

COMMENTARY

Chemogenomic approaches to drug discovery: similar receptors bind similar ligands

T Klabunde

Sanofi-Aventis Pharma Deutschland GmbH, Science & Medical Affairs, Drug Design, Industriepark Hoechst, Frankfurt am Main, Germany

Within recent years, a paradigm shift from traditional receptor-specific studies to a cross-receptor view has taken place within pharmaceutical research to increase the efficiency of modern drug discovery. Receptors are no longer viewed as single entities but grouped into sets of related proteins or receptor families that are explored in a systematic manner. This interdisciplinary approach attempting to derive predictive links between the chemical structures of bioactive molecules and the receptors with which these molecules interact is referred to as chemogenomics. Insights from chemogenomics are used for the rational compilation of screening sets and for the rational design and synthesis of directed chemical libraries to accelerate drug discovery.

British Journal of Pharmacology (2007) **152**, 5–7; doi:10.1038/sj.bjp.0707308; published online 29 May 2007

Keywords: chemogenomics; drug discovery; target hopping; lead finding; drug design

Abbreviations: CRTH2, chemoattractant receptor-homologous molecule expressed on Th2 cells; GPCR, G-protein-coupled receptor; 3D, three-dimensional

Sir James Black, winner of the 1988 Nobel Prize in Physiology or Medicine, said that 'the most fruitful basis for the discovery of a new drug is to start with an old drug' (Raju, 2000). Screening of known drugs on a given target and selective chemical optimization of observed 'side activities' can thus provide attractive chemical lead series for drug discovery programs (Wermuth, 2004). Chemogenomics-driven lead finding follows a similar line of thought. However, the pool of compounds for screening is extended from known drugs to a set of bioactive molecules, which has been rationally composed following the paradigm 'similar receptors bind similar ligands' (Klabunde, 2006). Or in other words, for a receptor as drug target of interest, known drugs and ligands of *similar* receptors, as well as compounds *similar* to these ligands, serve as a starting point for drug discovery.

How can receptor or ligand *similarity* be defined? What makes two receptors or two ligands become *similar*? What makes receptors bind *similar* ligands? How to predict ligands for a given receptor? Chemogenomics is an interdisciplinary field that attempts to answer these questions and exploit the answers for the accelerated discovery of novel chemical starting points or lead series (Caron *et al.*, 2001; Kubinyi and Müller, 2004). Within recent years, a paradigm shift has taken place within pharmaceutical research from traditional

receptor-specific studies to a cross-receptor view to increase the efficiency of modern drug discovery. Receptors are no longer viewed as individual and single entities but grouped into sets of related proteins or receptor families, for example, kinases, G-protein-coupled receptors (GPCRs), which are explored systematically. Compounds are profiled against a set of receptors and not tested against single targets. Large structure–activity databases that contain chemical structural information as well as biological activity data have been established by several pharmaceutical companies and commercial vendors. These databases can be mined to derive insights into common properties or structural features among ligands linked to common features of the receptors to which they bind. For instance, common motifs can be identified within the chemical structures of ligand sets binding to a specific receptor class or to a set of receptors sharing a common sequence motif. These insights are then used for the rational compilation of screening sets or the knowledge-based synthesis of chemical libraries to accelerate lead finding.

The review by Rognan (2007) published in this issue gives an overview on how chemogenomic approaches define receptor and/or ligand similarity (examples are given in Figure 1) and presents case studies on how this knowledge has been applied to rational drug design. Numerous chemogenomic approaches apply the classification of target families (such as ion channels, kinases, GPCRs) or protein subfamilies (such as purinergic GPCRs) without taking into account similarities of the assumed ligand-binding sites (in

Correspondence: Dr T Klabunde, Sanofi-Aventis Pharma Deutschland GmbH, Science & Medical Affairs, Drug Design, Industriepark Hoechst, Germany. E-mail: Thomas.Klabunde@sanofi-aventis.com

Received 12 March 2007; accepted 25 April 2007; published online 29 May 2007

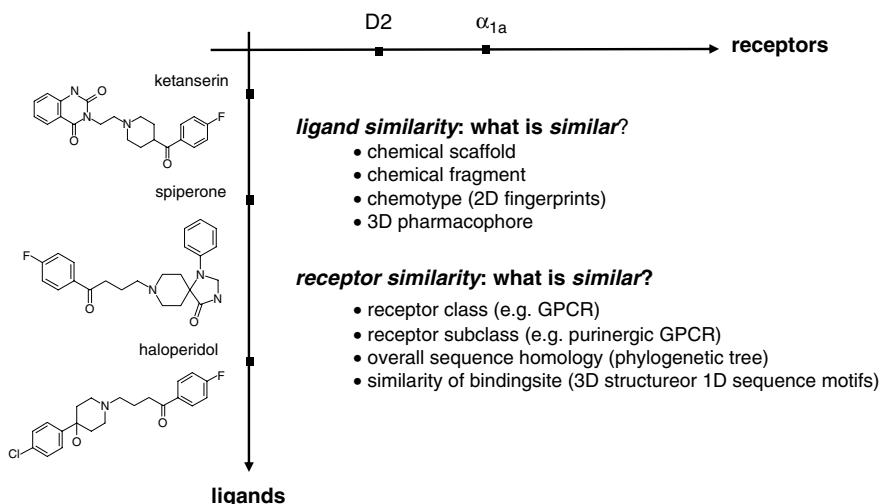


Figure 1 The similarity of ligands and receptors can be defined using different methods or 'descriptors'. The similarity of ligands (for example, ketanserin, spiperone, haloperidol) can be defined either by comparing the chemical scaffold or by looking for identical structural fragments. In addition, several other descriptors (for example, 2D fingerprints, 3D pharmacophores, and so on) and metrics for comparison have been found to be relevant when defining similarity of biologically active molecules. How to define the similarity of receptors (for example, α_{1a} adrenergic receptor and D2 dopamine receptor)? Proteins that belong to the same target family or class, for example, the family of GPCRs, can be considered as similar. A more detailed classification level is defining two receptors as being similar, if they bind the same class of ligands, for example, peptides. This method groups GPCRs in different subclasses such as chemokine receptors, peptide-binding GPCRs or purinergic GPCRs. Another classification level is based on sequence similarity of the receptors. And finally, a further relevant viewpoint for a chemogenomics-driven classification approach is the comparison of two receptors based on the similarity of their putative ligand-binding sites, regardless of their phylogenetic relationship. This is the best indication that a pair of receptors would bind similar ligands.

the review by Rognan, these applications are termed 'ligand-based' chemogenomics). Several practical examples are given in the review by Rognan. Two additional studies exemplifying this approach are worth mentioning. Researchers at Chemical Diversity Lab Inc., have used a scoring scheme based on physicochemical properties for the classification of 'GPCR-ligand-like' and 'non-GPCR-ligand-like' compounds (Balakin *et al.*, 2002). A neural network model was trained with several thousands of known GPCR ligands and non-GPCR ligands and was able to classify correctly more than 90% of randomly selected compound sets. Using this model, the company's compound collection was scored to select 30 000 compounds as a GPCR-focused collection. Another practical example is the design and knowledge-based synthesis of chemical libraries targeting the subfamily of purinergic GPCRs at Sanofi-Aventis (Klabunde, 2006). Common chemical scaffolds and three-dimensional (3D) pharmacophores within the ligands of purinergic GPCRs were identified and chemical libraries comprising 2400 compounds around 5 chemical scaffolds were synthesized. Screening of these libraries for the adenosine A1 receptor (as a member of the purinergic GPCR family) provided three novel A1 antagonist series.

Other chemogenomic applications (termed 'target-based chemogenomic approaches' in the Rognan review) compare and classify receptors based on ligand-binding sites by using sequence motifs or 3D structural information. In many cases, these approaches focus on those residues that are known from molecular recognition studies such as site-directed mutagenesis to be important for binding of the ligand or ligand fragment (termed 'chemoprints' in Klabunde, 2006). In particular, the work of Frimurer *et al.* (2005) exemplifies

the chemogenomics-driven 'target hopping' for the prostaglandin D₂-binding GPCR, CRTH2 (chemoattractant receptor-homologous molecule expressed on Th2 cells). The ligand-binding cavity of the CRTH2 receptor was found to resemble closely that of the angiotensin II type 1 receptor with respect to the physicochemical properties of the amino acids forming the binding site (although both receptors share only a low overall sequence homology). An *in silico* mining of a compound database of 1.2 million entries has been performed using a 3D pharmacophore model adapted from angiotensin II antagonists. Experimental testing of 600 molecules provided several potent hit series with antagonistic activity on the CRTH2 receptor.

Whereas identification of similar targets to the reference target and the identification of similar ligands to the reference ligands represent a two-step process for the ligand-based or receptor-based chemogenomic approaches, there are chemogenomic approaches that attempt to predict ligands for a target of interest in a single step (termed 'target-ligand' approaches). Merging descriptors of ligands and receptors describing putative ligand-receptor complexes and using matrices of biological activity data for a set of compounds profiled against a set of targets, machine learning models have been trained to predict ligands from the NCI database for 55 orphan receptors (Bock and Gough, 2005). Here, the predictions are awaiting experimental validation.

Chemogenomics has provided novel insights into receptor-ligand interaction and molecular recognition by the analysis of large biological activity data sets. In addition, chemogenomics-driven rational drug design often complements high-throughput screening (HTS) for finding

chemical starting points for novel drug discovery programs. The greatest impact of the chemogenomic approaches can be expected for targets with no or sparse ligand information as well as for targets lacking structural 3D data. For these targets, classical drug design strategies like ligand-based and structure-based virtual screening and/or *de novo* design cannot be applied.

References

Balakin KV, Tkachenko SE, Lang SA, Okun I, Ivashchenko AA, Savchuk NP (2002). Property-based design of GPCR-targeted library. *J Chem Inf Comput Sci* **42**: 1332–1342.

Bock JR, Gough DA (2005). Virtual screen for ligands of orphan G-protein coupled receptors. *J Chem Inf Model* **45**: 1402–1414.

Caron PR, Mullican MD, Mashal RD, Wilson KP, Su MS, Murcko MA (2001). Chemogenomic approaches to drug discovery. *Curr Opin Chem Biol* **5**: 464–470.

Frimurer TM, Ulven T, Elling CE, Gerlach LO, Kostenis E, Hogberg T (2005). A physicogenetic method to assign ligand-binding relationships between 7TM receptors. *Bioorg Med Chem Lett* **15**: 3707–3712.

Klabunde T (2006). Chemogenomic approaches to ligand design, In: Rognan D, (Ed). *Ligand Design for G Protein-coupled Receptors*. Wiley-VCH: Weinheim.

Kubinyi H, Müller G (Eds.) (2004). *Chemogenomics in Drug Discovery*. Wiley-VCH: Weinheim.

Raju TNK (2000). The Nobel chronicles. *Lancet* **355**: 1022.

Rognan D (2007). Chemogenomic approaches to rational drug design. *Br J Pharmacol* **152**: 38–52 (this issue).

Wermuth CG (2004). Selective optimization of side affinities: another way for drug discovery. *J Med Chem* **47**: 1303–1314.