DOI: 10.1002/cmdc.200600147

Predicting Compound Selectivity by Self-Organizing Maps: Cross-Activities of Metabotropic Glutamate Receptor Antagonists

Tobias Noeske,[a, c] Britta C. Sasse,[b] Holger Stark,[b] Christopher G. Parsons,[c] Tanja Weil,[c] and Gisbert Schneider*[a]

The self-organizing map (SOM) principle was introduced by Kohonen in 1982,[1] and has been applied to a variety of tasks in chemistry and chemical biology ever since.[2,3] In this study, we used the SOM algorithm for mapping known ligands according to a topological pharmacophore descriptor (CATS)[4] and predicting potential cross-activities. Our aim was to see whether 1) the descriptor is able to discriminate antagonists of metabotropic glutamate receptors (mGluR) 1 and 5, and 2) the SOM could be used for predicting potential additional binding behaviors of the ligands.

First, an mGluR reference collection containing 338 compounds was compiled including published and Merz in-house structures of noncompetitive group I mGluR antagonists. The collection comprises two subsets: allosteric mGlu1 receptor antagonists (213 compounds), and allosteric mGlu5 receptor antagonists (125 compounds).[5] These molecules cover a broad range of binding activities (K_i values between 1 nM and ≈10 μM) and represent different chemical classes. This mGluR reference library was complemented by the molecules from the COBRA database (v. 3.12; 5376 molecules) containing a broad set of known drugs, leads, and lead candidates affecting a large number of different drug targets.[6]

Subsequently, the molecules were converted to a vector representation giving the scaled occurrence frequencies of topological potential pharmacophore point pairs (CATS2D method).[4,7] In this study, intramolecular distances from zero to nine bonds were considered, resulting in a 150-dimensional vector representation of each molecular compound.

The complete COBRA database was subjected to clustering and mapping onto a two-dimensional grid by the SOM approach. The SOM provides a nonlinear two-dimensional projection of the 150-dimensional data space ("chemical space"), where local neighborhood is conserved. This means, that molecules that are located close to each other on the map are also close in the original high-dimensional space. For SOM training we applied a slightly modified version of the Kohonen algorithm as described previously.[6,8] As a result, all molecules from COBRA were distributed into 225 (15×15) clusters ("neurons" or "receptive fields"). The distribution of these compounds is shown in Figure 1a. It is evident that the SOM is devoid of large patches of empty clusters (<3%) and pronounced densities, which indicates successful mapping and also reflects the diversity of the COBRA entries. After SOM training we projected the mGluR data onto this map and analyzed the resulting distribution patterns. The two mGluR ligand classes form separate localized distributions, where the distribution of the mGluR5 antagonists (Figure 1c) appears to be slightly more focused than the mGluR1 data (Figure 1b). Notably, only 6% of the two ligand classes were clustered together. The SOM was

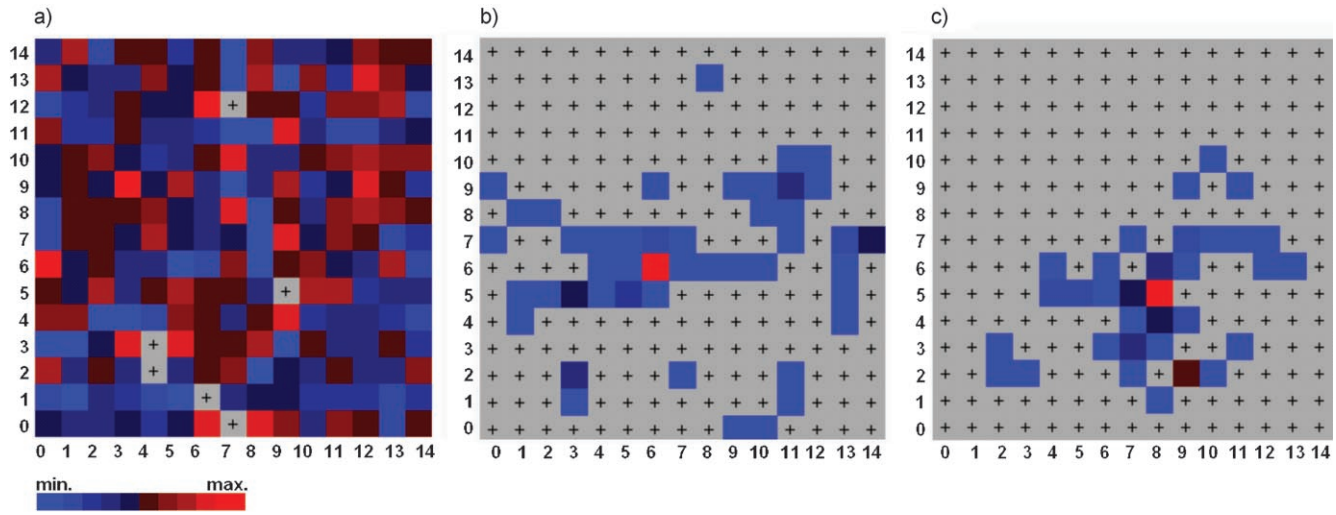

Figure 1. a) SOM projection of the complete COBRA data, b) the mGluR1 antagonists, and c) mGluR5 antagonists. The distribution of the compounds on the map was separately scaled for each figure. Field (6/6) was selected as the target "mGluR1 cluster", and field (8/5) as the target "mGluR5 cluster". Gray fields indicate empty clusters. Note that the map forms a torus.

[a] *T. Noeske, Prof. Dr. G. Schneider*
Johann Wolfgang Goethe University,
Institute of Organic Chemistry and Chemical Biology
ZAFES/CMP, Siesmayerstr. 70, 60323 Frankfurt (Germany)
Fax: (+49) 69-798-24880
E-mail: gisbert.schneider@modlab.de

[b] *B. C. Sasse, Prof. Dr. H. Stark*
Johann Wolfgang Goethe University, Institute of Pharmaceutical Chemistry
ZAFES/CMP, Max-von-Laue Str. 9, 60438 Frankfurt (Germany)

[c] *T. Noeske, Dr. C. G. Parsons, Dr. T. Weil*
Merz Pharmaceuticals GmbH
Altenhöfer Allee 3, 60438 Frankfurt (Germany)

able to discriminate between antagonists of the two mGluR subtypes. This result substantiates earlier findings that both the CATS descriptor and the SOM procedure are suited for clustering compounds according to their pharmacological activity.[9]

Clusters (6/6) and (8/5) revealed the highest density of mGluR1 and mGluR5 reference molecules, respectively (Figure 1). For prediction of potential side-effects or additional binding behavior of the mGluR antagonists, we listed the targets of those COBRA ligands that were co-located in these two clusters. Based on this analysis, mGluR1 antagonists of cluster (6/6) and mGluR5 antagonists of cluster (8/5) were predicted to interact with human dopamine D_2-like receptors, histamine H_1 receptor, and muscarinic acetylcholine (mACh) receptor. For pharmacological testing, we selected representative molecules from each cluster available from the Merz compound collection.

Representatives were defined as being closest to the cluster centroids in descriptor space. Four scaffold classes were found (Figure 2): Molecules **1** and **2** are known mGluR1 antagonists based on a quinoline scaffold,[10] **3** and **4** represent the known mGluR5 antagonists MTEP (3-(2-methyl-thiazol-4-ylethynyl)-pyridine)[11] and an analogue of MPEP (2-methyl-6-(phenylethynyl)-pyridine),[12] **5**–**7** are imidazo[1,2-a]pyrazine derivatives, and **8**–**12** mGluR5 antagonists based on an imidazo[1,2-a]pyrimidine scaffold.[13] In addition, for each of these test compounds individual predictions of target preference were made. This prediction was based on the relative occurrence frequencies of known COBRA compounds in the clusters, for example in the mGluR5 cluster we found $4\times H_1$, $2\times D_2$, $1\times$mACh ligands. Based on these crude statistics, compounds **1** and **2** were predicted to interact with all four targets, compounds **3**-**12** preferably with H_1. We also found aromatase inhibitors co-located in cluster (8/5), which was not further pursued but might be worthwhile to test.

The pharmacological assay results are summarized in Table 1. Although only weak binding constants in the low to medium micromolar range were determined, the results confirm the SOM predictions. Notably, all tested compounds exhibit binding affinity in our H_1 receptor assay, indicating a potential general interaction of mGluR1/5 antagonists with the histamine receptor. Thus, it might be meaningful to test affinities to other histamine receptor subtypes. With the exception of **5**, **8**, and **12**, the compounds showed moderate binding to dopamine D_3 receptors. Only **1** and **2** also exhibited dopamine D_2 receptor binding affinity, all others were inactive at dopamine D_2 receptors, meaning that three of the four substance classes represent D_2-selective chemotypes. The known mGluR1 antagonists **1** and **2** were most "promiscuous". These are the only compounds that comprise additional mACh activity and dopamine D_2 receptor binding affinity.

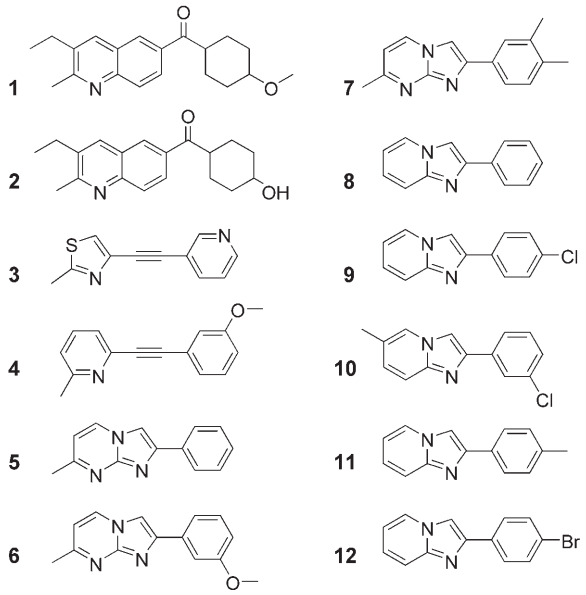

Figure 2. Known mGluR1 (**1**, **2**) and mGluR5 (**3–12**) antagonists, which were selected for activity testing based on the SOM results.

In summary, we successfully applied a topological pharmacophore descriptor and SOM-based clustering to predicting potential activities of known mGluR antagonists. The tested compounds exhibited binding constants in the micromolar range at predicted targets. Whether such activity is of actual pharmacological relevance remains to be shown. Irrespective of the outcome of such studies for the particular compounds presented here, our virtual screening concept might provide a

Table 1. IC_{50} values (mACh, mGluR1, mGluR5) and K_i values (D_2, D_3, H_1) of selected compounds.[a]

Identifier	mACh (M_1) [μM]	D_2 [μM]	D_3 [μM]	H_1 [μM]	mGluR1 [μM]	mGluR5 [μM]
1	54.7 (±.2.8)	91.6 (±13.6)	45.4 (±27.1)	22.0 (±3.7)	0.008[10]	–
2	n.d[b]	80.7 (±0.8)	25.4 (±9.4)	20.6 (±1.1)	0.080[10]	–
3	n.d.	n.d.	76.2 (±46.9)	26.3 (±4.1)	–	0.005[11]
4	n.d.	n.d.	75.6 (±23.5)	33.6 (±3.2)	–	0.036[12]
5	n.d.	n.d.	n.d.	22.9 (±1.9)	–	>30.0
6	n.d.	n.d.	90.9 (±78.4)	16.2 (±7.1)	–	>30.0
7	n.d.	n.d.	16.0 (±6.5)	4.6 (±1.7)	–	28.8
8	n.d.	n.d.	n.d.	36.4 (±32.6)	–	>30.0
9	n.d.	n.d.	53.4 (±37.2)	4.8 (±0.8)	-	>30.0
10	n.d.	n.d.	24.3 (±4.4)	6.5 (±3.4)	–	8.8[13]
11	n.d.	n.d.	32.9 (±16.7)	4.5 (±1.1)	–	>30.0
12	n.d.	n.d.	n.d.	4.3 (±0.6)	–	>30.0

[a] Underlined values indicate the predictions according to the SOM analysis. [b] n.d.: No detectable activity/affinity at the concentrations tested.

basis for early recognition of potential side-effects in lead discovery.

Experimental Section

D_{2short} and D_3 receptor binding assay. Membrane preparations of CHO-cells stably expressing human D_{2short}- and D_3 receptors were used for displacement studies.[14,15] [^{3}H]Spiperone (0.2 nM) served as a radioligand and non-specific binding was determined in the presence of BP 897 (10 μM). Stock solutions (10 mM) of test compounds were prepared with pure DMSO. They were diluted to give final concentration ranges either from 1 μM to 1 mM or from 10 nM to 10 μM, depending on the test compound's affinity. The assay was incubated for 2 h at RT and terminated by rapid filtration through PerkinElmer GF/B glass fiber filters (PerkinElmer Life Sciences, Rodgau, Germany) coated with 0.3% polyethylenimine (Sigma–Aldrich, Taufkirchen, Germany) using an Inotech cell harvester (Inotech AG, Dottikon, Switzerland). Radioactivity was counted using a PerkinElmer MicroBeta Trilux scintillation counter (PerkinElmer Life Sciences, Rodgau, Germany). For all compounds two independent experiments were performed in triplicate. Competition binding data were analyzed by GraphPad Prism (2000, version 3.02, San Diego, CA, USA), using non-linear least squares fit. K_i values were calculated from the IC_{50} values according to the Cheng–Prusoff equation.[16]

H_1 receptor binding assay. Membrane preparations of CHO-cells stably expressing human H_1 receptors were used for displacement studies.[17] [Pyridinyl-5-^{3}H]-pyrilamine (1 nM) served as a radioligand and non-specific binding was determined in the presence of chlorphenamine hydrogenmaleate (10 μM). Stock solutions (10 mM) of test compounds were prepared with pure DMSO. They were diluted to give final concentration ranging from 1 μM to 1 mM. The assay was incubated for 2 h at RT and terminated by rapid filtration through PerkinElmer GF/B glass fiber filters coated with 0.3% polyethylenimine using an Inotech cell harvester. Radioactivity was counted using a PerkinElmer MicroBeta Trilux scintillation counter. For all compounds two independent experiments were performed in triplicate. Competition binding data were analyzed by GraphPad Prism, using non-linear least squares fit. K_i values were calculated from the IC_{50} values according to Cheng–Prusoff equation.[16]

mAChR assay. The increase of intracellular calcium after stimulation with carbachol was measured using the fluorimetric imaging plate reader (FLIPR) and the Ca-Kit (both Molecular Devices). Cells were seeded in black 96-well plates with a clear bottom (CoStar) at a density of 60000 cells per well, and incubated in Ham's F12 medium for one night. Prior to addition of agonist or antagonist, the medium was aspirated, and cells were loaded for 1 h at 37 °C with 150 μL of loading buffer consisting of Ca-sensitive dye (Molecular Devices) reconstituted in HBSS, $MgCl_2$ (0.8 mM), $CaCl_2$ (1.8 mM), probenecid (2.5 mM), and HEPES (20 mM), pH 7.3. Subsequently, plates were transferred to FLIPR to detect calcium increase with the addition of CBC (50 mM final concentration) measured as relative fluorescence units (RFU). If antagonists were tested, these compounds were preincubated for 20 min at RT before addition of CBC. The fluorescence signal increase after addition of agonist reflects the increase of intracellular calcium. Inconsistencies in the amount of cells per well were normalized by using the spatial uniformity correction of the FLIPR software. The mean of replicated temporal data ($n = 5$) was calculated and used for graphical representation. For evaluation of the pharmacology, the calcium changes in response to different concentrations of agonist or antagonist were determined using maximum minus minimum (MaxMin) or an area under the curve (AUC) calculation. All responses (DPM- or RFU-values) were determined as percentage of control, which was defined as the maximum response at 50 nM CBC.

Acknowledgements

This research was supported by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften, Frankfurt. Tobias Noeske would like to thank Merz Pharmaceuticals GmbH for a PhD grant. We thank Dr. P. Sokoloff (INSERM, Paris), and Dr. J. Shine (The Garvan Institute of Medical Research, Sydney) for providing cell lines expressing dopamine D_3 and D_{2S} receptors, respectively. CHO-K1 cells stably expressing human H_1 receptors were generously provided by Dr. H. Timmermann and Dr. R. Leurs (Free University of Amsterdam, The Netherlands).

Keywords: bioinformatics · drug design · Kohonen network · topological pharmacophore · virtual screening

[1] T. Kohonen, *Biol. Cybern.* **1982**, *43*, 59–69.
[2] J. Zupan, J. Gasteiger, *Neural Networks in Chemistry and Drug Design*, Wiley-VCH, Weinheim, **1999**.
[3] G. Schneider, P. Wrede, *Prog. Biophys. Mol. Biol.* **1998**, *70*, 175–222.
[4] G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem.* **1999**, *111*, 3068–3070; *Angew. Chem. Int. Ed.* **1999**, *38*, 2894–2896.
[5] T. Noeske, A. Gutcaits, C. G. Parsons, T. Weil, *QSAR Comb. Sci.* **2005**, *25*, 134–146.
[6] P. Schneider, G. Schneider, *QSAR Comb. Sci.* **2003**, *22*, 713–718.
[7] U. Fechner, L. Franke, S. Renner, P. Schneider, G. Schneider, *J. Comput.-Aided Mol. Des.* **2003**, *17*, 687–698.
[8] G. Schneider, P. Schneider in *Chemogenomics in Drug Discovery* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 341–376.
[9] a) S. Anzali, G. Barnickel, M. Krug, J. Sadowski, M. Wagener, J. Gasteiger, J. Polanski, *J. Comput.-Aided Mol. Des.* **1996**, *10*, 521–534; b) J. Polanski, B. Walczak, *Comput. Chem.* **2000**, *24*, 615–625; c) G. Schneider, M. Netekoven, *J. Comb. Chem.* **2003**, *5*, 233–237; d) A. Teckentrup, H. Briem, J. Gasteiger, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 626–634; e) Y. D. Xiao, A. Clauset, R. Harris, E. Bayram, P. Santago 2nd, J. D. Schmitt, *J. Chem. Inf. Model.* **2005**, *45*, 1749–1758.
[10] D. Mabire, S. Coupa, C. Adelinet, A. Poncelet, Y. Simonnet, M. Venet, R. Wouters, A. S. Lesage, L. van Beijsterveldt, F. Bischoff, *J. Med. Chem.* **2005**, *48*, 2134–2153.
[11] N. D. Cosford, L. Tehrani, J. Roppe, E. Schweiger, N. D. Smith, J. Anderson, L. Bristow, J. Brodkin, X. Jiang, I. McDonald, S. Rao, M. Washburn, M. A. Varney, *J. Med. Chem.* **2003**, *46*, 204–206.
[12] F. Gasparini, K. Lingenhohl, N. Stoehr, P. J. Flor, M. Heinrich, I. Vranesic, M. Biollaz, H. Allgeier, R. Heckendorn, S. Urwyler, M. A. Varney, E. C. Johnson, S. D. Hess, S. P. Rao, A. I. Sacaan, E. M. Santori, G. Velicelebi, R. Kuhn, *Neuropharmacology* **1999**, *38*, 1493–1503.
[13] "Imidazo[1,2]pyridine derivatives as mGluR5-Antagonists": V. Mutel, J.-W. Peters, J. Wichmann, WO02092086, **2002**.
[14] G. Hayes, T. J. Biden, L. A. Selbie, J. Shine, *Mol. Endocrinol.* **1992**, *6*, 920–926.
[15] P. Sokoloff, M. Andrieux, R. Besancon, C. Pilon, M. P. Martres, B. Giros, J. C. Schwartz, *Eur. J. Pharmacol.* **1992**, *225*, 331–337.
[16] Y. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
[17] M. J. Smit, H. Timmerman, J. C. Hijzelendoorn, H. Fukui, R. Leurs, *Br. J. Pharmacol.* **1996**, *117*, 1071–1080.

Received: June 19, 2006
Published online on September 20, 2006

