From Astemizole to a Novel Hit Series of Small-Molecule Somatostatin 5 Receptor Antagonists via GPCR Affinity Profiling

Wolfgang Guba,*,† Luke G. Green,‡ Rainer E. Martin,‡ Olivier Roche,† Nicole Kratochwil,† Harald Mauser,† Caterina Bissantz,† Andreas Christ,§ and Martin Stahl‡

Discovery Chemistry, Lead Generation, F. Hoffman-LaRoche Ltd., 4070 Basel, Switzerland, Discovery Chemistry, Cheminformatics and Molecular Modeling, F. Hoffman-LaRoche Ltd., 4070 Basel, Switzerland, and Discovery, Metabolic and Vascular Diseases, F. Hoffmann-La Roche, Ltd., 4070 Basel, Switzerland

Received September 13, 2007

Abstract: The H1R antagonist astemizole was identified as a somatostatin 5 (SST5) receptor antagonist by a comparative sequence analysis of the consensus drug binding pocket of GPCRs. Subsequently, a similarity analysis of GPCR affinity profiles of astemizole versus a set of in-house GPCR-biased combinatorial libraries revealed new chemical entry points that led to a second lead series with nanomolar binding affinity.

Approaches labeled with the term chemogenomics systematize hit discovery by linking protein families (GPCRs,[a] kinases, ligand-gated ion channels, proteases, nuclear hormone receptors) to small molecule ligands via annotation schemes derived from common recognition motifs or affinity profiles. At Roche, we view chemogenomics as a multidimensional expansion of the medicinal chemistry similarity principle, where chemical similarity is augmented by the additional dimensions of sequence as well as biological similarity. Here we highlight the successful integration of this new paradigm into the lead generation process with the identification of a structurally diverse hit series for the somatostatin 5 receptor starting from the H1R antagonist astemizole.

The term "chemogenomics" appeared first in 2000 in a Vertex press release,[1] where it was described as an approach to "rapidly and simultaneously design multiple lead classes of drugs directed at protein targets in gene families". In a follow-up landmark publication in 2001, chemogenomics was defined as "the discovery and description of all possible drugs to all possible drug targets".[2] In 2001, there was only one further publication by Jacoby et al. that mentioned chemogenomics in the title.[3] Since then, the number of chemogenomics-related publications has been rising steadily, and related terms such as chemical genetics, chemical genomics, and chemical biology (by former Aventis) appeared and circulated with various definitions in the scientific community. Chemogenomics has become so popular in the pharmaceutical industry that two books have been devoted to this topic.[4,5]

Despite the popularity of chemogenomics, there is an obvious mismatch between the number of chemogenomics-related publications claiming an efficiency boost in drug discovery and the description of real applications. We are aware of very few publications dealing with the application of chemogenomics, such as a prospective study by Telik[6] validating the technology, another one from the same company disclosing a hit finding campaign for a drug discovery project,[7] and two recent publications from 7TM Pharma.[8,9] Therefore, we see a need to demonstrate to the scientific community how chemogenomics has been implemented in Lead Generation at Roche to support GPCR hit finding in bridging the gap between project initiation and data delivery from a high-throughput screening (HTS) campaign.

In the preceding publication, we described the discovery of the first nonpeptidic, small molecule, highly selective somatostatin (SST) 5 receptor antagonists by a chemogenomics search strategy.[10] This approach is based on identifying the most similar GPCRs with respect to the consensus drug binding site in the transmembrane region and testing the respective ligands against the target receptor. Among several biogenic amine receptor–ligands, the H1 receptor antagonist astemizole was chosen as a chemical starting point and transformed into a potent hSST5R antagonist, with nanomolar binding affinity being devoid of the original H1R activity. To minimize the risk of failure associated with optimizing a single hit series, we simultaneously initiated the search for a backup series. In this publication, we report on a second, novel series of SST5 receptor antagonists with the general structure **1**, which were discovered in the pre-HTS phase by similarity searches based on GPCR affinity fingerprints.

Both Weinstein et al.[11] and Kauvar et al.[12] pioneered the concept of affinity fingerprints, that is, the measurement of binding affinities toward a reference panel of proteins. The biological profiles are used as similarity descriptors for the identification of lead structures[13,14] and for establishing quantitative relationships between chemical structures and biological activity spectra.[15–19] Within the context of the GPCR chemogenomics initiative at Roche, approximately 5000 compounds containing privileged structure motifs from GPCR ligands were synthesized by parallel chemistry and submitted to Cerep[20] for measuring binding affinities against a panel of 15 GPCRs. The rationale for selecting this panel was founded on a principal component analysis of the ligand binding data of the Cerep BioPrint[21] database versus a group of 30 different GPCRs. A subset of 15 GPCRs representing the major portion of the information content of the original matrix was selected (M2, α2, β1, D1, H1, H2, 5-HT1A, 5-HT4e, 5-HT2A, MC4, Delta, Kappa, NK1, A1, CB1) for further profiling. Percent inhibition data were measured at 10, 1, and 0.1 μM concentrations, from which K_i values were estimated by applying a set of rules (to be described in a forthcoming publication). The relationship between structural similarity and affinity fingerprint diversity was systematically analyzed by neighborhood plots,[22] where the Tanimoto coefficient (Daylight fingerprints[23]) of a compound pair was plotted against the Euclidean distance of the corresponding affinity profiles.

* To whom correspondence should be addressed. Phone: (+41) 61 68 72932. Fax: (+41) 61 68 86459. E-mail: wolfgang.guba@roche.com.
† Discovery Chemistry, Cheminformatics and Molecular Modeling.
‡ Discovery Chemistry, Lead Generation.
§ Discovery, Metabolic & Vascular Diseases.
[a] Abbreviations: GPCR, G-protein-coupled receptor; HTS, high-throughput screening; SST, somatostatin; H1, histamine subtype 1 receptor; R, receptor; h, human; SST5, somatostatin subtype 5; M2, muscarinic subtype 2 receptor; 2α, alpha adrenoceptor subtype 2; β1, beta adrenoceptor subtype 1; receptor; D1, dopaminergic subtype 1 receptor; H2, histamine subtype 2 receptor; 5-HT1A, serotonin subtype 1A receptor; 5-HT4e, serotonin subtype 4e receptor; 5-HT2A, serotonin subtype 2A receptor; MC4, melanocortin subtype 4 receptor; Delta, delta opioid receptor; Kappa, kappa opioid receptor; NK1, neurokinin subtype 1 receptor; A1, adenosin subtype 1 receptor; CB1, cannabinoid subtype 1 receptor; K_i, inhibition constant; SAR, structure–activity relationship; hERG, human ether-a-go-go related gene.
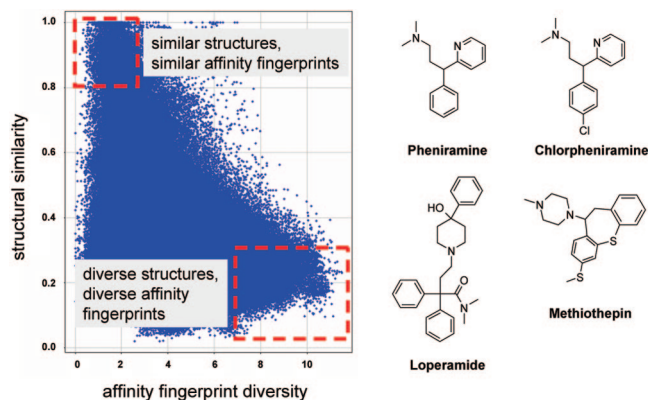
**Figure 1.** Neighborhood plot analyzing the relationship between chemical and biological similarity.
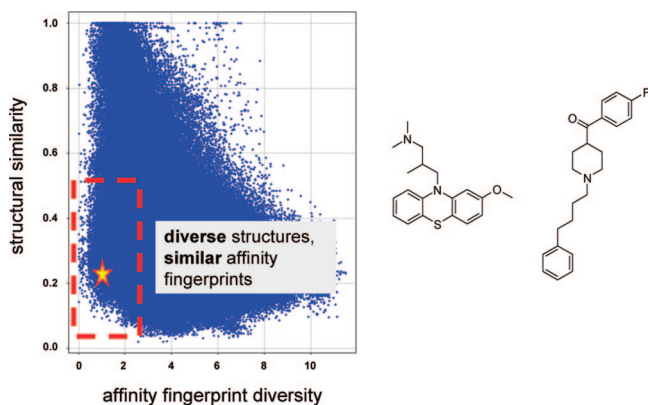
Figure 2. This neighbourhood plot illustrates the basis for a hit-finding strategy that aims at identifying topologically different, but biologically similar, chemical entry points for GPCR targets. The region encompassing biologically similar, but structurally diverse, compound pairs is highlighted with a dashed box. Structural similarity is defined by the Tanimoto coefficient of Daylight fingerprints, biological similarity via the Euclidean distance of the affinity profiles.

As highlighted in Figure 1, the upper left portion of the graph contains similar chemical structures with similar affinity fingerprints, whereras the lower right corner is populated by diverse compound pairs with greatly differing biological profiles. The triangular shape of the plot is indicative of the similarity principle, that is, structurally related ligands imply similar bioactivity profiles and affinity fingerprint diversity increases with structural dissimilarity.

However, the plot in Figure 2 also illustrates that there is a considerable proportion of chemically diverse structures with a similar bioactivity profile.

Thus, as already noted by others, the reverse of the similarity principle,[25] that is, similar biological properties determine structural similarity, is not valid.[12,24] Affinity fingerprints are a biological similarity metric that is not correlated with topological structural descriptors, and structurally diverse molecules may display similar affinity fingerprints. This observation is highly relevant for a typical Lead Generation scenario when additional but diverse chemical entry points are needed for a given target. From the above analysis, the following strategy for GPCR hit finding was defined which is based on detecting biologically similar, but structurally diverse molecules. A seed molecule, for which alternative structures are sought, is profiled against a reference panel of GPCRs. The target receptor of interest is not included in this panel. The affinity profiles of a previously profiled compound library are sorted by similarity to the affinity
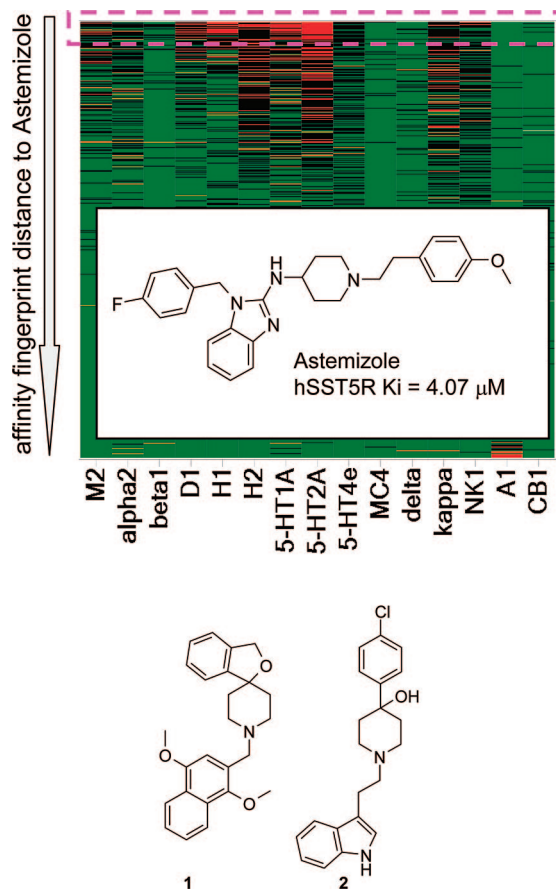

Figure 3. From astemizole to structurally diverse antagonists for the hSST5 receptor. The heatmap color coding reads as follows: red, <0.5 μM; green, >10 μM; black, <10 and >0.5 μM.

fingerprint of the query molecule. Among those biologically most similar compounds hits should be contained, which are structurally diverse compared to the original seed structure. The rationale is that compounds with similar affinity profiles, regardless of their molecular topology, should exploit similar conserved 3D recognition motifs, and therefore, an enrichment of topologically diverse hit structures toward the target of interest is expected. This strategy was applied for identifying an additional hit class in the SST5 project.

As described previously, the H1R antagonist astemizole was identified as hSST5R antagonist with micromolar binding affinity. Because the affinity fingerprint of astemizole was also contained in the Cerep BioPrint compound collection, the structures in the database were resorted with respect to the Euclidean distances of their affinity fingerprints to astemizole (Figure 3). The results of the SST5R radioligand binding assay revealed that the top 3.8% of the resorted compound list (applying an arbitrary cutoff of 4.0 for the Euclidean distance of affinity profiles) yielded a hit rate of 24% (36 hits out of 147 compounds), covering several different chemical classes. The boxed portion of the heatmap shows two representative hits (Figure 3), which are chemically diverse with respect to astemizole but which have similar GPCR affinity profiles.

Both compounds **1** and **2** have pairwise Daylight Tanimoto coefficients[21] with astemizole of 0.34 and 0.39, respectively. Because the usually applied cutoff value for 2D similarity searches is >0.8, these hits would have not been identified in a virtual screening campaign. This result clearly demonstrates that a chemogenomics search strategy based on affinity fingerprint similarity is able to retrieve structurally diverse hit classes.
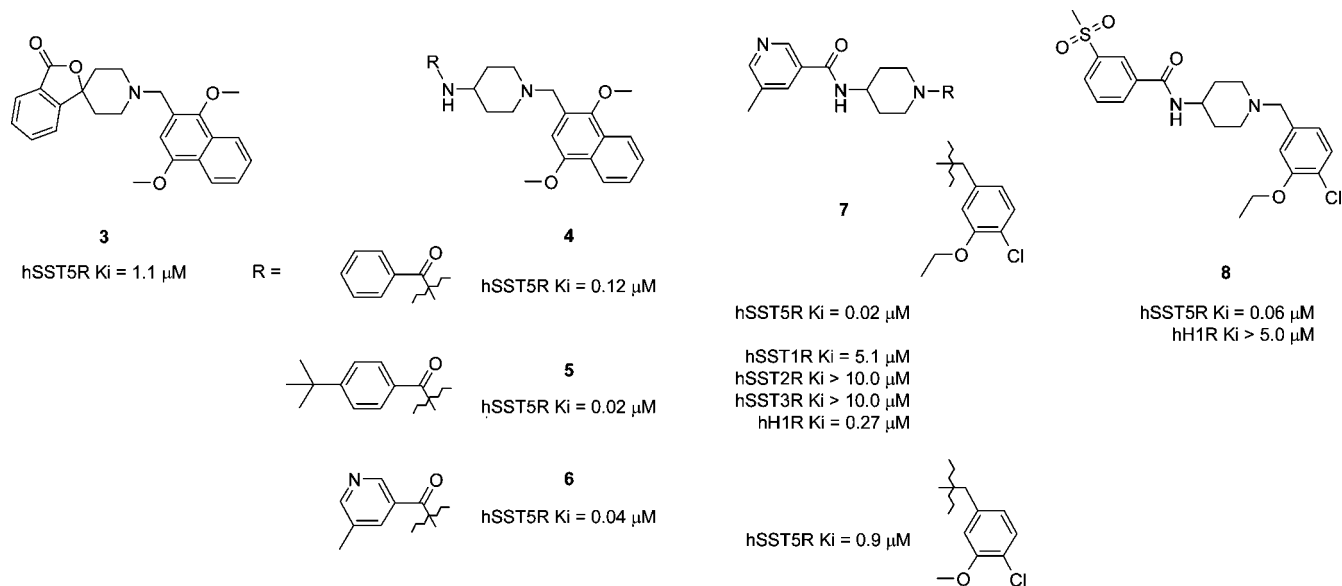
Figure 4. Hit optimization.

From those hits, the spiro compound **3** (hSST5R K_i = 1.1 μM) was chosen as a starting point for further optimization. Interestingly, a subsequent HTS of 600 k compounds yielded only 302 validated hits displaying more than 30% binding at a 3 μM concentration. In addition to the two hit classes derived from the similarity fingerprint analysis, another two hit series were identified from the HTS screen, which, however, had to be terminated in the subsequent optimization phase. Therefore, the chemogenomics strategy compared very favorably to the classical HTS approach.

The spiroketopiperidine **3** was resistant to optimization, however, "opening" of the spiro-group to 4-benzamidopiperidine **4** (Figure 4) afforded a much broader SAR and also showed an improved affinity for hSST5R (K_i 0.12 μM). The readily accessible 4-amino-piperidine scaffold allowed for a very rapid optimization of this compound class. Although further potency increases could be achieved by increasing the lipophilicity of benzamide **5**, the 5-methylnicotinamide **6** was preferred, as it offered greater potential to present a physicochemically more balanced molecule. The similarity to the emerging benzoxazole series communicated earlier was confirmed when, once again, a benzyl headgroup containing a *meta*-ethoxy substituent showed excellent potency (**7**, K_i = 15 nM). This compound shows good selectivity over the other SST receptors, however, a screen against a number of GPCRs demonstrated that there was still affinity for H1R (K_i = 0.27 μM). In addition to the strategy employed in the benzoxazole series, this undesirable cross-activity can be easily removed by the introduction of a polar group in the 3-position of the benzamide **8**. Thus, the series shows good potential for further optimization, where the remaining issue of hERG binding (hERG IC_{50} = 1.3 μM) needs to be addressed.

In conclusion, the comparison of affinity fingerprints of seed structures with a reference library of GPCR chemotypes, which have been profiled against the same target panel, has delivered a novel hit class of hSST5 receptor antagonists. From a broader perspective, this chemogenomics approach can be viewed as the successful application of a multidimensional similarity paradigm, where protein sequences and molecular structures are grouped with appropriate similarity metrics and both target and ligand clusters are linked together by ligand−target annotations, that is, via affinity fingerprints. However, a successful imple-

mentation of this strategy into the lead generation process requires a substantial investment in infrastructure to compile biological data, an enrichment of the corporate compound depository with target-specific chemotypes and setup of a data warehouse to link ligand with target information. With this example it has been demonstrated that chemogenomics is able to play a pivotal role in finding hits for novel GPCR targets without known small molecule ligands.

Supporting Information Available: Experimental details and analytical data for selected compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) See: http://www.vpharm.com.

(2) Caron, P. R.; Mullican, M. D.; Mashal, R. D.; Wilson, K. P.; Su, M. S.; Murcko, M. A. Chemogenomic approaches to drug discovery. *Curr. Opin Chem. Biol.* **2001**, *5*, 464–470.

(3) Jacoby, E. A novel chemogenomics knowledge-based ligand design strategy−Application to G-protein-coupled receptors. *Quant. Struct.-Act. Relat.* **2001**, *20*, 115–123.

(4) *Chemogenomics in Drug Discovery: A Medicinal Chemistry Perspective*; Kubinyi, H., Müller, G., Eds.; Wiley-VCH: Weinheim, Germany, 2004.

(5) *Chemogenomics: Knowledge-Based Approaches to Drug Discovery*; Jacoby, E., Ed.; Imperial College Press: London, 2006.

(6) Hsu, N.; Cai, D.; Damodaran, K.; Gomez, R. F.; Keck, J. G.; Laborde, E.; Lum, R. T.; Macke, T. J.; Martin, G.; Schow, S. R.; Simon, R. J.; Villar, H. O.; Wick, M. M.; Beroza, P. Novel cyclooxygenase-1 inhibitors discovered using affinity fingerprints. *J. Med. Chem.* **2004**, *47*, 4875–4880.

(7) Wadkins, R. M.; Hyatt, J. L.; Wei, X.; Yoon, K. J. P.; Wierdl, M.; Edwards, C. C.; Morton, C. L.; Obenauer, J. C.; Damodaran, K.; Beroza, P.; Danks, M. K.; Potter, P. M. Identification and characterization of novel benzil (diphenylethane-1,2-dione) analogues as inhibitors of mammalian carboxylesterases. *J. Med. Chem.* **2005**, *48*, 2906–2915.

(8) Frimurer, T. M.; Ulven, T.; Elling, C. E.; Gerlach, L.-O.; Kostenis, E.; Högberg, T. A physicogenetic method to assign ligand-binding relationships between 7TM receptors. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3707–3712.

(9) Receveur, J.-M.; Bjurling, E.; Ulven, T.; Little, P. B.; Norregaard, P. K.; Högberg, T. 4-Acylamino-and 4-ureidobenzamides as melanin-

concentrating hormone (MCH) receptor 1 antagonists. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5075–5080.

(10) Martin, R. E.; Green, L. G.; Guba, W.; Kratochwil, N.; Christ, A. Discovery of the first nonpeptide, small-molecule, highly selective somatostatin receptor subtype 5 antagonists: A chemogenomics approach. *J. Med. Chem.* **2007**, *50*, 6291−6294.

(11) Weinstein, J. N.; Myers, T. G.; O'Connor, P. M.; Friend, S. H.; Fornace, A. J., Jr.; Kohn, K. W.; Fojo, T.; Bates, S. E.; Rubinstein, L. V.; Anderson, N. L.; Buolamwini, J. K.; van Osdol, W. W.; Monks, A. P.; Scudiero, D. A.; Sausville, E. A.; Zaharevitz, D. W.; Bunow, B.; Viswanadhan, V. N.; Johnson, G. S.; Wittes, R. E.; Paull, K. D. An information-intensive approach to the molecular pharmacology of cancer. *Science* **1997**, *275*, 343–349.

(12) Kauvar, L. M.; Higgins, D. L.; Villar, H. O.; Sportsman, J. R.; Engqvist-Goldstein, A.; Bukar, R.; Bauer, K. E.; Dilley, H.; Rocke, D. M. Predicting ligand binding to proteins by affinity fingerprinting. *Chem. Biol.* **1995**, *2*, 107–118.

(13) Dixon, S. L.; Villar, H. O. Bioactive diversity and screening library selection via affinity fingerprinting. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1192–1203.

(14) Beroza, P.; Villar, H. O.; Wick, M. M.; Martin, G. R. Chemoproteomics as a basis for post-genomic drug discovery. *Drug Discovery Today* **2002**, *7*, 807–814.

(15) Krejsa, C. M.; Horvath, D.; Rogalski, S. M.; Penzotti, J. E.; Mao, B.; Barbosa, F.; Migeon, J. C. Predicting ADME properties and side effects: The BioPrint approach. *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 470–480.

(16) Fliri, A. F.; Loging, W. T.; Thadeio, P. F.; Volkmann, R. A. Biological spectra analysis: Linking biological activity profiles to molecular structure. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 261–266.

(17) Fliri, A. F.; Loging, W. T.; Thadeio, P. F.; Volkmann, R. A. Analysis of drug-induced effect patterns to link structure and side effects of medicines. *Nat. Chem. Biol.* **2005**, *1*, 389–397.

(18) Fliri, A. F.; Loging, W. T.; Thadeio, P. F.; Volkmann, R. A. Biospectra analysis: Model proteome characterizations for linking molecular structure and biological response. *J. Med. Chem.* **2005**, *48*, 6918–6925.

(19) Paolini, G. V.; Shapland, R. H. B.; van Hoorn, W. P.; Mason, H. S.; Hopkins, A. L. Global mapping of pharmacological space. *Nat. Biotechnol.* **2006**, *24*, 805–815.

(20) See: http://www.cerep.fr/Cerep/Users/index.asp.

(21) BioPrint is a registered trademark of Cerep SA. See the following website: www.cerep.com.

(22) Patterson, D. E.; Cramer, R. D.; Ferguson, A. M.; Clark, R. D.; Weinberger, L. E. Neighborhood behavior: A useful concept for validation of "molecular diversity" descriptors. *J. Med. Chem.* **1996**, *39*, 3049–3059.

(23) *Daylight Toolkit 4.71*, Daylight Chemical Information Systems, Inc., http://www.daylight.com.

(24) Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. Do structurally similar molecules have similar biological activity. *J. Med. Chem.* **2002**, *45*, 4350–4358.

(25) *Concepts and Applications of Molecular Similarity*; Maggiora G. M., Johnson, M. A., Eds.; Wiley-Interscience: New York, 1990.

JM701144E