 (1999).
  32. Hajduk, P. J. *et al.* Discovery of potent nonpeptide inhibitors of

- stromelysin using SAR by NMR. *J Am Chem Soc* **119**, 5818–5827 (1997).
33. Erlanson, D. A. *et al.* In situ assembly of enzyme inhibitors using extended tethering. *Nat Biotechnol* **21**, 308–314 (2003).
  34. Szczepankiewicz, B. G. *et al.* Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. *J Am Chem Soc* **125**, 4087–4096 (2003).
  35. Liu, G. *et al.* Fragment screening and assembly: A highly efficient approach to a selective and cell active protein tyrosine phosphatase 1B inhibitor. *J Med Chem* **46**, 4232–4235 (2003).
  36. Braisted, A. C. *et al.* Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J Am Chem Soc* **125**, 3714–3715 (2003).
  37. Jencks, W. P. On the attribution and additivity of binding energies. *Proc Natl Acad Sci U S A* **78**, 4046–50 (1981).
  38. Murray, C. W. & Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J Comput Aided Mol Des* **16**, 741–53 (2002).
  39. Swayze, E. E. *et al.* SAR by MS: A ligand based technique for drug lead discovery against structured RNA targets. *J Med Chem* **45**, 3816–3819 (2002).
  40. Hubbard, R. E. Fragment approaches in structure-based drug discovery. *J Synchrotron Radiat* **15**, 227–230 (2008).
  41. Roughley, S., Wright, L., Brough, P., Massey, A. & Hubbard, R. E. Hsp90 inhibitors and drugs from fragment and virtual screening. *Top Curr Chem* **317**, 61–82 (2012).
  42. Bertrand, S. M. *et al.* The Discovery of in Vivo Active Mitochondrial Branched-Chain Aminotransferase (BCATm) Inhibitors by Hybridizing Fragment and HTS Hits. *J Med Chem* **58**, 7140–7163 (2015).
  43. Boehm, H. J. *et al.* Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J Med Chem* **43**, 2664–2674 (2000).
  44. Fejzo, J. *et al.* The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem Biol* **6**, 755–69 (1999).
  45. Erlanson, D. A. *et al.* Site-directed ligand discovery. *Proc Natl Acad Sci* **97**, 9367–9372 (2000).
  46. Wendt, M. D. *et al.* Identification of Novel Binding Interactions in the Development of Potent, Selective 2-Naphthamidine Inhibitors of Urokinase. Synthesis, Structural Analysis, and SAR of N-Phenyl Amide 6-Substitution. *J Med Chem* **47**, 303–324 (2004).
  47. Van Dongen, M. J. P. *et al.* Structure-based screening as applied to human FABP4: A highly efficient alternative to HTS for hit generation. *J Am Chem Soc* **124**, 11874–11880 (2002).

48. Bollag, G. *et al.* Vemurafenib: the first drug approved for BRAF-mutant cancer. *Nat Rev Drug Discov* **11**, 873–886 (2012).
49. Isberg, V. *et al.* GPCRdb: an information system for G protein-coupled receptors. *Nucleic Acids Res* **44**, D356–364 (2015).
50. Chattopadhyay, A. GPCRs: Lipid-Dependent Membrane Receptors That Act as Drug Targets. *Adv Biol* **2014**, 1–12 (2014).
51. Rovati, G. E. *et al.* The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. *Mol Pharmacol* **71**, 959–64 (2007).
52. Nygaard, R. *et al.* The dynamic process of  $\beta(2)$ -adrenergic receptor activation. *Cell* **152**, 532–42 (2013).
53. Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc Natl Acad Sci* **108**, 13118–13123 (2011).
54. Rasmussen, S. G. F. *et al.* Crystal structure of the human  $\beta_2$  adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).
55. Michino, M. *et al.* What can crystal structures of aminergic receptors tell us about designing subtype-selective ligands? *Pharmacol Rev* **67**, 198–213 (2015).
56. Deupi, X., Standfuss, J. & Schertler, G. Conserved activation pathways in G-protein-coupled receptors. *Biochem Soc Trans* **40**, 383–388 (2012).
57. Rasmussen, S. G. F. *et al.* Structure of a nanobody-stabilized active state of the  $\beta_2$  adrenoceptor. *Nature* **469**, 175–180 (2011).
58. Ballesteros, J. A. & Weinstein, H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure–function relations in G protein-coupled receptors. *Methods Neurosci* **25**, 366–428 (1995).
59. Kang, D. S. & Tian, X. Role of  $\beta$ -arrestins and arrestin domain-containing proteins in G protein-coupled receptor trafficking. *Curr Opin Cell Biol* **27**, 63–71 (2014).
60. Hall, R. A., Premont, R. T. & Lefkowitz, R. J. Heptahelical receptor signaling: beyond the G protein paradigm. *J Cell Biol* **145**, 927–32 (1999).
61. Goodman, O. B. *et al.*  $\beta$ -Arrestin acts as a clathrin adaptor in endocytosis of the  $\beta_2$ -adrenergic receptor. *Nature* **383**, 447–450 (1996).
62. Wacker, D., Stevens, R. C. & Roth, B. L. How Ligands Illuminate GPCR Molecular Pharmacology. *Cell* **170**, 414–427 (2017).
63. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–45 (2000).
64. Kelemen, Á. A., Mordalski, S., Bojarski, A. J. & Keserű, G. M. in *Neuromethods* **132**, 419–461 (2018).
65. Upton, N., Chuang, T. T., Hunter, A. J. & Virley, D. J. 5-HT<sub>6</sub> receptor antagonists as novel cognitive enhancing agents for Alzheimer’s disease. *Neurotherapeutics* **5**, 458–469 (2008).

66. Messina, D. *et al.* Association of the 5-HT<sub>6</sub> receptor gene polymorphism C267T with Parkinson's disease. *Neurology* **58**, 828–9 (2002).
67. Codony, X., Vela, J. M. & Ramírez, M. J. 5-HT<sub>6</sub> receptor and cognition. *Curr Opin Pharmacol* **11**, 94–100 (2011).
68. Kendall, I. *et al.* E-6801, a 5-HT<sub>6</sub> receptor agonist, improves recognition memory by combined modulation of cholinergic and glutamatergic neurotransmission in the rat. *Psychopharmacology (Berl)* **213**, 413–430 (2011).
69. Dudek, M., Marcinkowska, M., Bucki, A., Olczyk, A. & Kołaczkowski, M. Idalopirdine - a small molecule antagonist of 5-HT<sub>6</sub> with therapeutic potential against obesity. *Metab Brain Dis* **30**, 1487–94 (2015).
70. Frantz, K. ., Hansson, K. ., Stouffer, D. . & Parsons, L. . 5-HT<sub>6</sub> receptor antagonism potentiates the behavioral and neurochemical effects of amphetamine but not cocaine. *Neuropharmacology* **42**, 170–180 (2002).
71. Morairty, S. R., Hedley, L., Flores, J., Martin, R. & Kilduff, T. S. Selective 5HT<sub>2A</sub> and 5HT<sub>6</sub> receptor antagonists promote sleep in rats. *Sleep* **31**, 34–44 (2008).
72. Ham, B.-J. *et al.* Serotonergic genes and personality traits in the Korean population. *Neurosci Lett* **354**, 2–5 (2004).
73. Lane, H.-Y. *et al.* Prefrontal executive function and D1, D3, 5-HT<sub>2A</sub> and 5-HT<sub>6</sub> receptor gene variations in healthy adults. *J Psychiatry Neurosci* **33**, 47–53 (2008).
74. LEE, S.-H. *et al.* Association between the 5-HT<sub>6</sub> receptor C267T polymorphism and response to antidepressant treatment in major depressive disorder. *Psychiatry Clin Neurosci* **59**, 140–145 (2005).
75. Sleight, A. J., Boess, F. G., Bös, M. & Bourson, A. The putative 5-HT<sub>6</sub> receptor: localization and function. *Ann N Y Acad Sci* **861**, 91–6 (1998).
76. Frederick, J. A. & Meador-Woodruff, J. H. Effects of clozapine and haloperidol on 5-HT<sub>6</sub> receptor mRNA levels in rat brain. *Schizophr Res* **38**, 7–12 (1999).
77. Holenz, J. *et al.* Medicinal Chemistry Driven Approaches Toward Novel and Selective Serotonin 5-HT<sub>6</sub> Receptor Ligands. *J Med Chem* **48**, 1781–1795 (2005).
78. Kobilka, B. K. G protein coupled receptor structure and activation. *Biochim Biophys Acta - Biomembr* **1768**, 794–807 (2007).
79. Ivachtchenko, A. V. Sulfonyl-containing modulators of serotonin 5-HT<sub>6</sub> receptors and their pharmacophore models. *Russ Chem Rev* **83**, 439–473 (2014).
80. Staroń, J., Warszycki, D., Kalinowska-Thuścik, J., Satała, G. & Bojarski, A. J. Rational design of 5-HT<sub>6</sub> R ligands using a bioisosteric strategy: synthesis, biological evaluation and molecular modelling. *RSC Adv* **5**, 25806–25815 (2015).

81. Cummings, J., Morstorf, T. & Lee, G. Alzheimer's drug-development pipeline: 2016. *Alzheimer's Dement Transl Res Clin Interv* **2**, 222–232 (2016).
82. U.S. National Library of Medicine. Home - ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/>. (Accessed: 19th December 2016)
83. Fillit, H; Friedman, L; Hara, Y; Koemeter-Cox, A.; McKeegan, N. Closing in on a cure: 2017 Alzheimer's clinical trials report. *Alzheimer's Drug Discovery Foundation* (2017). Available at: <https://www.alzdiscovery.org/assets/content/static/ADDF-2017-Alzheimers-Clinical-Trials-Report.pdf>.
84. Plassat, J. L., Amlaiky, N. & Hen, R. Molecular cloning of a mammalian serotonin receptor that activates adenylate cyclase. **44**, 229–236 (1993).
85. Mullins, U., Gianutsos, G. & Eison, A. S. Effects of Antidepressants on 5-HT<sub>7</sub> Receptor Regulation in the Rat Hypothalamus. *Neuropsychopharmacology* **21**, 352–367 (1999).
86. Sleight, A. J., Carolo, C., Petit, N., Zwingelstein, C. & Bourson, A. Identification of 5-hydroxytryptamine<sub>7</sub> receptor binding sites in rat hypothalamus: sensitivity to chronic antidepressant treatment. *Mol Pharmacol* **47**, 99–103 (1995).
87. Lopez-Rodriguez, M. *et al.* Serotonin 5-HT<sub>7</sub> Receptor Antagonists. *Curr Med Chem Nerv Syst Agents* **4**, 203–214 (2004).
88. Wesołowska, A., Nikiforuk, A., Stachowicz, K. & Tatarczyńska, E. Effect of the selective 5-HT<sub>7</sub> receptor antagonist SB 269970 in animal models of anxiety and depression. *Neuropharmacology* **51**, 578–586 (2006).
89. Galici, R., Boggs, J. D., Miller, K. L., Bonaventure, P. & Atack, J. R. Effects of SB-269970, a 5-HT<sub>7</sub> receptor antagonist, in mouse models predictive of antipsychotic-like activity. *Behav Pharmacol* **19**, 153–159 (2008).
90. Meneses, A. & Terrón, J. A. Role of 5-HT<sub>1A</sub> and 5-HT<sub>7</sub> receptors in the facilitatory response induced by 8-OH-DPAT on learning consolidation. *Behav Brain Res* **121**, 21–8 (2001).
91. Nikiforuk, A. Targeting the serotonin 5-HT<sub>7</sub>receptor in the search for treatments for CNS disorders: Rationale and progress to date. *CNS Drugs* **29**, 265–275 (2015).
92. Yang, Z., Liu, X., Yin, Y., Sun, S. & Deng, X. Involvement of 5-HT<sub>7</sub> receptors in the pathogenesis of temporal lobe epilepsy. *Eur J Pharmacol* **685**, 52–58 (2012).
93. Brenchat, A. *et al.* 5-HT<sub>7</sub>receptor activation inhibits mechanical hypersensitivity secondary to capsaicin sensitization in mice. *Pain* **141**, 239–247 (2009).
94. Wang, X. *et al.* 5-HT<sub>7</sub> Receptors Are Involved in Neurogenic Dural Vasodilatation in an Experimental Model of Migraine. *J Mol Neurosci* **54**, 164–170 (2014).
95. Ciranna, L. & Catania, M. V. 5-HT<sub>7</sub> receptors as modulators of neuronal

- excitability, synaptic transmission and plasticity: physiological role and possible implications in autism spectrum disorders. *Front Cell Neurosci* **8**, 250 (2014).
96. De Filippis, B. *et al.* Pharmacological Stimulation of the Brain Serotonin Receptor 7 as a Novel Therapeutic Approach for Rett Syndrome. *Neuropsychopharmacology* **39**, 2506–2518 (2014).
  97. Di Pilato, P. *et al.* Selective agonists for serotonin 7 (5-HT<sub>7</sub>) receptor and their applications in preclinical models: An overview. *Reviews in the Neurosciences* **25**, 401–415 (2014).
  98. López-Rodríguez, M. L. *et al.* First pharmacophoric hypothesis for 5-HT<sub>7</sub> antagonism. *Bioorg Med Chem Lett* **10**, 1097–1100 (2000).
  99. Catalyst 3.1; Molecular Simulations Inc.: San Diego, CA,. (1996).
  100. López-Rodríguez, M. L. *et al.* Optimization of the Pharmacophore Model for 5-HT<sub>7</sub>R Antagonism. Design and Synthesis of New Naphtholactam and Naphthosultam Derivatives. *J Med Chem* **46**, 5638–5650 (2003).
  101. Silberstein, S. D. & McCrory, D. C. Ergotamine and Dihydroergotamine: History, Pharmacology, and Efficacy. *Headache J Head Face Pain* **43**, 144–166 (2003).
  102. Launay, J.-M. *et al.* Function of the serotonin 5-hydroxytryptamine 2B receptor in pulmonary hypertension. *Nat Med* **8**, 1129–1135 (2002).
  103. Jaffré, F. *et al.* Involvement of the serotonin 5-HT<sub>2B</sub> receptor in cardiac hypertrophy linked to sympathetic stimulation: control of interleukin-6, interleukin-1beta, and tumor necrosis factor-alpha cytokine production by ventricular fibroblasts. *Circulation* **110**, 969–74 (2004).
  104. Alberti, C. Drug-induced retroperitoneal fibrosis: short aetiopathogenetic note, from the past times of ergot-derivatives large use to currently applied bio-pharmacology. *G Chir* **36**, 187–91 (2015).
  105. Johnson, K. W., Phebus, L. A. & Cohen, M. L. Serotonin in migraine: theories, animal models and emerging therapies. *Prog Drug Res* **51**, 219–244 (1998).
  106. Fozard, J. R. & Kalkman, H. O. 5-Hydroxytryptamine (5-HT) and the initiation of migraine: new perspectives. *Naunyn-Schmiedeberg's Archives of Pharmacology* **350**, 225–229 (1994).
  107. Schmuck, K., Ullmer, C., Kalkman, H. O., Probst, A. & Lübbert, H. Activation of Meningeal 5-HT<sub>2B</sub> Receptors: An Early Step in the Generation of Migraine Headache? *Eur J Neurosci* **8**, 959–967 (1996).
  108. Janssen, W. *et al.* 5-HT<sub>2B</sub> receptor antagonists inhibit fibrosis and protect from RV heart failure. *Biomed Res Int* **2015**, 438403 (2015).
  109. Zhou, Y. *et al.* Structure-Based Discovery of Novel and Selective 5-Hydroxytryptamine 2B Receptor Antagonists for the Treatment of Irritable Bowel Syndrome. *J Med Chem* **59**, 707–720 (2016).
  110. Bonhaus, D. W. *et al.* RS-127445: a selective, high affinity, orally

- bioavailable 5-HT<sub>2B</sub> receptor antagonist. *Br J Pharmacol* **127**, 1075–1082 (1999).
111. Thomas, D. R., Gager, T. L., Holland, V., Brown, A. M. & Wood, M. D. m-Chlorophenylpiperazine (mCPP) is an antagonist at the cloned human 5-HT<sub>2B</sub> receptor. *Neuroreport* **7**, 1457–1460 (1996).
  112. Kovács, A. *et al.* Effects of EGIS-7625, a Selective and Competitive 5-HT<sub>2B</sub> Receptor Antagonist. *Cardiovasc Drugs Ther* **17**, 427–434 (2003).
  113. Audia, J. E. *et al.* Potent, selective tetrahydro- $\beta$ -carboline antagonists of the serotonin 2B (5HT<sub>2B</sub>) contractile receptor in the rat stomach fundus. *J Med Chem* **39**, 2773–2780 (1996).
  114. Forbes, I. T. *et al.* N-(1-Methyl-5-indolyl)-N'-(3-pyridyl)urea hydrochloride: the first selective 5-HT<sub>1C</sub> receptor antagonist. *J Med Chem* **36**, 1104–1107 (1993).
  115. Russel, J. S. Oxindoles and Spirocyclic Variations: Strategies for C3 Functionalization. *Top Heterocycl Chem* **26**, 397–432 (2011).
  116. Bindra, J. S. Oxindole alkaloids. *Alkaloids Chem Physiol* **14**, 83–121 (1973).
  117. ChEMBL. Available at: <https://www.ebi.ac.uk/chembl/>. (Accessed: 12th October 2017)
  118. Laronze, J.-Y., Bascop, S.-I., Sapi, J. & Lévy, J. On the Synthesis of the Oxindole Alkaloid: ( $\pm$ )-Horsfiline. *Heterocycles* **38**, 725 (1994).
  119. Overman, L. E. & Rosen, M. D. Total synthesis of (-)-spirotryprostatin B and three stereoisomers. *Angew Chemie - Int Ed* **39**, 4596–4599 (2000).
  120. Finch, N. & Taylor, W. I. The Conversion of Tetrahydro- $\beta$ -Carboline Alkaloids into Oxindoles. The Structures and Partial Syntheses of Mitrephylline and Rhyn Cophylline. *Journal of the American Chemical Society* **84**, 1318–1320 (1962).
  121. Stahl, R. & Borschberg, H. -J. A Reinvestigation of the Oxidative Rearrangement of Yohimbane-Type Alkaloids. Part A. Formation of pseudoindoxyl (= 1,2-dihydro-3H-indol-3-one) derivatives. *Helv Chim Acta* **77**, 1331–1345 (1994).
  122. Shavel, J. & Zinnes, H. Oxindole Alkaloids. I. Oxidative-Rearrangement of Indole Alkaloids to their Oxindole Analogs. *J Am Chem Soc* **84**, 1320–1321 (1962).
  123. Martin, S. F. & Mortimore, M. New methods for the synthesis of oxindole alkaloids. Total syntheses of isopteropodine and pteropodine. *Tetrahedron Lett* **31**, 4557–4560 (1990).
  124. Takayama, H. *et al.* Synthesis of Na-methoxyindole and Na-methoxyoxindole alkaloids having yohimbine skeleton. *Heterocycles* **33**, 121–125 (1992).
  125. Yu, P. & Cook, J. M. Diastereospecific synthesis of ketooxindoles. Potential intermediates for the synthesis of alstonisine as well as for voachalotine related oxindole alkaloids. *Tetrahedron Lett* **38**, 8799–8802

- (1997).
126. Pellegrini, C., Strässler, C., Weber, M. & Borschberg, H. J. Synthesis of the oxindole alkaloid (-)-horsfiline. *Tetrahedron: Asymmetry* **5**, 1979–1992 (1994).
  127. Edmondson, S. D. & Danishefsky, S. J. The total synthesis of spirotryprostatin A. *Angew Chemie - Int Ed* **37**, 1138–1140 (1998).
  128. Edmondson, S., Danishefsky, S. J., Sepp-Lorenzino, L. & Rosen, N. Total synthesis of spirotryprostatin A, leading to the discovery of some biologically promising analogues. *J Am Chem Soc* **121**, 2147–2155 (1999).
  129. Wang, H. & Ganesan, A. A biomimetic total synthesis of (-)-spirotryprostatin B and related studies. *J Org Chem* **65**, 4685–4693 (2000).
  130. Pellegrini, C., Strässler, C., Weber, M. & Borschberg, H. J. Synthesis of the oxindole alkaloid (-)-horsfiline. *Tetrahedron: Asymmetry* **5**, 1979–1992 (1994).
  131. Sebahar, P. R. & Williams, R. M. The asymmetric total synthesis of (+)- and (-)-spirotryprostatin B [16]. *Journal of the American Chemical Society* **122**, 5666–5667 (2000).
  132. Jones, K. & Wilkinson, J. A total synthesis of horsfiline via aryl radical cyclisation. *J Chem Soc Chem Commun* **5**, 1767 (1992).
  133. Abelman, M. M., Oh, T. & Overman, L. E. Intramolecular Alkene Arylations for Rapid Assembly of Polycyclic Systems Containing Quaternary Centers. A New Synthesis of Spirooxindoles and Other Fused and Bridged Ring Systems. *Journal of Organic Chemistry* **52**, 4130–4133 (1987).
  134. Fuji, K., Node, M., Nagasawa, H., Naniwa, Y. & Terada, S. Asymmetric Induction via Addition-Elimination Process: Nitroolefination of  $\alpha$ -Substituted Lactones. *J Am Chem Soc* **108**, 3855–3856 (1986).
  135. Miyake, F. Y., Yakushijin, K. & Horne, D. A. Preparation and synthetic applications of 2-halotryptophan methyl esters: Synthesis of spirotryprostatin B. *Angew Chemie - Int Ed* **43**, 5357–5360 (2004).
  136. Meyers, C. & Carreira, E. M. Total Synthesis of (-)-Spirotryprostatin B. *Angew Chemie Int Ed* **42**, 694–696 (2003).
  137. Kulkarni, M. G. *et al.* Total synthesis of ( $\pm$ )-coerulescine and ( $\pm$ )-horsfiline. *Beilstein J Org Chem* **6**, 876–879 (2010).
  138. De, S., Das, M. K., Bhunia, S. & Bisai, A. Unified Approach to the Spiro(pyrrolidinyloxindole) and Hexahydropyrrolo[2,3-b]indole Alkaloids: Total Syntheses of Pseudophrynamines 270 and 272A. *Org Lett* **17**, 5922–5925 (2015).
  139. Sendai Fukusokan Kagaku Kenkyūjo., H. *et al.* *Heterocycles*. *Heterocycles* **33**, (Sendai Institute of Heterocyclic Chemistry).
  140. Edmondson, S., Danishefsky, S. J., Sepp-Lorenzino, L. & Rosen, N. Total synthesis of spirotryprostatin A, leading to the discovery of some

- biologically promising analogues. *J Am Chem Soc* **121**, 2147–2155 (1999).
141. Pictet, A. & Spengler, T. Über die Bildung von Isochinolin Derivaten durch Einwirkung von Methylal auf Phenylethylamin, Phenylalanin und Tyrosin. *Berichte der Dtsch Chem Gesellschaft* **44**, 2030–2036 (1911).
  142. Zhao, Y. *et al.* A potent small-molecule inhibitor of the MDM2-p53 interaction (MI-888) achieved complete and durable tumor regression in mice. *J Med Chem* **56**, 5553–5561 (2013).
  143. Kelemen, Á. A., Ferenczy, G. G. & Keserű, G. M. A desirability function-based scoring scheme for selecting fragment-like class A aminergic GPCR ligands. *J Comput Aided Mol Des* **29**, 59–66 (2015).
  144. Andrews, S. P., Brown, G. A. & Christopher, J. A. Structure-Based and Fragment-Based GPCR Drug Discovery. *ChemMedChem* **9**, 256–275 (2014).
  145. Balakin, K. V. *et al.* Property-based design of GPCR-targeted library. *J Chem Inf Comput Sci* **42**, 1332–1342 (2002).
  146. JChem for Excel, ChemAxon for Knime, ChemAxon fingerprint, version 5.2, Chemaxon: Budapest, Hungary. (2014).
  147. D. Segall, M. Multi-Parameter Optimization: Identifying High Quality Compounds with a Balance of Properties. *Curr Pharm Des* **18**, 1292–1310 (2012).
  148. Harrington, E. The desirability function. *Ind Qual Control* **21**, 494–498 (1965).
  149. Wager, T. T., Hou, X., Verhoest, P. R. & Villalobos, A. Moving beyond rules: the development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem Neurosci* **1**, 435–49 (2010).
  150. Staroń, J. *et al.* Pyrano[2,3,4-*cd*]indole as a Scaffold for Selective Nonbasic 5-HT<sub>6</sub> R Ligands. *ACS Med Chem Lett* **8**, 390–394 (2017).
  151. Harris, R. N. *et al.* Highly potent, non-basic 5-HT<sub>6</sub> ligands. Site mutagenesis evidence for a second binding mode at 5-HT<sub>6</sub> for antagonism. *Bioorg Med Chem Lett* **20**, 3436–3440 (2010).
  152. The PubChem Project. Available at: <https://pubchem.ncbi.nlm.nih.gov/>. (Accessed: 24th January 2018)
  153. Zhao, W., Hevener, K. E., White, S. W., Lee, R. E. & Boyett, J. M. A statistical framework to evaluate virtual screening. *BMC Bioinformatics* **10**, 225 (2009).
  154. Kelemen, Á. A. *et al.* Structure-Based Consensus Scoring Scheme for Selecting Class A Aminergic GPCR Fragments. *J Chem Inf Model* **56**, 412–422 (2016).
  155. Topiol, S. & Sabio, M. X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem Pharmacol* **78**, 11–20 (2009).
  156. Bohacek, R. S., McMartin, C. & Guida, W. C. The art and practice of structure-based drug design: A molecular modeling perspective. *Med Res*

- Rev* **16**, 3–50 (1996).
157. Totrov, M. & Abagyan, R. Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol* **18**, 178–184 (2008).
  158. Huang, S.-Y. & Zou, X. Ensemble docking of multiple protein structures: Considering protein structural variations in molecular docking. *Proteins Struct Funct Bioinforma* **66**, 399–421 (2006).
  159. Spyrakis, F. & Cavasotto, C. N. Open challenges in structure-based virtual screening: Receptor modeling, target flexibility consideration and active site water molecules description. *Arch Biochem Biophys* **583**, 105–119 (2015).
  160. Glide, Version 5.7, Schrödinger, LLC, New York, NY. (2011).
  161. Nicholls, A. Confidence limits, error bars and method comparison in molecular modeling. Part 1: The calculation of confidence intervals. *J Comput Aided Mol Des* **28**, 887–918 (2014).
  162. McGann, M., Nicholls, A. & Enyedy, I. The statistics of virtual screening and lead optimization. *J Comput Aided Mol Des* **29**, 923–936 (2015).
  163. Shimamura, T. *et al.* Structure of the human histamine H1 receptor complex with doxepin. *Nature* **475**, 65–70 (2011).
  164. Kruse, A. C. *et al.* Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* **504**, 101–106 (2013).
  165. Chien, E. Y. T. *et al.* Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **330**, 1091–5 (2010).
  166. Ring, A. M. *et al.* Adrenaline-activated structure of  $\beta$ 2-adrenoceptor stabilized by an engineered nanobody. *Nature* **502**, 575–579 (2013).
  167. LigPrep, Version 2.5, Schrödinger, LLC, New York, NY. (2011).
  168. Whitty, A. Growing PAINS in academic drug discovery. *Future Med Chem* **3**, 797–801 (2011).
  169. Baell, J. B. & Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J Med Chem* **53**, 2719–2740 (2010).
  170. Maher-Edwards, G. *et al.* SB-742457 and donepezil in Alzheimer disease: a randomized, placebo-controlled study. *Int J Geriatr Psychiatry* **26**, 536–544 (2011).
  171. Vass, M. *et al.* Dynamics and structural determinants of ligand recognition of the 5-HT<sub>6</sub> receptor. *J Comput Aided Mol Des* **29**, 1137–49 (2015).
  172. Kelemen, Á. A., Satala, G., Bojarski, A. J. & Keseru, G. M. Spiro[pyrrolidine-3, 3'-oxindoles] and their indoline analogues as new 5-HT<sub>6</sub> receptor chemotypes. *Molecules* **22**, 2221 (2017).
  173. mcule. Available at: <https://mcule.com/>. (Accessed: 12th October 2017)
  174. Kiss, R., Sandor, M. & Szalai, F. A. <http://Mcule.com>: a public web service

- for drug discovery. *J Cheminform* **4**, P17 (2012).
175. Powell, N. A. *et al.* Novel and selective spiroindoline-based inhibitors of sky kinase. *Bioorganic Med Chem Lett* **22**, 190–193 (2012).
  176. Li, C., Chan, C., Heimann, A. C. & Danishefsky, S. J. On the rearrangement of an azaspiroindolenine to a precursor to phalarine: Mechanistic insights. *Angew Chemie - Int Ed* **46**, 1444–1447 (2007).
  177. Ottoni, O., Cruz, R. & Krammer, N. H. Regioselective nitration of 3-acetylindole and its N-acyl and N-sulfonyl derivatives. *Tetrahedron Lett* **40**, 1117–1120 (1999).
  178. Lakshmaiah, G., Kawabata, T., Shang, M. & Fuji, K. Total synthesis of (-)-horsfiline via asymmetric nitroolefination. *Journal of Organic Chemistry* **64**, 1699–1704 (1999).
  179. Hostetler, G., Dunn, D., McKenna, B. A., Kopec, K. & Chatterjee, S. 1-Thia-4,7-diaza-spiro[4.4]nonane-3,6-dione: A Structural Motif for 5-hydroxytryptamine 6 receptor antagonism. *Chem Biol Drug Des* **83**, 149–153 (2014).
  180. Protiva, M., Vejdšek, Z. J., Jílek, J. O. & Macek, K. Synthetische modelle hypotensiv wirksamer alkaloides V. Einige weitere derivates des tryptamins und 1,2,3,4-tetrahydronorharmans. *Collect Czechoslov Chem Commun* **24**, 3978–3987 (1959).
  181. Bell, S. E. V., Brown, R. F. C., Eastwood, F. W. & Horvath, J. M. An Approach to Some Spiro Oxindole Alkaloids Through Cycloaddition Reactions of 3-Methylideneindolin-2-one. *Aust J Chem* **53**, 183 (2000).
  182. Zaidlewicz, M., Brown, H. C. & Neelamkavil, S. F. in *Encyclopedia of Reagents for Organic Synthesis* (John Wiley & Sons, Ltd, 2008).
  183. Allen, J. *et al.* The use of biocatalysis in the synthesis of labelled compounds. in *Journal of Labelled Compounds and Radiopharmaceuticals* **50**, 342–346 (John Wiley & Sons, Ltd., 2007).
  184. Manda, S. *et al.* Synthesis, antileishmanial and antitrypanosomal activities of N-substituted tetrahydro- $\beta$ -carboline. *Bioorganic Med Chem Lett* **24**, 3247–3250 (2014).
  185. Sevrin, M.; Morel, C.; Menin, J.; George, P. 2-(Methyl(4-piperidinyl))-1,2,3,4-tetrahydro-9H-pyrido(3,4-b)indole derivatives, their preparation and therapeutical use - Patent EP0302788 - PubChem. (1992).
  186. Shah, P. & Westwell, A. D. The role of fluorine in medicinal chemistry. *J Enzyme Inhib Med Chem* **22**, 527–540 (2007).
  187. Kikuchi, C., Nagaso, H., Hiranuma, T. & Koyama, M. Tetrahydrobenzindoles: Selective antagonists of the 5-HT<sub>7</sub> receptor. *Journal of Medicinal Chemistry* **42**, 533–535 (1999).
  188. Kikuchi, C. *et al.* 2a-[4-(Tetrahydropyridoindol-2-yl)butyl]tetrahydrobenzindole Derivatives: New Selective Antagonists of the 5-Hydroxytryptamine 7 Receptor. *J Med Chem* **45**, 2197–2206 (2002).

189. Medina, R. A. *et al.* Synthesis of New Serotonin 5-HT<sub>7</sub> Receptor Ligands. Determinants of 5-HT<sub>7</sub>/5-HT<sub>1A</sub> Receptor Selectivity. *J Med Chem* **52**, 2384–2392 (2009).
190. Kikuchi, C., Nagaso, H., Hiranuma, T. & Koyama, M. Tetrahydrobenzindoles: Selective antagonists of the 5-HT<sub>7</sub> receptor. *Journal of Medicinal Chemistry* **42**, 533–535 (1999).
191. Perrone, R. *et al.* Synthesis and structure-affinity relationships of 1-[ $\omega$ -(4-aryl-1-piperazinyl)alkyl]-1-aryl ketones as 5-HT<sub>7</sub> receptor ligands. *J Med Chem* **46**, 646–649 (2003).
192. Leopoldo, M., Lacivita, E., Berardi, F., Perrone, R. & Hedlund, P. B. Serotonin 5-HT<sub>7</sub>receptor agents: Structure-activity relationships and potential therapeutic applications in central nervous system disorders. *Pharmacol Ther* **129**, 120–148 (2011).
193. Pittala, V. & Pittala, D. Latest Advances Towards the Discovery of 5-HT<sub>7</sub> Receptor Ligands. *Mini-Reviews Med Chem* **11**, 1108–1121 (2011).
194. Deau, E. *et al.* Rational Design, Pharmacomodulation, and Synthesis of Dual 5-Hydroxytryptamine 7 (5-HT<sub>7</sub>)/5-Hydroxytryptamine 2A (5-HT<sub>2A</sub>) Receptor Antagonists and Evaluation by [18F]-PET Imaging in a Primate Brain. *J Med Chem* **58**, 8066–8096 (2015).
195. Volk, B. *et al.* Optimization of (Arylpiperazinylbutyl)oxindoles Exhibiting Selective 5-HT<sub>7</sub> Receptor Antagonist Activity. *J Med Chem* **54**, 6657–6669 (2011).
196. Na, Y. H. *et al.* Novel quinazolinone derivatives as 5-HT<sub>7</sub> receptor ligands. *Bioorg Med Chem* **16**, 2570–8 (2008).
197. Zhuang, R. *et al.* Exploration of novel piperazine or piperidine constructed non-covalent peptidyl derivatives as proteasome inhibitors. *Eur J Med Chem* **126**, 1056–1070 (2017).
198. Rataj, K. *et al.* Fingerprint-Based Machine Learning Approach to Identify Potent and Selective 5-HT<sub>2BR</sub> Ligands. *Molecules* **23**, 1137 (2018).
199. Vass, M., Ágai-Csongor, É., Horti, F. & Keserű, G. M. Multiple Fragment Docking and Linking in Primary and Secondary Pockets of Dopamine Receptors. *ACS Med Chem Lett* **5**, 1010–1014 (2014).
200. Rataj, K.; Czarnecki, W.; Podlowska, S.; Bojarski, A. J. Substructural Connectivity Fingerprint and Extreme Entropy Machines—A New Method of Compound Representation and Analysis. *Molecules* **23**, 1242 (2018).
201. Witek, J., Smusz, S., Rataj, K., Mordalski, S. & Bojarski, A. J. An application of machine learning methods to structural interaction fingerprints - A case study of kinase inhibitors. *Bioorganic Med Chem Lett* **24**, 580–585 (2014).
202. Bento, A. P. *et al.* The ChEMBL bioactivity database: an update. *Nucleic Acids Res* **42**, D1083–D1090 (2014).

203. Wang, C. *et al.* Structural basis for molecular recognition at serotonin receptors. *Science* **340**, 610–4 (2013).
204. Wacker, D. *et al.* Structural Features for Functional Selectivity at Serotonin Receptors. *Science* **340**, 615–619 (2013).
205. Liu, W. *et al.* Serial Femtosecond Crystallography of G Protein–Coupled Receptors. *Science* **342**, 1521 LP-1524 (2013).
206. Kristiansen, K. *et al.* A highly conserved aspartic acid (Asp-155) anchors the terminal amine moiety of tryptamines and is involved in membrane targeting of the 5-HT(2A) serotonin receptor but does not participate in activation via a ‘salt-bridge disruption’ mechanism. *J Pharmacol Exp Ther* **293**, 735–746 (2000).
207. Moss, N. *et al.* A new class of 5-HT<sub>2B</sub> antagonists possesses favorable potency, selectivity, and rat pharmacokinetic properties. *Bioorg Med Chem Lett* **19**, 2206–2210 (2009).
208. ChEMBL GPCR SARfari Home Page. Available at: <https://www.ebi.ac.uk/chembl/sarfari/gpcrsarfari>.
209. Bento, A. P. *et al.* The ChEMBL bioactivity database: an update. *Nucleic Acids Res* **42**, D1083–D1090 (2014).
210. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res* **40**, D1100-7 (2012).
211. Berthold, M. R. *et al.* {KNIME}: The {K}onstanz {I}nformation {M}iner. in *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)* (Springer, 2007).
212. Nissink, J. W. M. Simple Size-Independent Measure of Ligand Efficiency. *J Chem Inf Model* **49**, 1617–1622 (2009).
213. Hann, M. M., Leach, A. R. & Green, D. V. S. in *Chemoinformatics in Drug Discovery* **23**, 43–57 (Wiley-VCH Verlag GmbH & Co. KGaA, 2005).
214. {RDKit}: Open-source cheminformatics.
215. Haga, K. *et al.* Structure of the human M<sub>2</sub> muscarinic acetylcholine receptor bound to an antagonist. *Nature* **482**, 547–551 (2012).
216. Hanson, M. A. *et al.* A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β<sub>2</sub>-Adrenergic Receptor. *Structure* **16**, 897–905 (2008).
217. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human beta<sub>2</sub>-adrenergic G protein-coupled receptor. *Science* **318**, 1258–65 (2007).
218. Wacker, D. *et al.* Conserved Binding Mode of Human β<sub>2</sub> Adrenergic Receptor Inverse Agonists and Antagonist Revealed by X-ray Crystallography. *J Am Chem Soc* **132**, 11443–11445 (2010).
219. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist-β<sub>2</sub> adrenoceptor complex. *Nature* **469**, 236–240 (2011).
220. Zou, Y., Weis, W. I. & Kobilka, B. K. N-Terminal T4 Lysozyme Fusion

- Facilitates Crystallization of a G Protein Coupled Receptor. *PLoS One* **7**, e46039 (2012).
221. Rasmussen, S. G. F. *et al.* Crystal structure of the human  $\beta$ 2 adrenergic G-protein-coupled receptor. *Nature* **450**, 383–387 (2007).
  222. Bokoch, M. P. *et al.* Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **463**, 108–112 (2010).
  223. Weichert, D. *et al.* Covalent agonists for studying G protein-coupled receptor activation. *Proc Natl Acad Sci U S A* **111**, 10744–8 (2014).
  224. Wang, C. *et al.* Structural basis for molecular recognition at serotonin receptors. *Science* **340**, 610–4 (2013).
  225. Wacker, D. *et al.* Structural features for functional selectivity at serotonin receptors. *Science* **340**, 615–9 (2013).
  226. Liu, W. *et al.* Serial femtosecond crystallography of G protein-coupled receptors. *Science* **342**, 1521–4 (2013).
  227. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res* **28**, 235–42 (2000).
  228. Protein Preparation, Version X, Schrödinger, LLC, New York, NY. (2011).
  229. Clustering of Conformers, Schrödinger, LLC, New York, NY. (2014).
  230. Godden, J. W., Xue, L. & Bajorath, J. Combinatorial Preferences Affect Molecular Similarity/Diversity Calculations Using Binary Fingerprints and Tanimoto Coefficients. *J Chem Inf Comput Sci* **40**, 163–166 (2000).
  231. ChEMBL GPCR SARfari Home Page. (2014).
  232. Berman, H. M. The Protein Data Bank. *Nucleic Acids Res* **28**, 235–242 (2000).
  233. Schrödinger Release 2017-4: Schrödinger Suite 2017-4 Protein Preparation Wizard; Schrödinger, LLC, New York, NY. (2017).
  234. Yung-Chi, C. & Prusoff, W. H. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22**, 3099–3108 (1973).

## APPENDIX

Drug/Compound	Company	Target - Indication(s)	Fragment optimization	Year
---------------	---------	------------------------	-----------------------	------

### Approved drug

Verubecestat	Merck	beta-secretase 1 (BACE1) - Alzheimer's Disease	growing, NMR/X-ray based	2017
Vemurafenib (Zelboraf)	Plexxikon /Roche	B-Raf(V600E) - skin-cancer	growing, X-ray based	2011
Venetoclax (Venclexta)	AbbVie /Genentech	Selective Bcl-2 - chronic lymphocytic leukemia	linking, target-detected NMR	2016

### Phase 3

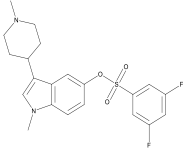
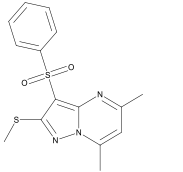
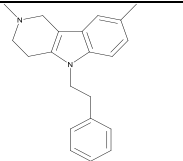
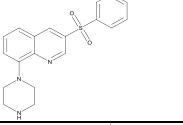
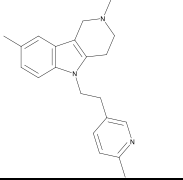
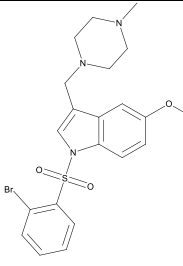
PLX3397	Plexxikon	colony-stimulating factor 1(CSF1R) - tenosynovial giant-cell tumor	growing, X-ray based	2015
Lanabecestat (AZD3293, LY3314814)	AstraZeneca /Astex /Lilly	beta-secretase 1 (BACE1) - Alzheimer's Disease	growing, NMR/X-ray/calorimetry	2014

### Phase 2

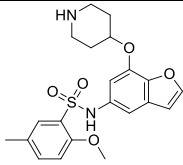
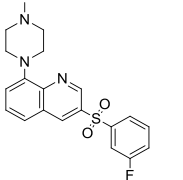
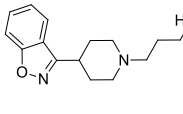
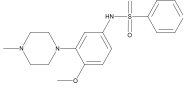
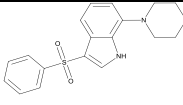
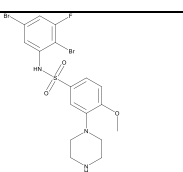
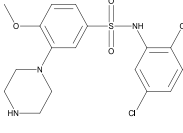
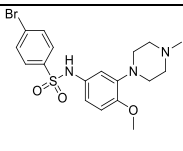
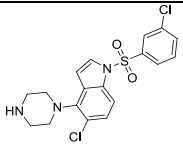
AT7519	Astex	cyclin dependent kinase (CDK1,2,4,5,9) - cancer therapy	growing, X-ray based	2008
AT9283	Astex	Aurora, JAK2 – cancer therapy	growing, X-ray based	2009
AZD5363	AstraZeneca /Astex/CR-UK	Akt-kinase – breast cancer	growing, X-ray based	2013
Erdafitinib	J&J /Astex	FGFR1-4	growing, X-ray based	2017
Indeglitazar	Plexxikon	peroxisome proliferator-activated receptors (pan-PPAR) agonist – diabetes	growing, X-ray based	2009
LY2886721	Lilly	beta-secretase 1 (BACE1) - Alzheimer's Disease	co-crystallization	2011
LY517717	Lilly /Protherics	activated Factor Xa (FXa) - thrombosis	X-ray based	2007
Navitoclax (ABT-263)	Abbott	Bcl-2/Bcl-xL	NMR-based	2011
NVP-AUY922	Vernalis /Novartis	HSP90	ligand-observed NMR	2008
Onalespib (AT13387)	Astex	HSP90	growing, NMR/X-ray based	2010

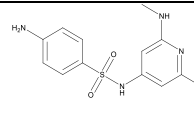
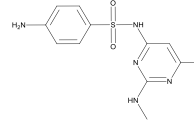
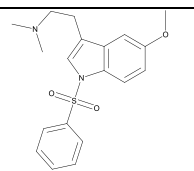
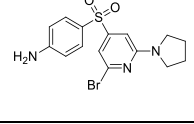
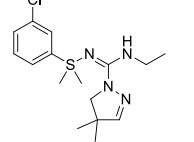
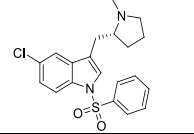
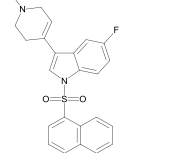
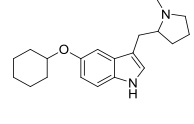
Phase 1				
Mivebrisib (ABBV-075)	AbbVie	BET	merging, X-ray screening	2018
eFT508	eFFECTOR Therapeutics	MNK1/2	merging, X-ray based	2018
ABL001	Novartis	BCR-ABL	X-ray based	2014
ABT-518	Abbott	MMP-2 & 9	SAR by NMR, LC-MS	2009
ABT-737	Abbott	Bcl-2/Bcl-xL	NMR-based screening	2013
ASTX660	Astex	XIAP/cIAP1	growing	2015
AT13148	Astex	AKT, p70S6K, ROCK	ligand-observed NMR, X-ray	2012
AZD3839	AstraZeneca	BACE1	X-ray based	2012
AZD5099	AstraZeneca	Bacterial topoisomerase II	merging (X-Ray screening)	2014
BCL201	Vernalis /Servier /Roche	BCL-2	n.a.	2015
DG-051	deCODE	LTA4H	merging, X-ray screening	2010
IC-776	Lilly /ICOS	LFA-1	NMR-based screening	2012
LP-261	Locus	Tubulin	n.a.	2012
LY2811376	Lilly	BACE1	merging, X-ray screening	2011
PF06650833	Pfizer	IRAK4	n.a.	2017
PLX5568	Plexxikon	kinase	X-ray based	2009
SGX-393	SGX	BCR-ABL	X-ray based	2008
SGX-523	SGX	Met	X-ray based	2009
SNS-314	Sunesis	Aurora	LC-MS	2009

**Table Appendix 1-** List of drugs developed by fragment-based drug discovery approaches<sup>21,22</sup>

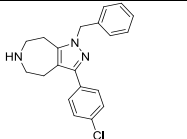
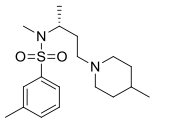
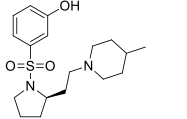
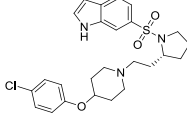
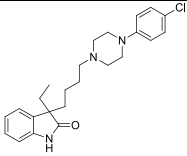
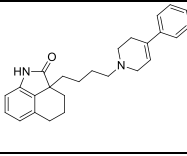
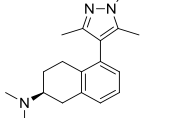
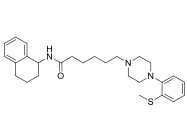
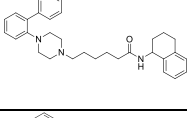
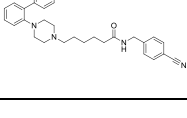
Structure	Name/ Synonyms	Developer	5-HT <sub>6</sub> K <sub>i</sub> (nM)	Clinical Phase	Year
	<b>9</b> , SGS-518, Ly 483518	Lundbeck /Lilly	-	phase 2	2007
	<b>112</b> , AVN-211	Avineuro	2.1	phase 2 active	2011
	<b>113</b> , AVN-101	Avineuro	2.0	phase 2 active	2016
undisclosed	SYN-120, landipirdine	Roche /Biotie Ther.	0.2	phase 2	-
undisclosed	SAM760, PF-05212377, WYE-103760	Pfizer	-	phase 2	-
undisclosed	ABT-354	Abbvie	-	phase 1	0
	<b>8</b> , SB-742457, RVT-101, Intepirdine	GSK /Axovant	0.234	phase 3 failed 2017	2008
	<b>114</b> , Dimebon (latrepirdine)	Pfizer	26.0	phase 3 failed	1983
	<b>10</b> , SUVN-502	Suven LifeSci.	1.71	phase 3 failed	2011

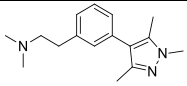
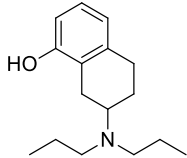
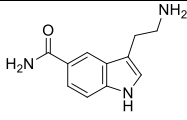
	<b>115</b> , PF-05212365, SAM-531, cerlapirdine	Pfizer	1.3	phase 3 failed	2011
	<b>116</b> , Idalopirdine, Lu_AE_58054,	Lundbeck Pharma	0.83	phase 3 failed	2010
	<b>117</b> , AVN-322	Avineuro	0.39	phase 2 failed	2011
	<b>118</b> , PRX-07034	Epix Pharma	4.0	phase 2 terminated	2008
	<b>119</b> , SYN-114	Synosia Therap. /Biotie Therap.	0.3	phase 1 failed	2014
	<b>120</b> , SB-271046	GSK	1.26	phase 1 failed	1999
	<b>121</b> , R-1485	Roche	1.26	phase 1 failed	2007
	<b>122</b> , SAM-315	Pfizer	1.1	preclinical	
	<b>123</b> , AVN-216	Avineuro	9.59	preclinical	2014

	<b>124</b> , BVT-74316	Biovitrum	1.2	preclinical	2006
	<b>125</b> , 11C-LuAE60157	Lundbeck Pharma	0.2	preclinical	2010
undisclosed	11C-GSK215083	GSK	0.34	preclinical	2012
	<b>126</b> , ADN-1184	Adamed	16.0	preclinical	2014
	<b>127</b> , SB-258585	GSK	2.51	preclinical	-
	<b>128</b> , Ro 4368554	Roche	0.5	preclinical	2005
	<b>129</b> , SB-357134	GSK	3.16	preclinical	2002
	<b>130</b> , SB-399885	GSK	0.7	preclinical	2006
	<b>131</b> , SB-214111	GSK	5.0	preclinical	1999
	<b>132</b> , SB-699929	GSK	0.3	preclinical	-

	<b>133</b> , Ro 63-0563	Roche	12.0	preclinical	1998
	<b>134</b> , Ro 04-6790	Roche	50.12	preclinical	1998
	<b>135</b> , MS-245	Merck	2.3	preclinical	2000
	<b>136</b> , Ro-66-0074	Roche	1.0	preclinical	2003
	<b>137</b> , CMP X	Abbott	25.0	preclinical	2011
	<b>138</b> , ALX-1161	Allelix-NPS Pharma.	1.4	preclinical	1999
	<b>139</b> , ALX-1175	Allelix-NPS Pharma.	1.0	preclinical	2000
	<b>140</b> , ALX-0440	Allelix-NPS Pharma.	0.93	preclinical	1999

**Table Appendix 2** – Summary of prominent 5-HT<sub>6</sub>R antagonists (compounds in shaded boxes have not yet reached the clinic). Green highlight: successful clinical phase active status; Red highlight: failed clinical phase;

Structure	Name/ Synonyms /ID	Developer	5-HT <sub>6</sub> Ki (nM)	Mechanism of action	Clinical Phase	Year
	<b>11</b> , JNJ-18308683	Janssen R&D	-	selective antagonist	phase 2 recruiting, 2017	2012
	<b>141</b> , SB-258719	GSK	31.6	selective inverse agonist	preclinical	1998
	<b>142</b> , SB-269970	GSK	0.34	selective antagonist	preclinical	2000
	<b>12</b> , SB-656104	GSK	1.99	selective antagonist	preclinical	2003
	<b>143</b>	EGIS	0.38	selective antagonist	investigational	2008
	<b>13</b> , DR-4004	Kikuchi et al	2.14	non-selective antagonist	investigational	1999
	<b>144</b> , AS-19	Johansson et al	0.6	selective agonist	investigational	2003
	<b>145</b> , LP-44	Leopoldo et al	0.22	selective agonist	investigational	2004
	<b>146</b> , LP-12	Leopoldo et al	0.13	selective agonist	investigational	2007
	<b>147</b> , LP-211	Leopoldo et al	15	selective agonist	investigational	2008

	<b>148, E-55888</b>	Esteve	2.5	selective agonist	preclinical	2008
	<b>149, 8-OH-DPAT</b>	Pfizer	8.81	non-selective agonist	chemical probe	2004
	<b>150, 5-CT</b>	Eglen et al	1.8	non-selective agonist	chemical probe	1997

**Table Appendix 3** – List of selective 5-HT<sub>7</sub>R antagonists and agonists (compounds in shaded boxes have not yet reached the clinic); Green highlight: successful clinical phase active status;

Rank number cut-off	Rank(x)%	3PBL	3RZE	4IB4	4IAQ	3UON	4MQT	4LDE	2RH1	3NY9
57	1	1.6	5.7	4.6	4.6	4.8	4.8	0.9	4.6	3.8
285	5	1.8	3.9	4.0	3.7	3.4	4.7	1.4	3.8	3.8
569	10	1.8	3.1	3.3	2.9	3.0	4.1	1.6	3.3	3.7
1138	20	1.6	2.4	2.5	2.3	2.3	3.0	1.6	2.8	3.0
1708	30	1.5	2.1	2.1	2.0	2.0	2.2	1.6	2.3	2.4

**Table Appendix 4** - Enrichments of standalone structures – Training Set

5HT <sub>1R</sub>										
Rank number										
criterion	Rank(x)%	3PBL	3RZE	4IB4	4IAQ	3UON	4MQT	4LDE	2RH1	3NY9
31	1	3.3	9.8	<b>16.3</b>	6.5	3.3	3.3	0.0	3.3	0.0
156	5	0.6	6.4	<b>7.7</b>	3.2	2.6	4.5	1.9	4.5	3.9
313	10	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
626	20	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
939	30	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TAAR <sub>i</sub>										
Rank number										
criterion	Rank(x)%	3PBL	3RZE	4IB4	4IAQ	3UON	4MQT	4LDE	2RH1	3NY9
170	1	3.0	3.0	1.8	<b>3.5</b>	1.8	1.2	3.0	3.5	2.4
853	5	2.5	2.0	2.1	<b>2.5</b>	1.2	2.7	3.0	1.7	1.9
1706	10	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0	1.0
3413	20	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0	1.0
5120	30	1.0	1.0	1.0	1.0	1.0	1.2	1.0	1.0	1.0

**Table Appendix 5** - Enrichments of standalone structures – PubChem Validation Sets

5HT <sub>1</sub> R															
F	Rank(1)%			Rank(5)%			Rank(10)%			Rank(20)%			Rank(30)%		
	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)
9	<b>3</b>	<b>1</b>	<b>33.7 (1.78)</b>	137	9	6.6 (6.2E-02)	1644	16	0.98 (5.9E-03)	1855	18	0.98 (5.2E-03)	1855	18	0.98 (5.2E-03)
8	<b>12</b>	<b>1</b>	<b>8.4 (0.3)</b>	425	12	2.9 (2.2E-02)	2900	29	1.01 (1.7E-03)	3114	31	1.0 (0.0)	3114	31	1.0 (0.0)
7	<b>30</b>	<b>5</b>	<b>16.8 (0.2)</b>	754	16	2.1 (1.3E-02)	2999	31	1.0 (0.0)	3126	31	1.0 (0.0)	3126	31	1.0 (0.0)
6	<b>57</b>	<b>6</b>	<b>10.6 (0.1)</b>	1120	17	1.5 (8.7E-03)	3028	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)
5	<b>111</b>	<b>9</b>	<b>8.2 (7.5E-02)</b>	1508	22	1.5 (5.9E-03)	3052	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)
4	198	12	6.1 (4.6E-02)	1886	24	1.3 (4.3E-03)	3065	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)
3	358	14	4.0 (2.7E-02)	2221	26	1.2 (3.2E-03)	3077	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)
2	660	15	2.3 (1.5E-02)	2601	28	1.1 (2.2E-03)	3097	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)
1	1280	20	1.6 (7.3E-02)	2920	30	1.0 (1.2E-03)	3122	31	1.0 (0.0)	3127	31	1.0 (0.0)	3127	31	1.0 (0.0)

**Table Appendix 6** – FrACS enrichments achieved on PubChem 5-HT<sub>1</sub>R validation sets - F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval

TAAR <sub>1</sub>																
F	Rank(1)%			Rank(5)%			Rank(10)%			Rank(20)%			Rank(30)%			
	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	
9	0	0	n.a.	7	0	0.0 (n.a.)	42	3	7.2 (6.0E-02)	289	14	4.9 (1.8E-02)	1054	37	3.6 (7.5E-03)	
8	2	0	0.0 (n.a.)	<b>32</b>	<b>2</b>	<b>6.3 (6.6E-02)</b>	175	7	4.0 (2.2E-02)	1019	41	4.0 (8.0E-03)	3006	66	2.2 (3.1E-03)	
7	5	0	0.0 (n.a.)	<b>110</b>	<b>6</b>	<b>5.5 (3.2E-02)</b>	435	20	4.6 (1.4E-02)	2012	63	3.2 (4.6E-03)	5116	89	1.8 (1.9E-03)	
6	12	0	0.0 (n.a.)	246	9	3.7 (1.8E-02)	877	34	3.9 (8.7E-03)	3338	80	2.4 (2.9E-03)	7269	103	1.4 (1.3E-03)	
5	33	0	0.0 (n.a.)	475	16	3.4 (1.2E-02)	1544	47	3.1 (5.6E-03)	4950	93	1.9 (1.9E-03)	9342	120	1.3 (9.4E-04)	
4	86	4	4.7 (3.4E-02)	886	25	2.9 (7.7E-03)	2571	66	2.6 (3.7E-03)	6973	103	1.5 (1.4E-03)	11384	139	1.2 (6.5E-04)	
3	201	10	5.0 (2.2E-02)	1723	43	2.5 (4.9E-03)	4183	80	1.9 (2.3E-03)	9234	126	1.4 (9.2E-04)	13170	149	1.1 (4.8E-04)	
2	510	16	3.2 (1.1E-02)	3327	64	1.9 (2.9E-03)	6666	95	1.4 (1.5E-03)	11881	142	1.2 (6.0E-04)	14803	159	1.1 (3.1E-04)	
1	1859	36	2.0 (4.3E-03)	6765	98	1.5 (1.4E-03)	10649	5	13 (7.3E-04)	14593	161	1.1 (2.8E-04)	16281	166	1.0 (1.6E-04)	

**Table Appendix 7** - FrACS enrichments achieved on PubChem TAAR<sub>1</sub> validation set - F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval

F	Rank(1)%			Rank(5)%			Rank(10)%			Rank(20)%			Rank(30)%		
	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)	S	A	EF (95% CI)
9	0	0	n.a.	0	0	n.a.	7	1	21.4 [2.9 (3.9E-02)]	103	10	14.6 [1.9 (7.1E-03)]	<b>456</b>	<b>56</b>	<b>18.4</b> [2.5 (3.6E-03)]
8	0	0	n.a.	3	1	50.0 [6.8 (13.8)]	45	2	6.7 [0.9 (7.7E-03)]	484	45	13.9 [1.9 (3.2E-03)]	<b>1538</b>	<b>145</b>	<b>14.1</b> [1.9 (1.7E-03)]
7	0	0	n.a.	39	3	11.5 [1.6 (1.1E-02)]	207	13	9.4 [1.3 (4.1E-03)]	1219	101	12.4 [1.7 (1.9E-03)]	<b>3137</b>	<b>269</b>	<b>12.9</b> [1.7 (1.1E-03)]
6	3	0	0.0 (n.a.)	106	9	12.7 [1.7 ± 6.7E-03]	541	40	11.1 [1.5 (2.7E-03)]	2298	187	12.2 [1.7 (1.3E-03)]	<b>5251</b>	<b>382</b>	<b>10.9</b> [1.5 (7.3E-04)]
5	8	0	0.0 (n.a.)	267	17	9.5 [1.3 ± 3.7E-03]	1105	87	11.8 [1.6 (1.9E-03)]	3936	317	12.1 [1.6 (9.1E-04)]	<b>7657</b>	<b>524</b>	<b>10.3</b> [1.4 (5.4E-04)]
4	36	0	0.0 (n.a.)	680	51	11.2 [1.5 ± 2.4E-03]	2198	173	11.8 [1.6 (1.3E-03)]	6134	438	10.7 [1.4 (6.5E-04)]	<b>10453</b>	<b>653</b>	<b>9.4</b> [1.3 (4.1E-04)]
3	135	4	4.4 [0.6 (3.6E-03)]	1481	107	10.8 [1.5 ± 1.6E-03]	3843	291	11.4 [1.5 (9.1E-04)]	8909	575	9.7 [1.3 (4.7E-04)]	13673	774	8.5 [1.2 (3.1E-04)]
2	451	26	8.6 [1.2 (2.7E-03)]	3195	234	11.0 [1.5 ± 1.0E-03]	6684	454	10.2 [1.4 (6.0E-04)]	12703	733	8.7 [1.2 (3.3E-04)]	17360	909	7.9 [1.1 (2.2E-04)]
1	2071	140	10.1 [1.4 (1.3E-03)]	7792	469	9.0 [1.2 ± 5.2E-04]	12504	728	8.7 [1.2 (3.3E-04)]	18498	1022	8.3 [1.1 (1.9E-04)]	21899	1165	8.0 [1.1 (1.1E-04)]

**Table Appendix 8** - Sequential screening enrichments achieved by the consecutive application of FrAGS and FrACS on the ChEMBL fragment set - FrAGS pre-filtering step EF = 6.7 (4.1E-04) (95% Confidence Interval); F: FrACS cut-off, S: number of selected compounds, A: number of actives, (95% CI): 95% Confidence Interval; Best enrichments highlighted as bold entires.

## NYILATKOZAT

Alulírott Kelemen Ádám Andor kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Budapest, 2018. május 14. ....

Kelemen Ádám Andor