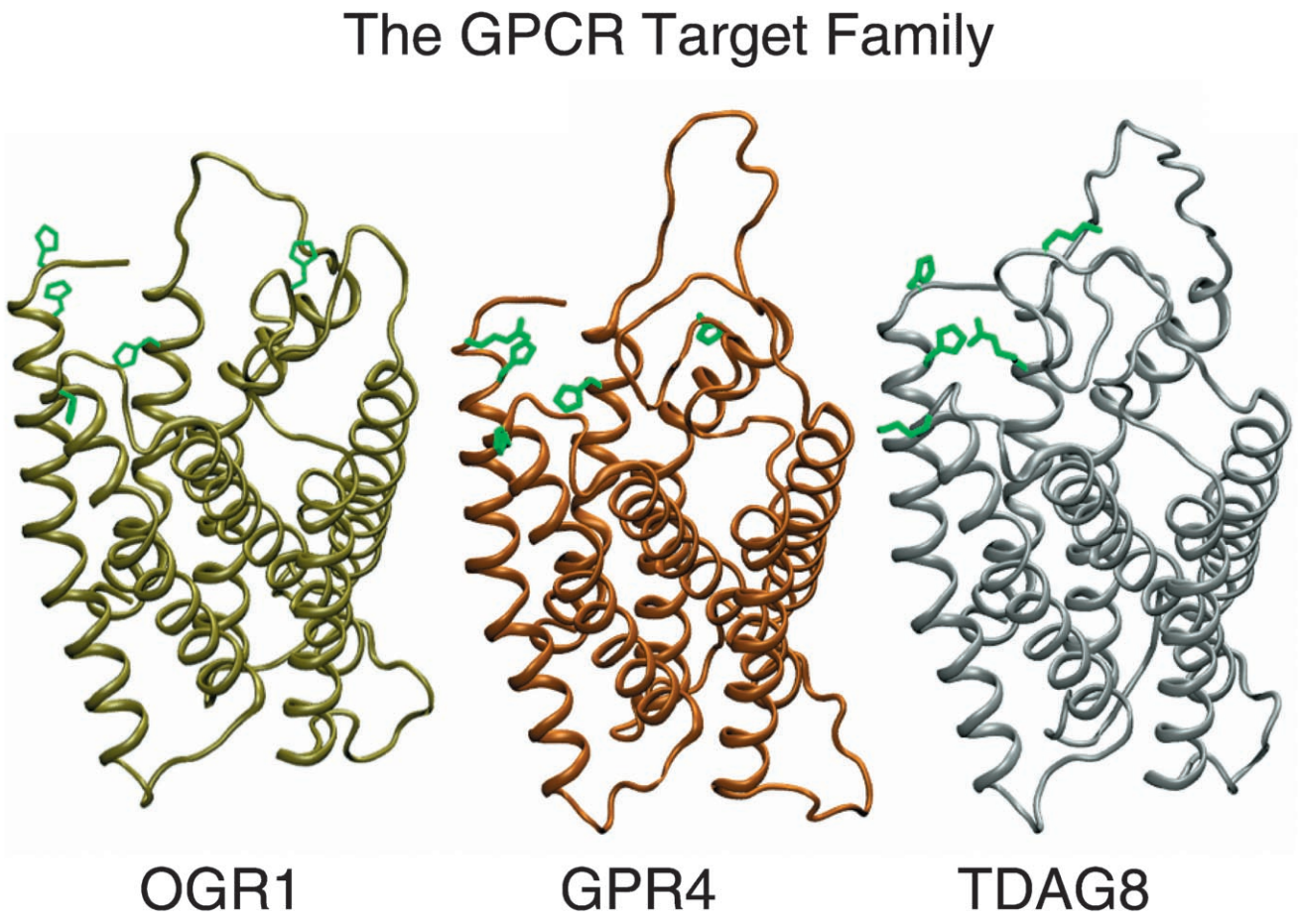


 DOI: 10.1002/cmdc.200600134

DOI: 10.1002/cmdc.200600134

The 7 TM G-Protein-Coupled Receptor Target Family**

Edgar Jacoby,* Rochdi Bouhelal, Marc Gerspacher, and Klaus Seuwen[a]

Chemical biology approaches have a long history in the exploration of the G-protein-coupled receptor (GPCR) family, which represents the largest and most important group of targets for therapeutics. The analysis of the human genome revealed a significant number of new members with unknown physiological function which are today the focus of many reverse pharmacology drug-discovery programs. As the seven hydrophobic transmembrane segments are a defining common structural feature of these receptors, and as signaling through heterotrimeric G proteins is not demonstrated in all cases, these proteins are also referred to as seven transmembrane (7 TM) or serpentine receptors. This review summarizes important historic milestones of GPCR research, from the beginning, when pharmacology was mainly descriptive, to the age of modern molecular biology, with the cloning of the first receptor and now the availability of the entire human GPCR repertoire at the sequence and protein level. It shows how GPCR-directed drug discovery was initially based on the careful testing of a few specifically made chemical compounds and is today pursued with modern drug-discovery approaches, including combinatorial library design, structural biology, molecular informatics, and advanced screening technologies for the identification of new compounds that activate or inhibit GPCRs specifically. Such compounds, in conjunction with other new technologies, allow us to study the role of receptors in physiology and medicine, and will hopefully result in novel therapies. We also outline how basic research on the signaling and regulatory mechanisms of GPCRs is advancing, leading to the discovery of new GPCR-interacting proteins and thus opening new perspectives for drug development. Practical examples from GPCR expression studies, HTS (high-throughput screening), and the design of monoamine-related GPCR-focused combinatorial libraries illustrate ongoing GPCR chemical biology research. Finally, we outline future progress that may relate today's discoveries to the development of new medicines.

1. Introduction

GPCRs are the largest known gene superfamily of the human genome. Around 30% of all marketed prescription drugs act on GPCRs; in addition, they include around 30% of all targets investigated so far, which makes this class of proteins the historically most successful therapeutic target family.[1] As illustrated in Table 1, GPCR-directed drugs cover a wide range of therapeutic indications.[1,2]

"Before cloning", GPCRs were originally defined as receptors that transduce signals from the extracellular compartment to the interior via biochemical processes involving GTP-binding proteins. Molecular cloning of the first receptor genes suggested protein structures similar to rhodopsin, with seven transmembrane α-helical domains (hence "7 TM receptors"). Today, GPCRs are known as extremely versatile receptors for extracellular messengers as diverse as biogenic amines, purines and nucleic acid derivatives, lipids, peptides and proteins, odorants, pheromones, tastants, ions like calcium and protons, and even photons in the case of rhodopsin. GPCRs can form homo- and heterodimers, as well as complex receptosomes, which in a case-by-case-dependent manner can incorporate additional intra- and extracellular soluble and transmembrane proteins.[3,4]

As illustrated in Figure 1, three main families of human GPCRs are known. The rhodopsin-like family A is the largest and the best-studied from the structural and functional points of view. The other two main subfamilies are the secretin-like receptor family B, which bind several neuropeptides and other peptide hormones, and the metabotropic glutamate receptor (mGluR)-like family C. A still separate group is constituted by the receptors of the frizzled family, for which the direct coupling to heterotrimeric G proteins is still a matter of debate.[5]

The human GPCRs have recently been reclassified using phylogenetic analyses into five different groups named GRAFS, which is the acronym for the groups: glutamate, rhodopsin, adhesion, frizzled/taste2, and secretin.[9] The GRAFS system shows some distinct differences to the classification given above: 1) the adhesion receptors, which are expressed on leukocytes and in the CNS, are formed by secretin-like receptors that have a long N-terminal domain including adhesion molecule repeats like epidermal growth factor (EGF) domains, and are likely involved in cell–cell interactions; 2) the taste receptors were reclassified into two subgroups, one within the glutamate group and one together with the frizzled/taste2 group.

[a] *Dr. E. Jacoby, Dr. R. Bouhelal, Dr. M. Gerspacher, Dr. K. Seuwen*
Novartis Institutes for Biomedical Research
4002 Basel (Switzerland)
Fax: (+41) 61-32-46261
edgar.jacoby@novartis.com

[**] *This review will appear later as a chapter in the book* Chemical Biology—From Small Molecules to Systems Biology and Drug Design *(Eds.: S. Schreiber, T. Kapoor, G. Wess), Wiley-VCH, Weinheim, which is scheduled for publication in November 2006.*

While the GRAFS classification is useful, in this chapter for historic reasons we will maintain the A B C nomenclature, as described above.

In the last decades, several GPCR subfamilies were explored systematically in a way that today selective ligands and drugs are known for a large number of the receptors of these families.[10] The elucidation of the human genome with the discovery of the sequences of many novel orphan GPCRs with unknown function provided the basis for further systematization of the exploration of the GPCR superfamily for drug discovery. As a result of the evolutionary conserved commonalities present within a homogenous subgroup of GPCRs, especially for aspects of molecular recognition, it is a very rational expectation that through the further focus within subfamilies it will be possible to find ligands of the new receptors and to discover innovative medicines.[11,12]

The following chapter summarizes milestones of GPCR research and shows how modern chemical biology disciplines and discovery technologies are used today to explore this highly important target family and contribute to new and better medicines.

2. History and Development

Seen in their unparalleled significance for medicine, the history of GPCR chemical biology is in principle as old as the history of pharmacology.[13] Since the beginning of the 19th century, pharmacologists like Ariens, Furchgott, Schild, Blake, and others investigated animal models, isolated organs, and tissues to study the dose-dependent activity of neurotransmitters and peptide hormones as well as natural and synthetic drugs. The targets for most of these molecules later turned out to be GPCRs and ion channels. Many essential concepts were established, such as the binding site and receptor theory, the definitions of agonists and antagonists, affinity and efficacy, as well as the use of radioligands for binding studies and receptor quantification (Table 2). Several methods emerged to analyze the dose response of compounds quantitatively.

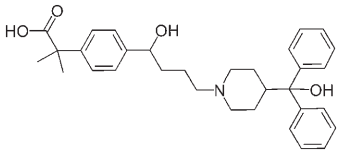
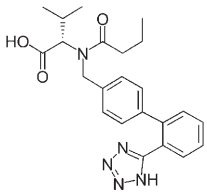
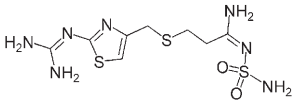
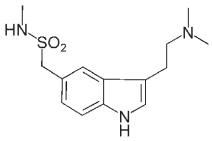
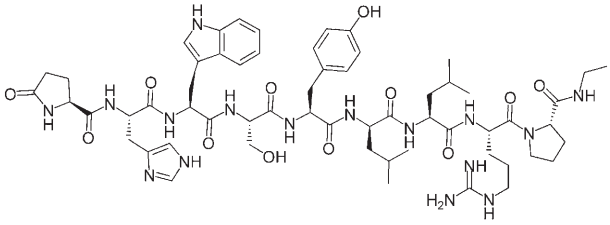
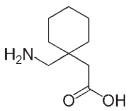
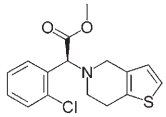
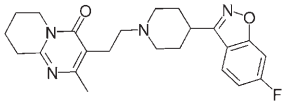
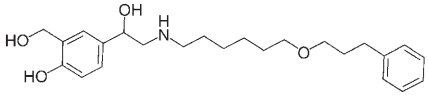
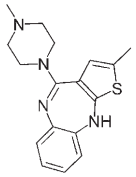
Edgar Jacoby holds a Licence en Sciences Chimiques (1989) from the University of Louvain and completed his Dr.rer. nat. degree under the tutorship of Professor Axel Wollmer in chemistry at the RWTH Aachen in 1993. After postdoctoral work as a NATO and Fulbright Fellow with Professor Michael Weiss in molecular biophysics and structural biology at Harvard Medical School (1994) and the University of Chicago (1995), he joined Servier as Cadre de Recherches in molecular modeling in 1996. In 1999, he joined the combinatorial chemistry group at Novartis Central Technologies as laboratory head for the in silico design of combinatorial compound libraries. In 2001, he was nominated Technology Program Head of Molecular and Library Informatics, and since 2004 has led the in silico group in the Lead Discovery Center of the Novartis Institute for Biomedical Research in Basel.

Marc Gerspacher studied chemistry at the University of Bern. He performed his graduate thesis research at the University of Bern under the direction of Professor Hanspeter Pfander, and received his PhD in 1989. He then moved to the University of California, Berkeley, where he worked as a Swiss National Science Foundation postdoctoral fellow with Professor Henry Rapoport. In 1991 he joined the pharmaceuticals research department of Ciba–Geigy AG in Basel as a medicinal chemist. He continued working as a medicinal chemist, contributing to or leading a number of research projects in various disease areas, first at Ciba–Geigy and after 1996, at Novartis Pharma AG, where he is currently working as drug-discovery project team head in oncology research at the Novartis Institute for Biomedical Research in Basel.

Rochdi Bouhelal received his PhD from the University of Science of Montpellier in 1987 under the direction of Professor Joël Bockaert. In 1987 he moved to Novartis (formerly Sandoz), where he did postdoctoral research in the laboratory of Dr. Anis Mir and Dr. John Fozard in the former cardiovascular department. He joined the Discovery Technologies department in 1992 as a research scientist in assay development and in the high-throughput screening of GPCRs and ion channels. Dr. Bouhelal is also involved in novel functional assay approaches for GPCR in HTS and orphan receptor research.

Klaus Seuwen received a PhD in cell biology and biophysics from the University of Konstanz in 1985. For postdoctoral training he joined the Centre de Biochimie in Nice, where he worked on cell-proliferation control, with a specific focus on G proteins and G-protein-coupled receptors in signal transduction. In 1989 he became a tenured research investigator at the French National Institute for Biomedical Research (INSERM). On leave from INSERM, Dr. Seuwen joined the Sandoz Preclinical Research Center in Basel in 1992, where he continued to work on GPCRs and stem-cell biology in the bone metabolism field. Since 2000 he has directed efforts at Novartis aimed at the characterization of new GPCRs identified through genome analysis.

Table 1. Examples of top-selling GPCR drugs.[a]

Structure	Action	Trade Name	Molecular Entity	Company	Therapeutic Indication	World Sales [US$ millions]
	H_1 antagonist	Allegra/ Telfast®	fexofenadine	Sanofi–Aventis	allergies	1792
	AT_1 antagonist	Diovan®	valsartan	Novartis	hypertension	2214
	H_2 antagonist	Gaster®	famotidine	Yamanouchi	gastric ulcer	656
	$5HT_{1D}$ agonist	Imigran®	sumatriptan	GlaxoSmithKline	migraine	1454
	LH-RH agonist	Leuplin/ Lupron®	leuprorelin	Takeda/Abbott	cancer	904
	$GABA_B$ agonist	Neurontin®	GABApentin	Pfizer	neurological pain	2480
	$P2Y_{12}$ antagonist	Plavix®	clopidogrel	Bristol–Myers Squibb	stroke	5277
	mixed $5HT_2/D_2$ antagonist	Risperdal®	risperidone	Johnson & Johnson	schizophrenia	371
	β_1 agonist	Serevent®	salmeterol	GlaxoSmithKline	asthma	679
	mixed $5HT_2/D_1/D_2$ antagonist	Zyprexa®	olanzapine	Elli Lilly	schizophrenia	4905

[a] Source: IMS Knowledge Link; reported world sales are 12 months to the end of Q1, FY 2005. GPCR drugs cover many therapeutic indications and represent a substantial part of today's marketed medicines.

The molecular nature of the receptors, however, remained unrevealed long after pioneers of biochemistry—including Krebs, Rodbell, and Gilman, working on adrenoceptors—had discovered important elements of the signaling cascade in the 1960s and 1970s.[15,16] The early milestones for the elucidation of the signaling cascades, which couple hormones via the re-

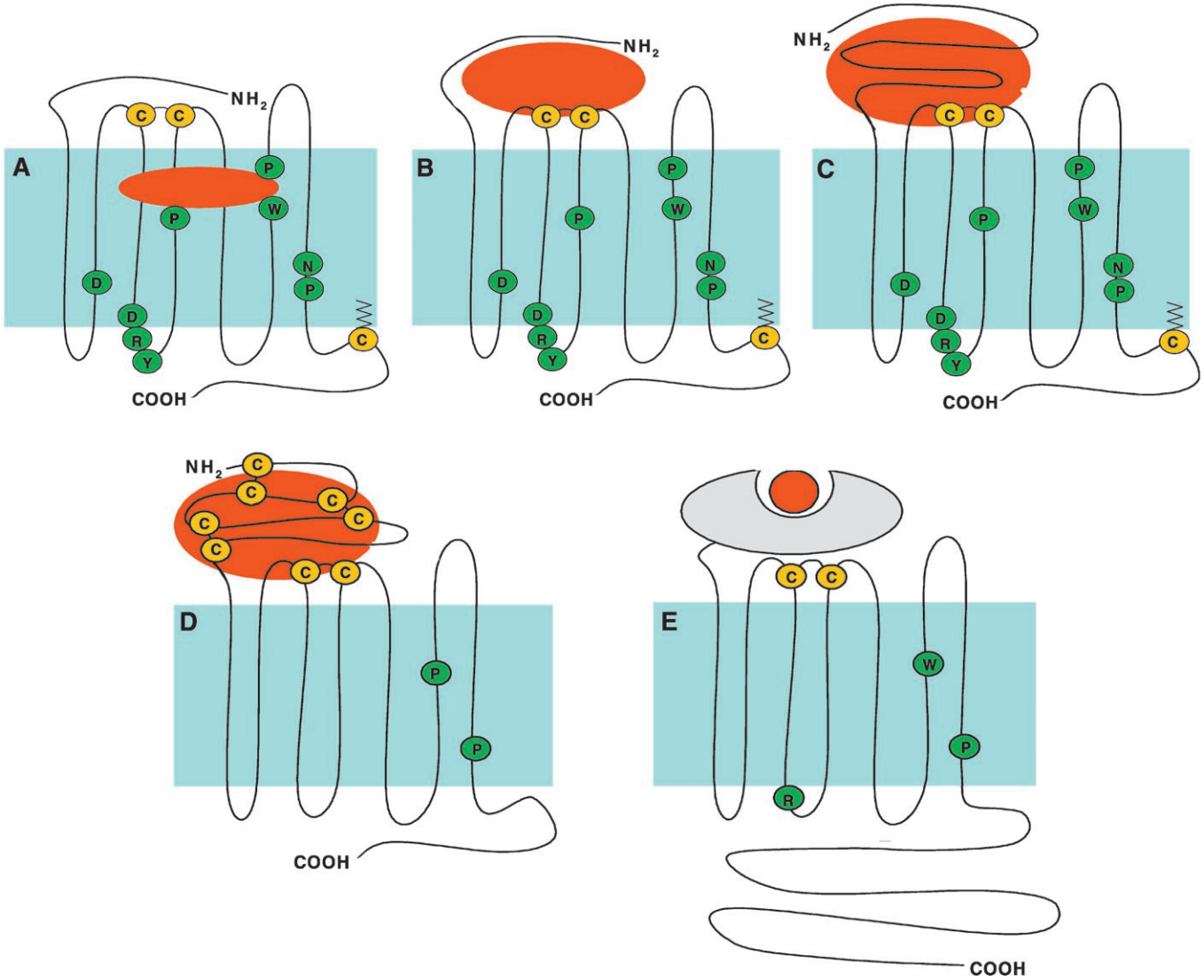


Figure 1. Classification of human GPCRs. As described in greater detail below, the human genome contains 720–800 genes coding for functional GPCRs. Based on their primary sequence, these genes were historically classified into three main families, A, B, C or 1, 2, 3, respectively.[6] Sequences within each family generally share over 25% sequence identity in the 7TM core region. The rhodopsin-like family A is by far the largest subgroup and contains the opsins, olfactory GPCRs, small-molecule/peptide hormone GPCRs, and glycoprotein hormone GPCRs. Family A GPCRs are characterized by several highly conserved amino acids in the 7TM bundle, and there is usually a disulfide bridge linking extracellular loops E1 and E2 (rare exceptions include the melanin stimulating/adrenocorticotropic hormone (MSH/ACTH) and the cannabinoid receptors). Most of the family A receptors have a palmitoylated cysteine residue in the intracellular C-terminal tail. The binding sites of the endogenous small-molecule hormone ligands of class A GPCRs are located within the 7TM bundle (Part A, the ligand binding site is indicated in orange). For peptide and glycoprotein hormone receptors (Parts B and C, respectively), binding occurs at the N terminus, the extracellular loop segments, and the superior parts of the TM helices. Family B comprises 50 GPCRs for peptides such as secretin, calcitonin, and parathyroid hormone. Family B GPCRs are characterized by a relatively long N-terminal tail, which together with the juxtamembrane 7TM regions is implicated in ligand binding and which contains a network of three conserved disulfide bridges defining a globular domain structure (Part D). The 3D structure of the N-terminal ligand binding domain of the mouse CRF_2 receptor was recently determined by high-resolution NMR spectroscopy. As in family A, the family B receptors show a number of conserved proline residues within the TM segments which are thought to be essential for the conformational dynamics of the receptors. Family B receptors appear to couple preferentially to the activation of the effector adenylate cyclase through the G protein G_s, and in general to a lesser extent, through G_i and G_q.[7] Family C GPCRs include the mGluR, the γ-aminobutyric acid type B ($GABA_B$) and Ca^{+2}-sensing (CaR) receptors. This group has 17 members in the human genome, including the pheromone receptors, which form a small family in humans, but a much larger one in rodents. The majority of family C receptors are characterized by very large N- and C-terminal tails, a disulfide bridge connecting the first and second extracellular loops, together with a very short and well-conserved third intracellular loop (Part E). A number of the strongly conserved residues of class A GPCRs are also strongly conserved in class C GPCRs; this is consistent with class A and class C receptors sharing a common ancestor. The ligand binding site is located in the N-terminal domain, which is composed of the so-called venus flytrap module (VFTM) that shares sequence similarity with bacterial periplasmic amino acid binding proteins. In all class C GPCRs except the $GABA_B$ receptor, a cysteine-rich domain (CRD), which contains nine conserved cysteine residues, links the VFTM to the 7TM domain. For $mGluR_1$, the VFTM domain was crystallized in the liganded and unliganded state and was shown to form a disulfide-linked homodimer that undergoes considerable reorganization upon ligand binding.[8] The 11 human frizzled/smoothened receptors control cell development and proliferation mediated by the secreted glycoproteins Wnt and Hedgehog. The N-terminus contains a CRD ligand binding domain with 10 conserved cysteines, all of which form disulfide bonds. The names frizzled and smoothened refer to specific *Drosophila* phenotypes that were linked to mutations in the *Drosophila* orthologs. In general, the N-terminal domains of GPCRs contain N-glycosylation sites for post-translational modification which ensure correct folding in the endoplasmic reticulum and proper cell-surface expression.

Table 2. General pharmacological terms used in this Review to describe compound action at GPCRs.[a]

Term	Definition
Receptor	A cellular macromolecule, or an assembly of macromolecules, that is concerned directly and specifically with chemical signaling between and within cells. The combination of a hormone, neurotransmitter, drug, or intracellular messenger with a cell's receptor(s) initiates a change in cell function.
Agonist	A ligand that binds to a receptor and alters the receptor state resulting in a biological response. Conventional agonists increase receptor activity. *Full agonists* stimulate the maximum response capacity of the system; *partial agonists* do not reach the maximum response capacity. The designation of *full* versus *partial* agonist is system-dependent, and a full agonist for one tissue or measurement may be a *partial agonist* in another. *Inverse agonists* reduce the *constitutive* biological response. A non-endogenous agonist may combine either with the same site as the endogenous agonist (*primary* or *orthosteric site*), or with a different *allosteric* site on the receptor (*allosteric* or *allotopic site*).
Antagonist	A drug that decreases the action of another drug, generally an agonist. Many antagonists act at the same receptor macromolecule as the agonist. In *competitive antagonism*, the binding of the agonist and antagonist is mutually exclusive, either because the agonist and antagonist compete for the same binding site or combine with spatially adjacent and overlapping binding sites (synoptic interaction); a third possibility is that different binding sites are involved, but influence the macromolecule in such way that simultaneous binding is impossible.
Allosteric (allotopic) modulator	A ligand that increases or decreases the action of a (primary or orthosteric) agonist or antagonist by combining with a distinct (allosteric or allotropic) site on the receptor macromolecule.
Desensitization	Decline in the response to continuous or repeated application of agonist.

[a] Adapted from the recommendation of the IUPHAR Committee on Receptor Nomenclature and Drug Classification.[14]

ceptors to the intracellular effector proteins, include the discovery of cyclic adenosine monophosphate (cAMP) by Sutherland as the first characterized secondary messenger,[17] the enzyme adenylate cyclase responsible for its synthesis, and heterotrimeric G proteins as transducers. Intracellular free calcium and inositol phosphates were later characterized as further secondary messengers, and phospholipases, kinases and ion channels emerged as important effector systems downstream of GPCR activation. The list of effectors is ever expanding (Figure 2).[18,19]

Long before the G-protein-coupled receptor proteins were isolated and sequenced, many important therapeutic classes were successfully introduced clinically, including the β blockers, antihistaminics, anticholinergics, analgesic opiates, and neuroleptics.[20] These compounds were developed from discovery to market very rapidly and were successful in the pharmaceutical industry. The sales provided funds to fuel further research in the field. A critical success factor for their discovery was the presence of relatively well-established knowledge of the physiology of the related hormone, and that new chemical compounds were systematically tested in biological models of multiple disease areas in parallel, allowing a complete understanding of their mode of action. Binding profiles of drugs and reference compounds were generated on membrane preparations from different organs, leading to the first clear evidence for receptor subtypes expressed in different tissues.

The development of new protein chemistry technologies such as affinity labeling and affinity chromatography allowed access to enriched and purified sources of receptors and finally introduced the molecular age of GPCR research. With access to a broad range of adrenergic ligands and by coupling the new affinity chromatography procedures with more conventional chromatographic procedures, the Lefkowitz research group was first able to purify the β_2-adrenoceptor in 1979.[21] The proof-of-concept experiment that showed the purified β_2-adrenoceptor protein is indeed the functional receptor was carried out by reconstitution experiments in phospholipid vesicles with purified G protein and the catalytic moiety of adenylate cyclase.[22] The progress in molecular cloning techniques provided access to the DNA sequence of the receptors. Microsequencing of small peptide stretches obtained from the purified adrenoceptors enabled the design of oligonucleotide probes. This allowed researchers at Merck in 1986 to clone the gene and cDNA encoding the hamster β_2-adrenoceptor by using a genomic cDNA library and by identifying overlapping clones that encoded all the peptide stretches defining the full sequence.[23]

The cloning of the β_2-adrenoceptor was a historic breakthrough and catalyzed molecular GPCR research. The analysis of the sequence revealed the homology to bovine rhodopsin, which, since the beginning of the 1980s, was a model system for the study of membrane proteins and the investigation of the molecular basis of vision. Given its remarkably easy access from retinal rod preparations, the sequence of bovine rhodopsin was determined in 1982 by Ovchinnikov using conventional protein sequencing, and cloned in 1983.[24] A structure–function relationship was established for bacteriorhodopsin, the photon-driven and retinal-binding proton pump from the purple membrane of *Halobacterium halobium*, for which Henderson and Unwin, had already determined a 7TM topology using electron microscopy techniques in 1975.[25] Given that the investigation of the signaling mechanisms of rhodopsin had revealed its linkage to the G protein transducin, the knowledge of the β_2-adrenoceptor sequence and signaling mecha-

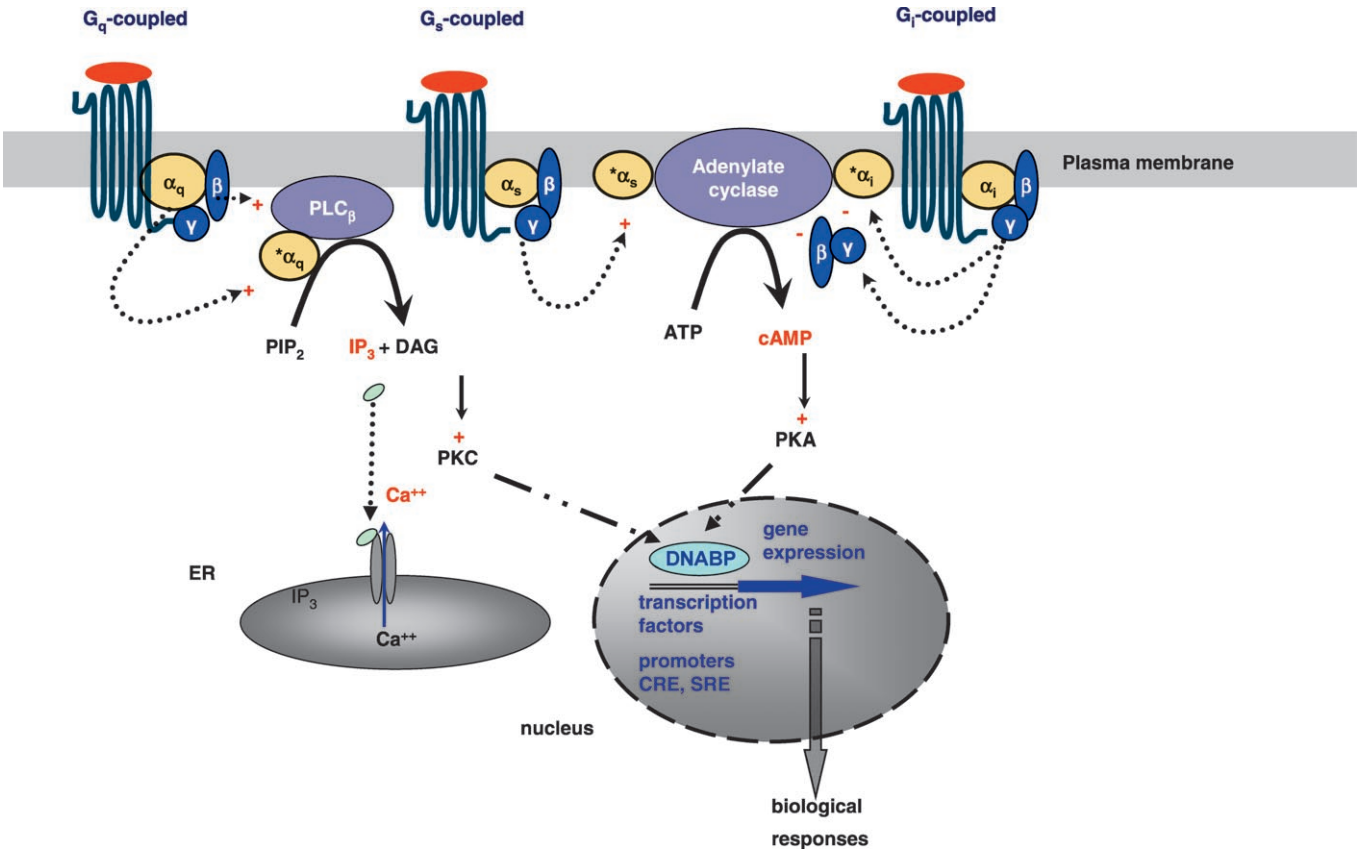


Figure 2. Classical GPCR signaling. Receptors couple to heterotrimeric G proteins to regulate a variety of cell responses. Agonist binding at the receptor leads to exchange of G-protein-bound GDP to GTP. The activated heterotrimer dissociates into the α subunit (symbolized as α*) and the βγ dimer, both of which have an independent capacity to signal forward through the activation or inhibition of effectors. Hydrolysis of GTP to GDP leads to signal termination and re-association of the heterotrimer; regulators of G-protein-signaling (RGS) proteins enhance the intrinsic GTPase activity of the G_{α} subunit. Some G-protein subunits and effectors are expressed ubiquitously; others, only in specific tissues. The 16 mammalian G protein α subunits fall into four broad families based on primary structure and the dependent signaling cascade. Members of the stimulatory $G_{\alpha s}$ family couple to adenylate cyclase to cause an increase of intracellular cAMP levels. The eight members of the $G_{\alpha i/o}$ family inhibit adenylate cyclase and trigger other signaling events. The three members of the $G_{\alpha q/11}$ family activate phospholipase C_{β} (PLC_{β}), resulting in the intramembrane hydrolysis of phosphatidylinositol-4,5-bisphosphate (PIP_2) to inositol-1,4,5-triphosphate (IP_3) and diacylglycerol (DAG). DAG increases the activity of protein kinase C (PKC), and IP_3 triggers the release of Ca^{+2} ions from intracellular stores. Finally, the two members of the $G_{\alpha 12/13}$ family regulate Rho proteins. $G_{\beta\gamma}$ dimers are combinations of five known isoforms of the G_{β} subunit and 13 known isoforms of the G_{γ} subunit. Each individual isoform can associate with a set of effectors and regulators. $G_{\beta\gamma}$ dimers signal a large number of effectors including ion channels, phospholipases, phosphoinositide kinases, and the ras/raf/ERK (extracellular signal-regulated kinase) pathways. Examples of effectors include G-protein-regulated inward rectifying K^+ channel ($GIRK_{1\text{-}4}$), voltage-dependent Ca^{+2} channels (VDCC), phospholipase A_2 (PLA_2), PLC_{β}, and the Na^+/H^+ exchanger (NHE1). The specific function of individual $G_{\beta\gamma}$ dimers is not fully explored. A single GPCR can activate more than one type of G protein. For more detail, see references [18, 19].

nism consolidated the view that rhodopsin would provide an ideal model system for other GPCRs. The speculation on the presence of a large family of such receptors with the 7TM arrangement as a fold characteristic was then confirmed in the following years by the successful cloning of essentially all monoamine GPCRs and several peptide class A GPCRs showing all the characteristic 7TM signatures in hydrophobicity plot analyses.

To this end, cDNA libraries, prepared from cells or tissues known to be rich in certain receptors, were screened by low-stringency hybridization or used for polymerase chain reaction (PCR) amplification of candidate genes using degenerate primers. Proof-of-function was obtained after the expression of the cloned receptor in heterologous cells, which elicited an agonist response. In many cases, however, the functional identity of the cloned receptors could not be matched, and receptors of unknown function were identified. Since the end of the 1990s, such orphan GPCRs became the object of reverse-pharmacology-based drug-discovery programs.[26,27] The successfully completed deorphanization projects resulted in relevant patent and intellectual property claims by the researchers. The elucidation of the human genome in 2001 motivated additional projects in this direction because almost all members of the 7TM GPCR target family became visible at the DNA sequence level, and advanced gene-expression analysis and bioinformatics methods became available for mining and classification purposes. Around 60 orphan receptors were since ligand-paired, and progress was, in many cases, achieved for entire subfamilies such as the trace amine receptors or the endothelial differentiation gene (EDG) receptors.

Besides the systematic exploration in drug-discovery programs, other branches of GPCR research focused on the de-

tailed investigation of receptor signaling and regulation. Generally, direct signaling via secondary messengers resulting in immediate cell responses can be distinguished from the persistent activation of gene expression in the nucleus. The discovery of mutations conferring constitutive receptor activation led to the identification of receptor signaling in the absence of agonist ligands; these were later related to a number of diseases.[28] For example, in Jansen's disease[29] the hypercalcemia and skeletal dysplasia found in many cases is the result of a constitutively overactive parathyroid hormone/parathyroid hormone related protein (PTH/PTHrP) receptor, which carries a point mutation. Studies of the constitutive activity of receptors led to the demonstration of inverse agonism both in vitro and in vivo. In the extended ternary complex model of receptor activation, inverse agonists are ligands that preferably bind to and stabilize the inactive conformational state of the receptor and therefore reduce background signaling.[30] Many receptors show a weak constitutive activity in specific cell systems following overexpression, and this can be used to determine the coupling mechanisms engaged downstream.

Using mutagenesis and chimeric receptors, the ligand binding domains and intracellular domains interacting with G proteins and other effectors were determined.[31,32] Multiple signaling roles and signal-switching mechanisms were discovered for many GPCRs. For instance, the β_2-adrenoceptor signals upon initial agonist binding via the G_s pathway. Protein kinase A (PKA)-mediated phosphorylation within the third intracellular loop switches the signaling specificity toward G_i signaling pathways. A subsequent change of the signaling properties occurs through G-protein-coupled receptor kinase (GRK)-mediated phosphorylation of the receptor C-terminal tail, resulting in binding of β-arrestin proteins, which mediate receptor down-regulation via clathrin-coated pits. The internalized complexes subsequently undergo regulated endosomal sorting either toward lysosomal degradation or recycling back to the plasma membrane. β-Arrestin also acts as a scaffold protein for other signaling pathways and recruits, for instance, the c-Src kinase via the poly-Pro-SH3 domain and thus activates mitogen-activated protein (MAP) kinase signaling. Also, G-protein-independent signaling toward the NHE1 ion exchanger was observed. This occurs via the Na^+/H^+ exchanger regulatory factor (NHERF) protein interacting through its post-synaptic density-95, disc large, zonulla occludens-1 (PDZ) domain with the PDZ-binding motifs found at the C terminus of several GPCRs.[33]

The investigation of the mechanism of agonist-induced receptor signaling, desensitization, internalization, trafficking, and recycling resulted in the discovery of many proteins that interact with GPCRs and which are collectively called G-protein-coupled receptor interacting proteins (GIPs).[34,35] The GIPs link GPCRs to large protein networks called receptosomes, the mechanistic investigation and exploration of which for drug discovery is the subject of intense research activity. We elaborate more on this topic at the end of the chapter.

3. General Considerations

3.1 GPCRs in human and other genomes

The human genome and genomes from several other species (mouse, rat, zebra fish, *Drosophila*, *C. elegans*) are now relatively well analyzed with respect to GPCRs, and it turns out that these receptors constitute the largest gene family in mammals. The most recent studies concluded—depending on the stringencies of the different bioinformatics data-mining methods used—the presence of 720–800 human GPCRs, which accounts for about 2% of the human genome. These include ≈380 unique functional nonolfactory/nonsensory GPCR sequences, for which endogenous ligands are expected and which are therefore referred to as *endo*-GPCRs.[9,36] The *endo*-GPCR group has attracted much attention in recent years. These receptors are expressed in different tissues and regulate various aspects of physiology. A recent comparative investigation of the human and mouse *endo*-GPCR repertoire[36] revealed 367 human and 392 mouse GPCRs; 343 were found in common to both species. The human receptors without orthologs in mice include several orphan receptors such as the melanin-concentrating hormone subtype 2 (MCH_2) receptor and the recently identified receptor for the eosinophil chemoattractant 5-oxo-eicosatetraenoic acid. Of the 367 human GPCRs, 284 belong to the rhodopsin-like class A, 50 to secretin receptor-like class B, 17 to class C, and 11 to the frizzled-smoothened receptor-like class F/S. Among the 392 mouse GPCRs, 313, 47, 17, and 10 belong to classes A, B, C, and F/S, respectively. The cataloguing of these receptors according to ligand specificities reported in the literature identified 224 human and 214 mouse GPCRs with known ligands. The remaining 143 human and 178 mouse GPCRs have no known ligands and are therefore orphan receptors. Among the orphan receptors, 98 human and 136 mouse receptors belong to class A, 34 human and 31 mouse receptors belong to class B, six receptors belong to class C in both species, and none belong to class F/S.

Olfactory receptor genes represent the largest mammalian subgroup. They are class A receptors encoded by single exons and are transcribed in the olfactory epithelium, where they interact specifically with the G protein G_{olf} to transduce odorant signals. They provided the basis for the understanding of odor recognition, which was awarded in 2004 with the Nobel Prize for Medicine and Physiology to Buck and Axel.[37] For some olfactory receptors, expressed sequence tags (ESTs) were picked up in peripheral organs (for example, the Prostate-Specific Gene Receptor (PSGR)). However, the significance of these findings remains unclear at present. Especially for the human olfactory receptor family, it is not yet entirely clear which of these receptors are functionally expressed, as about 50% of the genes identified likely represent pseudogenes. In mice, the majority of olfactory receptors appears to be functional. The annotation and functional characterization of olfactory receptors is rapidly evolving, and specific databases have been created that follow recent developments 'online' (Table 3). In addition to olfactory receptors, taste and pheromone receptors are identified as chemosensory.

Table 3. Publicly available internet-based molecular informatics resources that provide relevant information for GPCR chemical biology research.

Internet Resource URL	Specification of GPCR-Related Information Available
http://www.iuphar-db.org/iuphar-rd/index.html	Official database of the IUPHAR Committee on Receptor Nomenclature and Drug Classification, which includes information on name synonyms, structure, functional assays, ligands, agonist and antagonist potencies, radioligand assays, transduction mechanisms, receptor distribution, tissue function and phenotype.
http://kidb.bioc.cwru.edu/	Database of NIMH Psychoactive Drug Screening Program; pharmaco-informatics systems with a strong focus on GPCR pharmacology and profile structure–activity data.
http://www.gpcr.org/7tm	GPCRDB: Information system of CMBI in Nijmegen contains information about sequences, multiple sequence alignments, phylogenetic trees, 3D models, GPCR mutation data, and ligand binding constants.
http://bioinfo-pharma.u-strasbg.fr/gpcrdb/gpcrdb_form.html	hGPCRdb: The human druggable GPCR database at the University Louis Pasteur of Strasbourg provides searching capabilities for chemogenomics analyses of the 7TM and binding cavity domains of human GPCRs.
http://senselab.med.yale.edu/senselab/ORDB/	Olfactory Receptor Database of the SenseLab project at Yale University which is a long-term effort to build integrated, multidisciplinary models of neurons and neural systems using the olfactory pathway as a model. The database provides metadata of gene and protein sequences of olfactory receptors.
http://tinygrap.uit.no	The GRAP database at the University of Tromsø contains information of mutants of family A GPCRs with a detailed description of the ligand binding and signal-transduction properties.
http://umber.sbs.man.ac.uk/dbbrowser/gpcrPRINTS/ http://bioinformatics.biol.uoa.gr/PRED-GPCR/ http://www.soe.ucsc.edu/research/compbio/gpcr-subclass/	A diagnostic bioinformatics resource at the University of Manchester that profiles a query sequence against the PRINTS fingerprint database to determine most similar families or receptor subtypes. Additional bioinformatics classifiers of GPCRs exist at the University of Athens and the University of California, Santa Cruz, and are based on hidden Markov model (HMM) and SVM methods, respectively.
http://chembank.med.harvard.edu/ http://pubchem.ncbi.nlm.nih.gov/	ChemBank at Harvard University and Pubchem at the NCBI are chemoinformatics databases for small molecules and their biological activities. Both systems are supported by the NCI's initiative for chemical genetics.
http://www.ebi.ac.uk/interpro/	InterPro at EBI is a general bioinformatics database of protein families, domains, and functional sites in which identifiable features found in known proteins can be applied to unknown protein sequences.

The pheromone receptors play an important role in modulating behavior in rodents; whether they are involved in human behavior is a matter of debate. Pheromone receptors belong to class C. They are specifically expressed in the vemeronasal organ in rodents, which is a specific structure separate from, but in proximity to, the main olfactory epithelium. Whereas there are more than 100 active receptors in mice, only 11 have been identified in humans, and their ligands are unknown.

Taste receptors come in two families, which are rather well conserved between human and mouse. One group belongs to class C and has three members (T1R1, 2, and 3); these receptors form heterodimers like $GABA_B$ receptors, and the different entities formed are responsible for detecting sugars and the amino acid glutamate. The second group of taste receptors are class A-like (T2Rs) and in humans comprises more than 30 receptors, which appear to be involved in detecting bitter tastes. All taste receptors are expressed exclusively in the tongue, and there is a separation between cells that express T1- and T2-type receptors.

The opsins represent the highly interesting small family of light-detecting GPCRs.[38] In addition to the four well-known opsins operating in rod and cone cells, there are four additional opsin-related receptors (RGR opsin, peropsin, melanopsin, encephalopsin) that are likely able to bind chromophores and appear to play interesting roles in light-sensing outside the well-described primary phototransduction processes. For instance, melanopsin may be involved in the control of circadian rhythms. ESTs for encephalopsin were isolated from several tissues, including brain and skin.

The genome of the nematode *C. elegans* was the first to be sequenced in full, followed by that of *Drosophila* shortly after. These very distantly related organisms share with mammals the presence of receptor systems for monoamines, acetylcholine, $GABA_B$, glutamate, Wnt glucoproteins, and several neuropeptides, inferring their potential use as model organisms to explore the biology of the conserved receptor systems.[39]

Virally encoded GPCRs might have a direct role in human diseases. Indeed, the GPCR from Kaposi's sarcoma-associated herpes virus has recently been implicated in Kaposi's sarcomagenesis, and the human cytomegalovirus-encoded GPCRs have been implicated in atherosclerosis. Given the versatility of GPCR signaling and its wide involvement in physiological processes, it is not surprising that viruses have evolved to exploit these receptors to their advantage.[40]

3.2 Strategies for the deorphanization of GPCRs

Deorphanization, the identification of activating ligands for previous orphan receptors, is a key task in reverse molecular pharmacology. Identifying receptor–agonist pairs usually allows the rapid elucidation of the physiological role of both partners, sometimes putting them in unexpected context. Thus, the identification of orexin unexpectedly led to an understanding of narcolepsy; the discovery of pH-sensing receptors triggered new experimental approaches in several areas of biology. Although bioinformatics methods were initially helpful to successfully direct ligand-pairing experiments as illustrated by examples given in Section 3.5, deorphanization strategies rely on biological screening of orphan GPCRs expressed in specific recombinant expression systems such as immortalized mammalian cells, yeast, and *Xenopus* melanophores.[26,27]

The agonist ligand libraries used for deorphanization include small molecules, peptides, proteins, lipids or tissue extracts, which are specifically selected as described in Section 3.4. The identification of an activating agonistic ligand of the cell-surface-expressed receptors is dependent on the activation of an intracellular signaling cascade. The difficulty for the assay design is that the signaling cascade is not known for a new orphan receptor. Therefore, generic assay systems amenable for HTS needed to be designed to allow the screening of large surrogate ligand collections.

One of the most successful approaches for deorphanization uses fluorescent imaging plate reader (FLIPR) screening technology, which detects ligand-induced intracellular Ca^{+2} mobilization. To direct the signaling via the PLC Ca^{+2} readout, the receptors are transiently expressed in mammalian cells in the presence of one or more cocktails of promiscuous G proteins such as $G_{q15/16}$, which couple to the majority of GPCRs.[41] Alternatively, cocktails of the engineered chimeric G proteins such as $G_{\alpha qi5\text{-}6}$ or $G_{\alpha qs5\text{-}6}$, in which the C-terminal five or six amino acids of $G_{\alpha q}$ have been replaced by the corresponding amino acids of $G_{\alpha i}$ or $G_{\alpha s}$ in order to redirect coupling of $G_{\alpha i}$- or $G_{\alpha s}$-specific receptors via PLC_{β} can be used.[42] Through mechanisms that are not yet fully understood, pre-stimulation of some cell types with agonists of G_q-coupled receptors dramatically sensitizes these cells to stimulation by G_i and G_s-coupled receptors, again linking such receptors to the calcium signaling system.

GPCRs have been successfully expressed in yeast and coupled to the endogenous mating response pathway. Yeast-based assays use a variety of stable expressed synthetic G proteins, and the readout is linked, for example, to the expression of β-galactosidase or other reporter genes. The use of *Xenopus* melanocytes (frog skin cell) for transfection with mutant orphan GPCRs, which increased constitutive activity, represents an alternative to the mammalian and yeast expression systems. In response to selective GPCR signaling via $G_{\alpha s}$ or $G_{\alpha q}$, the melanosomes disperse the melanin pigment and cause a darkening of the cells. Conversely, when signaling is via $G_{\alpha i}$, the melanosomes aggregate and cause a lightening of the cells. Activation and signaling can thus specifically be determined by simple measurements of light transmittance. The so-called constitutively activating receptor technology (CART), which is limited in compound throughput, provides the advantage of identifying agonist and inverse agonist in the same experiment.[43]

The validation of the hits includes testing of the possible interference with endogenous receptors of the heterologous expression system and is followed by selectivity screening on other GPCRs and further investigation with cell-based, tissue, or in vivo models. These experiments help to determine the physiological role of the newly discovered ligands and receptors.

There are limitations in such screening strategies, as the heterologous expression systems may not provide a permissive context for signaling. For example, the class B GPCR calcitonin receptor-like receptor (CRLR) requires the presence of single-transmembrane domain receptor-activity-modulating proteins (RAMPs), which regulate the transport to the membrane and ligand specificity properties of the receptor. Depending on the RAMP subtype, CRLR can act a calcitonin-gene-related peptide or adrenomedullin receptor.[44] Other receptors are active only as heterodimers, as was initially demonstrated for the $GABA_B$ receptors, requiring the co-expression of the two partner receptor proteins.

C5L2, a receptor which shares homology to the C3a and C5a anaphylatoxin receptors, is currently thought to work simply as a ligand sink without any classical signaling activity. Similarly, the chemokine receptor D6 is thought to bind several chemokines with the only purpose to internalize and degrade them. Unfortunately, it is not possible at present to predict with reasonable certainty from the primary sequence of a receptor whether or not it is involved in signaling. This raises the possibility for other orphan receptors that either do not signal or use alternative G-protein-independent signaling pathways. These examples illustrate the need for the development of novel screening and imaging technologies that report, for instance, on receptor translocation of proteins between subcompartments of living cells using light resonance energy transfer either based on fluorescence (FRET) or bioluminescence (BRET).[45,46]

Other receptors such as viral GPCRs (like ORF74 of Kaposi's sarcoma-associated herpes virus), are highly constitutively active and function in the absence of ligand, which raises the possibility for the presence of other ligandless orphans. Again, other orphan receptors might only play roles in intracellular mechanism, like acting as trafficking factors via heterodimerization or being expressed in the membranes of intracellular organelles; exogenously applied non-membrane-permeable ligands will not activate such receptors. The correct plasma membrane localization of orphan receptors studied should be controlled using immunocytochemistry methods. There are several cases in which the reported receptor agonists may not be those that are physiologically relevant. For instance, the receptors HM74 and HM74a respond to nicotinic acid (niacin), a clinically useful molecule that normalizes dyslipidemia, but the physiological first messenger(s) remain to be discovered. In some cases, the original reports describing new receptor–ligand pairings were not reproducible. For instance, the related receptors OGR1, GPR4, and TDAG8 were described several

years ago as receptors for lipid messengers. Later, it was demonstrated that OGR1, GPR4, and TDAG8 may in fact be considered as genuine pH-sensing receptors.[47]

3.3 Structural biology of GPCRs and molecular modeling of ligand–receptor interactions

Until the year 2000, when the first crystal structure of bovine rhodopsin was solved by Palczewski and co-workers at a resolution of 2.8 Å (which was later refined in 2004 to 2.2 Å and for which in total seven crystallographic conformational states are deposited at the Protein Data Bank (PDB)[48]), structural biology investigations of GPCRs were limited to indirect mutagenesis and second-generation affinity labeling methods based on the substituted-cysteine accessibility method (SCAM) in which sulfhydryl-reactive affinity reagents are combined with either wild-type or a series of substituted-cysteine mutant receptors (such as D2 receptors). The first 3D molecular models were based on the analysis of the 2D projection maps generated from cryo-electron microscopic data of 2D crystals of rhodopsin and the analogous bacteriorhodopsin, for which an X-ray crystal structure at 2.5 Å resolution became available in 1997 using microcrystals grown in lipid cubic phases.[49] The comparison of the 3D structures of rhodopsin and bacteriorhodopsin clearly showed differences in the length of the loop and helix segments and of the relative arrangement, tilts, and kinks of the individual helices among the two proteins; such differences had been already inferred to exist based on the 2D projection maps. Whereas the early 3D models based on the bacteriorhodopsin template were able to explain the data generated from mutagenesis experiments[50,51] to some degree, the quality of these analyses became clearly improved when the 3D structure of the bovine rhodopsin became available. This applies especially for the class A GPCRs, which, despite a sequence similarity to rhodopsin of only 20–30%, share characteristic signature motifs in each TM helix.[52] The main ligand binding site of small-molecule hormones and nonpeptidic agonists and antagonists is located within the central crevice of the 7TM bundle, in analogy to the lipophilic binding pocket of retinal in the light-sensing proteins. This is a remarkable similarity, especially for the overlap between the positions of the proposed ligand contact residues and the positions of the retinal contact residues in rhodopsin. The extracellular side involved in ligand binding appears to form a receptor-specific binding site, while the cytoplasmic side and the ends of the transmembrane helices toward the cytoplasm are significantly more conserved.

Illustrative examples include work on the β_2-adrenergic, serotonin $5HT_{1A}$ (Figure 3), neurokinin-1 (NK_1), adenosine A_3, purine $P2Y_1$, angiotensin AT_1, and chemokine CCR_2 receptors,[2,52–54] for which 3D models helped the understanding of detailed aspects of the observed structure–activity relationships (SARs) based on the analysis of the ligand–receptor interactions probed especially by two-dimensional mutagenesis experiments (experiments in which both the ligand and the receptor are simultaneously modified according to the presumed nature of the specific molecular interaction). Such experiments are expected to be of better quality than the more frequent one-dimensional mutations of the receptor, whereby the described effect on the binding might not exclusively result from a direct ligand–receptor contact, but also from long-range structural perturbations. Such studies demonstrated that antagonists of small-molecule hormone receptors bind isosterically to the endogenous ligands, whereas nonpeptide antagonists may bind rather differently than the peptide agonists. Mutation experiments in combination with molecular modeling of ligand–receptor interactions were also useful in understanding the species differences for ligand affinities and specificities (for example, NK1 antagonists in human and rat).

More prospectively, the models were used to provide a conceptual framework for combinatorial library design strategies such as the Novartis and Biofocus chemogenomics knowledge-based approaches described below and for the optimization of selectivity aspects of lead series. For instance, modeling of the $5HT_{2C}$ receptor–ligand interaction in combination with ligand-derived comparative molecular field analysis (CoMFA) was crucial for the discovery and optimization of $5HT_{2C/B}$-selective indoline urea ligands that do not target $5HT_{2A}$ receptors.[55] Rhodopsin-based models of the 7TM domain were also instrumen-

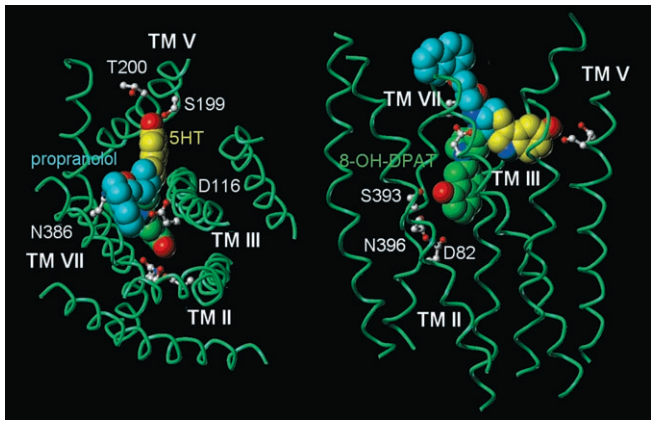


Figure 3. Three-ligand-binding-sites model for monoamine-related GPCRs illustrated by a rhodopsin-based 3D model of the $5HT_{1A}$ receptor (left: extracellular view; right: side view). We recently proposed a three-binding-sites hypothesis for the molecular recognition of ligands at monoamine GPCRs by combining: 1) analyses of the architectures of known monoamine GPCR ligands (see Figure 8); 2) analyses of molecular models of the ligand–receptor interactions; and 3) structural bioinformatics analyses of the sequence similarities of the three distinct binding regions of "one-site-filling" ligand fragments within the monoamine GPCR family. For the $5HT_{1A}$ receptor, which provided a template for the discussion of other related ligand–GPCR interactions, mutagenesis studies have mapped three spatially distinct binding regions that correspond to the binding sites of the "small, one-site-filling" ligands serotonin (5HT, yellow), propranolol (cyan), and 8-hydroxy-*N*,*N*-dipropylaminotetralin (8-OH-DPAT, green). All three binding sites are located within the highly conserved 7TM domain of the receptor and overlap at the residue Asp 3.32 (D116) in TM3, which constitutes the key anchor site for basic monoamine ligands. The three distinct binding sites are also reflected by the architectures of known high-affinity ligands, which cross-link two or three "one-site filling" fragments around a basic amino group. For further detail, see references [51, 53]. Throughout this Review, the residue positions are numbered according to van Rhee and Jacobson:[32] the first digit gives the transmembrane domain and the following number indicates the position of the residue relative to position 50, which is arbitrarily assigned to the most conserved residue in each helix.

tal to researchers at Novo Nordisk in understanding the molecular recognition of privileged structures used in generalized library design approaches, which provide a way to target orphan GPCRs in the absence of the knowledge of the endogenous or surrogate ligand.[56] Modeling the interactions of three sets of privileged motif-based ligands into their receptors (2-arylindole-based ligands in the serotonin $5HT_6$ and melanocortin-4 (MC_4) receptors, spiropiperidine-indane-based ligands in the growth hormone secretagogue (GHS) and MC_4 receptors, and 2-tetrazole-biphenyl-based ligands in the AT_1 and GHS receptors) showed the correlation of conserved patterns of residues in the ligand binding pockets of the receptors with the recognition of specific privileged fragments. These findings imply that any one particular privileged structure can target a specific subset of receptors and that motif-based searches can be used for subsetting the receptor repertoire including the orphan receptors. The models also showed that only parts of the privileged structures are accommodated within the conserved subpocket; some contacts are between substructure elements of the full privileged motif and the nonconserved part of the pocket, which suggests the possibility for the design of selective ligands based on privileged motifs. A broad spectrum of homology-modeling techniques ranging from strict, template-based methods to de novo prediction methods (such as the PREDICT method)[57] are used to build GPCR models. Although some reports suggest that rhodopsin template-based approaches can be adapted to the entire GPCR repertoire,[58] one needs to carefully investigate the underlying sequence alignments of such models, which for some helices in some subfamilies are not clear.[9] Whereas most of the time these models neglect the long intracellular loops and N- and C-terminal domains, some studies emphasized the role of the second extracellular loop E2 in ligand specificity. In the bovine rhodopsin structure the E2 loop, which is bridged through a conserved disulfide link to the residue Cys 3.25 at the top of TM3, covers parts of the central binding crevice in a lid-like manner. One of the two β strands that define the fold of the loop comes into direct contact with the retinal ligand. As the length of the loop varies significantly within the class A family, general conclusions are difficult. Recently, based on random saturation mutagenesis experiments of the C5a receptor, it was proposed that the E2 loop acts as a negative regulator of receptor activation and stabilizes the nonsignaling receptor conformation in the absence of the agonist ligand.[59] Also, the E2 loop has been implicated in ligand–ligand allosteric interactions, which were experimentally investigated by the SCAM approach.[60] For instance, in the interaction of the muscarinic M_1 receptor with the allosteric modulator gallamine, an acidic sequence segment just before the loop cysteine residue could be linked to these effects. The potential role of the E2 loop in the allosteric effects observed for amiloride on the action of antagonists of the α_{1A} and α_{2A} adrenoceptors and dopamine receptors is reported.

Recently, the potential value of GPCR models for in silico screening applications became of interest. Using a 3D model of the NK_1 receptor generated by the MOBILE (modeling binding sites including ligand information explicitly) approach in combination with 2D and 3D database searches, novel submicromolar NK_1 antagonists were discovered.[61] As shown in another study,[62] models of the dopamine D_3, muscarinic M_1, and vasopressin V_{1a} receptors based on the rhodopsin template seem to be of sufficient accuracy to be useful (20- to 40-fold enrichment relative to random screening) for protein-based virtual screening experiments. This procedure used standard docking software like DOCK, FlexX or GOLD, and searched for GPCR antagonists starting from antagonist-bound models shaped by minimizing manually docked antagonist into the binding site. The same procedure, however, was not applicable when a single agonist ligand was used for the binding site shaping step, indicating that the structural changes that can be achieved by minimization to expand the binding site are not sufficient for stimulating the conformational changes that occur in receptor activation. Instead, a multi-agonist ligand pharmacophore-based receptor refinement method needed to be used to generate useful models for the virtual screening of agonists. Corroborative findings were described for models generated with the PREDICT method and using the DOCK software in prospective virtual screening for the D_2, $5HT_{1A}$, $5HT_4$, NK_1, and CCR_3 GPCRs.[63]

Given the differences in the length of the intra- and extracellular loops, the latter are expected to contribute to ligand entry, binding, and/or modulation, especially for the peptide- and protein-binding GPCRs. Given that the currently available structures of rhodopsin in the inactive state can, at best, be a reference for an antagonist state of related class A GPCRs, there are many significant unknowns for the understanding of the structure–function relationship of GPCRs. In this respect, the modeling and indirect structural experiments of GPCRs also revealed the functional role of structural microdomains as opposed to simply considering individual residues. An important microdomain is the so-called DRY domain, which refers to a conserved sequence patch at the cytoplasmic end of TM3 in class A GPCRs and which involves residues in TM2, TM6, and TM7 as well.[64] The overall picture common to many class A GPCRs is that residue Arg 3.50 is hydrogen bonded to a carboxylate side chain at position Asp 3.49 and to one or two residues in TM6 equivalent to residues Glu 6.30 and Thr 6.34 in rhodopsin. Removal of these interactions often results in constitutive activation of the receptor, and based on this and findings of analysis of structural intermediates of the photocycle of rhodopsin, the emerging theory for receptor activation suggests a mechanism involving a separation of the TM3 and TM6 domains together with a twist in TM6 which pulls the third intracellular I3 loop into the cell, uncovering residues related to G-protein coupling. Since the DRY microdomain is not conserved in other GPCR families (exceptions are some class C GPCRs), one may conclude that the conformational changes and signaling mechanisms are not strictly conserved. Importantly, as the active conformations generated through constitutively activating mutations and specific agonist ligands seem to be nonidentical, the concept of protean ligands was defined by Kenakin to explain that each specific ligand–receptor pair defines a functional entity with distinct signaling and functional properties.[65] Clearly, this concept raises questions on the generality

of the above-mentioned virtual screening studies for GPCR agonists.

Significantly fewer modeling studies are reported for class B and class C GPCRs. For class B GPCRs, a general two-sites model has emerged for peptide binding.[7] In this mechanism, the C-terminal ligand region binds the extracellular N-terminal domain of the receptor. This interaction acts as an affinity trap, promoting the interaction of the N-terminal region of the ligand with the juxtamembrane 7TM domain of the receptor. For instance, molecular models were generated for the interaction of peptide agonists with the CFR_2 and PTH receptors, placing emphasis on α helix recognition sites.[66,67] Nonpeptide ligands bind the juxtamembrane or the N-terminal domain and, in most cases, allosterically modulate peptide–ligand binding.[7] Also noteworthy is the modeling work around the allosteric binding sites of the class C CaR[68] and $mGluR_1$ and $mGluR_5$ receptors,[69] for which site-directed mutagenesis and rhodopsin-based homology modeling showed a novel antagonist binding site within the 7TM bundles clearly separated from the agonist binding site located in the N-terminal domains of these receptors.

Oligomerization of GPCRs appears to further contribute to the complexity of the picture,[70,71] and recently a structural hypothesis was provided using molecular modeling to describe how the G protein transducin docks with the dimeric and tetrameric states of rhodopsin to reveal structural details of this critical interface in the signal-transduction process.[72] Results of biophysical studies of the leukotriene B4 BLT1 receptor reconstituted with a heterotrimeric G protein using a combination of mass spectrometry after chemical cross-linking together with neutron scattering in solution sustain this hypothesis by providing evidence for the overall assembly of a pentameric complex formed by two BLT1 units and one trimeric G protein.[73]

Ultimately, it will require high-resolution structures of multiple receptors bound to multiple ligands including agonist, inverse agonist, and antagonist coupled to G proteins and other modulators to fully understand the conformational dynamics of GPCRs. Therefore, the development of systematic approaches for X-ray and NMR spectrometric analysis of GPCR structures is currently a major scientific challenge that will require further progress in the expression, purification, and crystallization of GPCRs and their interacting proteins.[74]

3.4 Designing compound libraries targeting GPCRs

In recent years, the design of GPCR-directed compound libraries has become an intense activity of drug-discovery chemistry.[75–77] Generally, the design of deorphanization libraries can be distinguished from targeted lead-finding libraries. Given the broad chemical diversity of the hormones that are recognized by GPCRs, deorphanization libraries try to cover as many known active chemical classes as possible. The term surrogate agonist library is also appropriate given that the purpose of these libraries is to find a chemical compound that selectively activates a given orphan receptor of interest.[26,27] Typically, compounds identical or similar to previously identified GPCR agonists are included together with approved drugs and other reference compounds with known bioactivity, such as primary metabolites like the KEGG compound set, or commercially available compilations like the Tocris LOPAC, the Prestwick, or the Sial Biomol sets. In addition to HPLC fractionations of tissue extracts to identify new peptides and metabolites, protein mimetic libraries including β-turn/α-helix mimetics together with random or designed peptide libraries based on the bioinformatics analysis of putatively secreted peptides and protein hormones defined in the genome are of interest. Typically, the size of deorphanization or surrogate sets is on the order of a few thousand well-characterized compounds amenable for medium-throughput screening.

The design of lead-finding libraries follows the same molecular mimicry principles and makes best use of the substantial medicinal chemistry knowledge generated during the last decades around GPCR compounds together with more modern concepts, including lead/drug likeness and computational combinatorial library design.[78,79] Although focused library design concepts target the classical binding sites in general, design concepts of bivalent ligands and allosteric ligands are expected to become more important in the future given the anticipated progress in the understanding of the GPCR oligomerization phenomenon.[80] Divalent ligands that selectively target δ–κ opioid receptor heterodimers are a recent example.[81] The general experience with focused libraries and screening sets for GPCRs is very positive, and hit rates of 1–10% can be expected with library sizes of 500–2500 compounds, if the libraries are designed toward new members with expected conserved molecular recognition. Peptide and protein mimetics libraries including β-turn/α-helix mimetics are of central importance.[82,83] A number of important hormones such as angiotensin, bradykinin, cholecystokinin (CCK), melanin stimulating factor (MSF), and somatostatin (SST) make their key recognition through specific β-turn motifs. Others such as corticotrophin releasing factor (CRF), PTH/PTHrP, neuropeptide Y (NPY), vasoactive intestinal peptide (VIP), and growth hormone releasing factor (GHRF) interact through α-helix motifs.[7,84] Whereas the design of organic druglike α-helix mimetics is still in its infancy, the design of orally active β-turn mimetics based on organic druglike scaffolds, cyclic α peptides, or β/γ peptides has advanced to a level of routine methodology. The work of Garland and Dean[85,86] defined a set of triangular distance constraints that the substitution points of a scaffold must satisfy to mimic the specific C^{α} atoms of the peptide template. This provided a generalized framework for the design of novel β-turn mimetic scaffolds and was, in combination with database searches, successfully applied toward the design of CCK and SST antagonists.[84]

The use of privileged substructures or molecular master keys, whether target-class-specific or mimicking protein secondary-structure elements, is an accepted concept in medicinal chemistry. The privileged-structure approach emphasizes molecular scaffolds or selected substructures that are able to provide high-affinity ligands (agonist or antagonists) for diverse receptors. It originates from work carried out at Merck on the design of benzodiazepine-based CCK antagonists, from which the previously known κ opioid tifluadom was identified as a

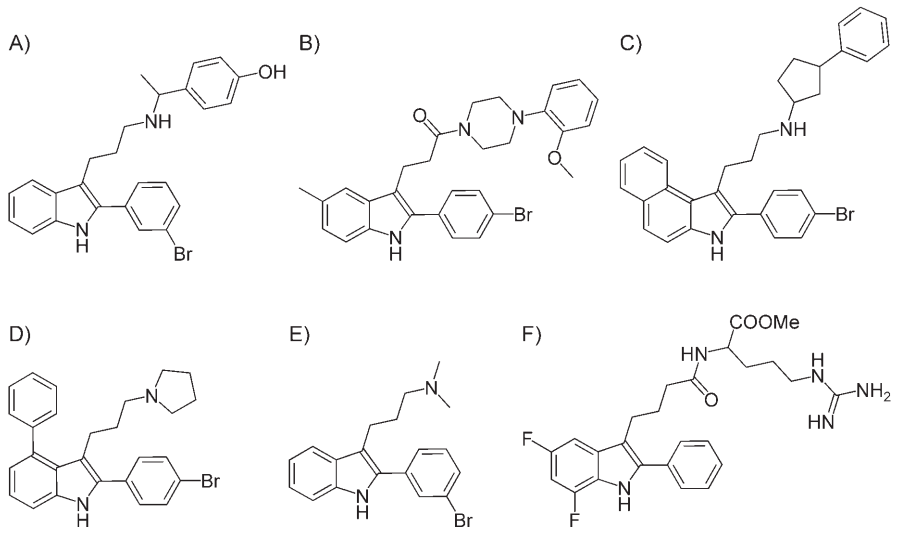


Figure 4. Examples of GPCR-active compounds based on the 2-arylindoles privileged scaffold identified from a focused combinatorial library at Merck. Screening of the library against several GPCRs led to the discovery of receptor antagonists toward A) NPY_5 ($IC_{50}=0.8$ nM); B) NK_1 ($IC_{50}=0.8$ nM); C) chemokine CCR_3 ($IC_{50}=920$ nM) and CCR_5 ($IC_{50}=1190$ nM); D) serotonin $5HT_{2A}$ ($IC_{50}=10$ nM); E) serotonin $5HT_6$ ($IC_{50}=0.7$ nM); and F) SST_4 ($K_i=0.7$ nM).[91]

lead structure.[87] A number of recent literature reviews provide impressive reference repertoires of empirically derived privileged structures, most notably the spiropiperidines, biphenyltetrazoles, benzimidazoles, and benzofurans.[88–90] The 2-arylindole scaffold illustrated in Figure 4 represents a particularly successful example and was shown at Merck to generate actives for diverse class A GPCRs.[91] In the view of the abovementioned modeling of the ligand–receptor interactions, the privileged structural classes require further analysis to allow a more directed use of such libraries for specific receptor subsets. The development of chemoinformatics methods and procedures enabling the automatic identification and extraction of privileged structures is especially needed in the context of generating knowledge from HTS data.[92] Based on the molecular framework approach developed by Bemis and Murcko,[93] we recently initiated a systematic analysis using reference compound and target information. Using the framework analysis as implemented in the Scitegic Pipeline Pilot software, we designed a data-pipelining protocol that generates frequency analysis based on the input of the various reference sets. The approach is illustrated in Figure 5 for the monoamine GPCRs.

A different type of fragment-based design method called thematic analysis was developed by researchers at Biofocus for the design of focused class A GPCR libraries.[77] This knowledge-based method is similar to a method developed at Novartis which is illustrated in Section 4.3.[53] SARs were analyzed in detail across the whole class A GPCR family, and family–activity relationships were used to develop a new classification process based on the pairing of sequence themes and ligand structural motifs. A sequence theme is a consensus collection of amino acids within the central binding cavity, and a motif is a specific structural element binding to such a particular microenvironment of the binding site. The analysis resulted in a compilation of themes and motifs, which to date are used at Biofocus to generate focused discovery libraries and to increase the lead optimization efficiency for these targets. The individual compound libraries target subsets of GPCRs, including orphans, that share a predefined combination of themes consisting of a central dominant theme and peripheral ancillary themes. The library scaffold is designed to complement the central theme and is amenable to incorporate a variety of structural motifs that address the individual sequence themes. Each library, consisting of approximately 1000 compounds, can thus be thought of as representing a number of predefined themes, which are either present or absent in any given receptor, allowing through such fingerprinting to compute a score of library appropriateness for each receptor. Thematic analysis is also used to aid lead optimization by the analysis of those themes which are or are not involved in the binding of a particular hit molecule and the exploitation of new combinations of used and unused themes to increase affinity and selectivity.

Compared with the fragment-based approaches, several research groups have developed knowledge-based library design strategies which are, in principle, based on Sir James Black's frequently quoted statement that "the most fruitful basis for the discovery of a new drug is to start with an old drug". The associated selective optimization of side activities (SOSA) approach is another very successful medicinal chemistry concept in which the atypical neuroleptics acting on a couple of GPCRs simultaneously provide a relevant illustration of the rationale.[94] The related computer-assisted drug design (CADD) methods make use of selected reference compound sets and molecular descriptors together with advanced chemoinformatics methods to compare and rank the similarity of designed candidate molecules.[95,96] Homology-based similarity searching was developed at Novartis as a chemoinformatics similarity searching method that is able to identify not only ligands binding to the same target as the reference ligand(s), but also potential ligands of other homologous targets for which no ligands are yet known.[97] The method is based on the Similog descriptor, which describes molecules as counts of pharmacophore triplets formed by the individual non-hydrogen atoms and uses a centroid of the reference compounds to describe the distance to the candidate molecule. In a retrospective analysis, the method was shown to be highly effective for monoamine GPCR and became an essential tool for the compilation of focused screening sets.

Related to the chemoinformatics similarity searching methods are machine learning methods such as artificial neural net-

	5HT1A	5HT1B	5HT1D	5HT1F	5HT2A	5HT2B	5HT2C	5HT4	A1A	A2A	B1	B3	D1	D2	D3	D4	M1	M3	H2	Frequency
PS1	0	0	0	0	2	0	0	0	0	4	0	0	1	11	0	0	0	0	0	4
PS2	0	0	0	0	0	1	1	0	0	0	0	0	0	2	0	14	0	0	0	4
PS3	6	0	0	0	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	4
PS4	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	4
PS5	2	0	0	0	0	0	0	0	1	0	0	0	0	2	1	0	0	0	0	4
PS6	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	3	0	0	0	4
PS7	2	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	4

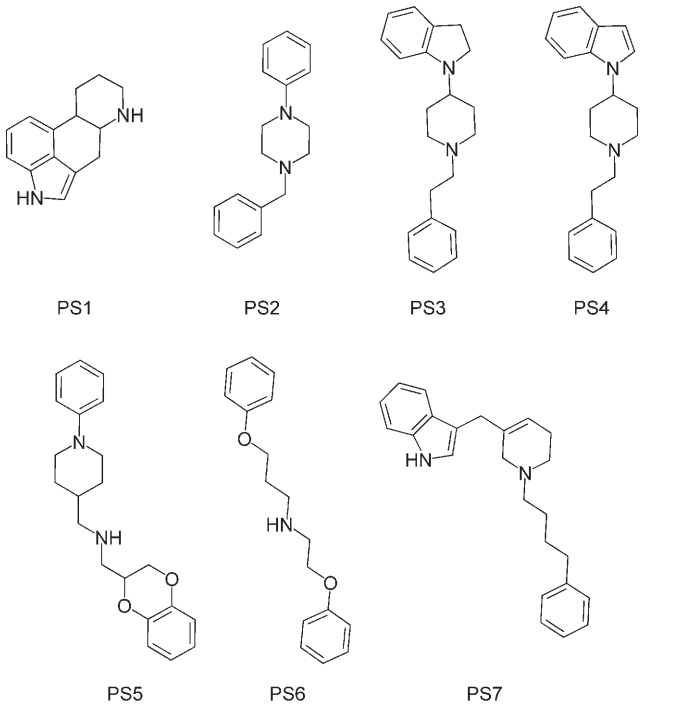


Figure 5. Analysis of privileged-scaffold–target matrix of monoamine GPCR ligands. For each GPCR ligand assigned in the MDL drug data report (MDDR) database to a specific monoamine GPCR subtype, the Bemis–Murcko frameworks were generated. The lists of frameworks were then combined, and duplicates were eliminated. The comprehensive list of unique frameworks define the row vector of the matrix, and the GPCR subtypes were arranged to the column vector. The matrix elements were assigned by the number of compounds reported including a given framework for a given subtype. In addition, for each framework the total number of monoamine GPCR subtypes addressed were added and summarized in the frequency column; the rows were then sorted by decreasing frequency. The structures of the seven most represented frameworks together with the addressed monoamine GPCR subtypes are shown.

works, Kohonen self-organizing maps, and support vector machines (SVMs), which try to align the chemical and biological spaces based on mapping procedures.[98] The goal in this case is to identify which parts (islands) of the chemical-property space correspond to specific target-family or therapeutic activities, and vice-versa. A number of research groups have applied such methods for the design of broad GPCR-focused libraries and more recently, to specifically distinguish family subgroups class A, B, and C GPCR ligands, or to identify specific GPCR ligands for the adenosine A2A, cannabinoid, CRF, and endothelin GPCRs.[99] De novo design methods using ligand-based pharmacophore models and abstract feature tree representations of GPCR ligands are reported, in which virtually generated molecules are evaluated and proposed for synthesis.[100,101]

3.5 The contribution of molecular informatics to GPCR chemical biology

Given the rapidly growing quantity of molecular data and information related to GPCRs, the need for molecular information systems that integrate bioinformatics and chemoinformatics systems was recently recognized.[11,102] The cross-linking of the chemical and biological GPCR knowledge spaces through classification and annotation schemes is an essential element of chemogenomics knowledge-based ligand-design strategies, which are based on the fact that similar ligands bind to similar targets. The systems allow the compilation of relevant reference sets for chemoinformatics-based similarity searches and for library design of target-class-focused collections. Similarly, the ligand similarity principle can be used in the reverse direction to infer putative molecular targets of compounds of interest. Most of the systematically generated information on GPCRs is publicly accessible today through the internet, and a selection of relevant information sites is summarized in Table 3. In addition, a growing number of chemogenomics knowledge-based companies such as Aureus Pharma, Inpharmatica, GVKBio, Evolvus, and Jubilant Biosys are developing molecular information systems which integrate GPCR data from patents and selected literature together with chemical and biological search engines in a comprehensive manner.

Molecular information systems like the Cerep Bioprint Matrix or Iconix DrugMatrix, which summarize the analysis of validated IC_{50} profiling data of drug and development compounds on a panel of GPCR and other targets together with ADMET (absorption, distribution, metabolism, excretion, and toxicity) data, are becoming important for lead prioritization and the design of safety pharmacology studies. Today, such data is used up front to identify the potential side effects in clinical investigations using both in vitro and in silico testing.[103–105]

Given the fast-growing complexity of the knowledge around GPCRs and their interacting effector and regulator proteins which is opening many new potential mechanisms for interaction with small drug molecules, the design of the data models of the molecular information systems will need to evolve further to enable integration and mining of knowledge within a broader systems biology and chemical genetic network concept space.

Bioinformatics analyses provide an essential contribution to GPCR chemical biology. The investigation of sequence similarities through phylogenetic, diagnostic fingerprint, or HMM analyses are a commonly used strategy to classify new orphan members and to facilitate the identification of the endogenous ligands.[106,107] Phylogenetic analyses predicted, for instance, that sphigosine-1-phosphate (S1P), the endogenous ligand of the EDG_1 GPCR, is also the ligand of the EDG_3, EDG_5, EDG_6, and EDG_8 GPCRs. Moreover, the ligand and the pharmacology of the human histamine H_4 GPCR was predicted through phylogeny, noting that it shares only 26% identity with the histamine H_1 receptor. Conversely, examples are known in which sequence homology can be misleading. For example, a receptor originally known as $P2Y_7$ (BLT1) was thought to be a nucleotide receptor based on its similarity to P_2Y purinoceptors, but it was shown to be activated by an unrelated ligand, leukotriene B_4.

Different types of bioinformatics analyses focus on specific sequence motifs and signatures which may lead to different conclusions than the analysis of the overall sequence identity. For instance, two orphan receptors, GPR61 and GPR62, were reported to share 30% overall sequence identity with the human $5HT_6$ receptor and were thus classified as monoamine-like receptors. Strikingly, both of them show mutations of the D3.32 residue and should therefore belong to a different subfamily.

Understanding the principles of molecular recognition in combination with residue- and motif-based 1D and 3D bioinformatics data mining are becoming essential elements for successful chemogenomics knowledge-based ligand-design strategies. Noteworthy in this perspective is the recent work done at Pfizer and Biofocus, where, based on the analysis of sequence data, mutation data, and physicochemical properties of the ligands, approaches were outlined to discover sequence patterns characteristic of specific ligand classes.[77,108] The potential of such computational methods was recently illustrated for the identification of ligands of the prostaglandin D2 receptor CRTH2 (chemoattractant receptor-homologous molecule expressed on T helper type 2). Using a computational strategy which emphasizes the classification of GPCRs with respect to physicochemical features of selected amino acid residues of the central binding cavity, researchers at 7TM Pharma showed that the angiotensin AT1 and AT2 references can be used to identify high-affinity ligands for the CRTH2 receptor. In the conventional phylogenetic analysis, the AT1 and AT2 receptors are not identified as close neighbors according to the conventional evolutionary relationship models.[109]

Other signature motifs direct the signaling interactions of the receptors with effector and regulator proteins. The identification of a conserved motif within second intracellular loops I2 and I3 of the somatostatin receptor subtypes SST_1, SST_3, and SST_4, the dopamine D_2, and the α_{2B}-adrenoceptors, which confers inhibitory coupling to the NHE_1 ion exchanger, is given as a recent example.[110]

4. Applications and Practical Examples

4.1 Biological expression of GPCRs

The analysis of the tissue distribution of the receptors provides valuable information related to the potential physiological function and therapeutic indication of a given GPCR, and is an essential part of the pharmacological target validation in the drug-discovery process. Validation is based on evidence that the target gene is expressed in cells relevant to the pathophysiological mechanism of the disease indication. This information is combined with epidemiological evidence that target gene expression is associated with the appearance and progression of the disease indication. Furthermore, evidence that target gene activity is necessary for a defined phenotypic response relevant to disease indication is tested by the inhibition of its expression or function, or by overexpression.

For instance, the undecapeptide Substance P is a neurotransmitter that mediates diverse biological responses in the nervous and immune systems mainly through the NK_1 GPCR. The specific response of the hormone depends on the location of the NK_1 receptor, and pain, neurogenic inflammation, asthma, and emesis are currently discussed as potential therapeutic indications for NK_1 antagonists. The knowledge of the tissue distribution is thus essential to predict potential main and side activities.

To this aim, specifically designed functional genomics experiments using oligonucleotide GPCR chips or reverse transcriptase-polymer chain reaction (RT-PCR) technologies in combination with immunochemistry approaches allow the identification of gene expression profiles across a wide variety of healthy versus diseased human and animal tissues.[36,43] The GPCR expression matrix generated by Vassilatis et al.,[36] represented in Figure 6, shows the expression of 100 randomly selected *endo*-GPCRs in peripheral and neural mouse tissues, demonstrating that most GPCRs are expressed in multiple tissues and that individual tissues express multiple receptors. Strikingly, over 90% of the analyzed GPCRs are expressed in the brain. The profiles of most GPCRs are unique, yielding thousands of tissue-specific receptor combinations for the potential modulation of physiological processes and design of therapeutics. Given that each tissue appears to have a unique combination of GPCRs indicates that secondary messenger pathways are used in different contexts to allow differentiation of cellular responses to hormone action. Expression profiling also contributes to the understanding of the functional significance of receptor subtypes, which in different tissues couple the same hormone to different G proteins and effector systems, and which might also show differences in their constitutive activity or regulatory aspects such as desensitization kinetics.

4.2 Advances in HTS of GPCRs

Since the birth of modern HTS in the mid-1980s, drug discovery experienced an explosion in novel assay methodologies and technologies. While around 10000 compounds were tested every year in few assays in the mid-1980s, these numbers rapidly increased in the major pharmaceutical companies over the past 20 years to reach 1–2 million compounds tested within 50–100 assays. The major challenge is to develop and implement simple assay methods to expedite HTS while maintaining high quality and generating relevant information at low cost. GPCRs are targets for which these criteria apply well, as their mode of activation by ligands offers many opportunities for assay design and miniaturization. As illustrated in Figure 2, in addition to basic ligand binding assays, the signaling cascade subsequent to GPCR activation opens versatile opportunities to develop HTS assays based on G protein activation, determination of secondary messengers, or nuclear activation. A variety of biophysical readout techniques and assay formats are routinely used today and have advantages and limitations as summarized in Table 4.

Every assay is selected based on a set of criteria including infrastructure, instrumentation, throughput requirement, and the type of information requested, among others. For cell-based

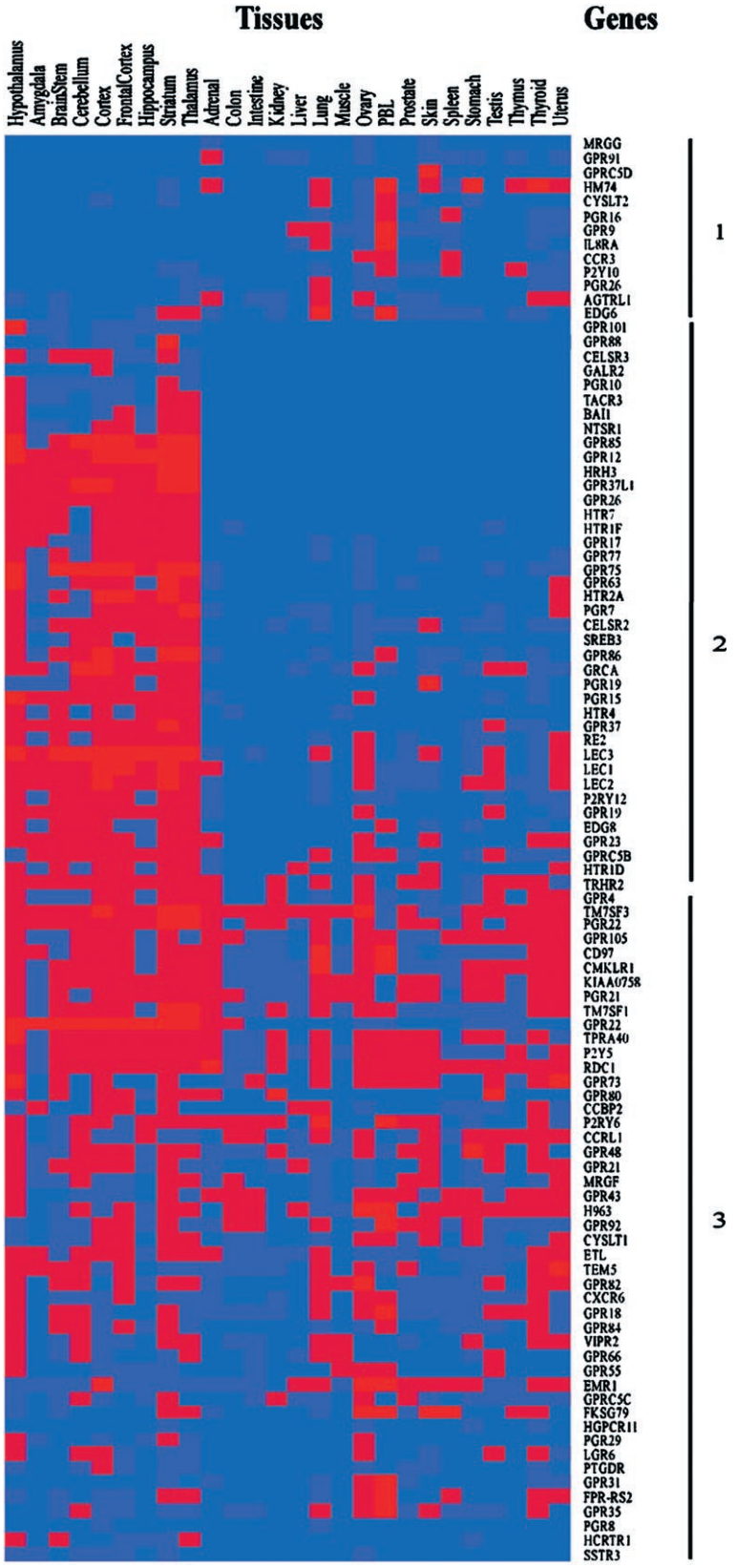


Figure 6. Cluster analysis of the expression of 100 randomly selected mouse *endo*-GPCR genes in 17 peripheral tissues and nine different brain regions. The genes were analyzed individually by RT-PCR as shown, and the intensity of the observed bands was determined by scanning. Each gene is represented by a single row of colored boxes with four different expression levels: no expression, blue; low expression, purple; moderate expression, dark red; strong expression, pure red. Three groups of *endo*-GPCRs with broadly related profiles were observed. The first group (1) contained genes expressed primarily in peripheral tissues. Seven of these genes were expressed exclusively in peripheral tissues and not in the brain. The second group (2) contained genes expressed primarily in brain. Of these 41 genes, 14 were solely expressed in brain and not in peripheral tissues. In the third group (3), the genes were broadly expressed in the brain and throughout the periphery. Figure reproduced with permission from [36].

Table 4. Commonly used assays for GPCR HTS.

Molecular Principles	Coupling	Assay Type	Plate Format	Comments/Limitations
ligand binding	G_i, G_s, G_q	radioligand filtration assay	96	safety, cost
		SPA[a] radioligand binding	384	cost
		FP[b]	384, 1536	ligand labeling
G-protein activation	G_i, G_s	GTPγ^{35}S filtration assay	96	safety, cost
		GTPγ^{35}S SPA	384	cost
2nd messenger	G_i, G_s	cAMP determination based on fluorescence approaches: FP, FI[c], HTRF/LANCE[d]	1536	FP: sensitivity
	G_i, G_q	IP_3 determination on binding and chromatographic approaches	96, 384	low throughput, mainly G_q
	G_i, G_s, G_q	Ca^{+2} determination using specific fluorescence reader technology and indicator dyes (FLIPR/Fluo-4, FDSS6000/Fluo-4, Fura-2) or proteins (Aequorin)	384	mainly with G_q and G_s with CNG2[e] channels
nuclear activation	G_i, G_s, G_q	reporter gene assays activated through CRE[f] and SRE[g] response elements and SEAP[h], luciferase, and β-lactamase readouts	384, 1536	may lead to signal variation based on cell quality; long incubation.

[a] SPA = scintillation proximity assay: a homogeneous assay that detects radioisotopes in close proximity to a solid scintillant. [b] FP = fluorescence polarization. [c] FI = fluorescence intensity. [d] HTRF/LANCE = homogeneous time-resolved fluorescence/lanthanide cryptate excitation. [e] CNG2 = cAMP-gated ion channel 2. [f] CRE = cAMP response element. [g] SRE = serum response element. [h] SEAP = secreted alkaline phosphatase. For further detail, see references [111–113].

GPCR assays, the question arises to measure affinity or efficacy;[114] both are two fundamental and distinct characteristics of the compound–receptor pair pharmacology.[115] Functional cellular assays can yield more information than can ligand-binding assays in searches for allosteric modulators that act at receptor sites other than the binding site of the endogenous agonist. The same is true if multiple measurements are required in the same well to provide additional activity and selectivity information. For example, Sabroe et al. showed in a single HTS run that dual CCR_1 and CCR_3 blockers are able to abrogate chemokine-induced cell chemotaxis and other functional parameters such as eosinophil shape changes and calcium mobilization.[116]

FLIPR duplex calcium mobilization assays were developed at Novartis to identify blockers of the chemokine $CXCR_4$ receptor. Screening compounds are tested in the same well against the $CXCR_4$ receptor and subsequently against the muscarinic M_5 receptor expressed in the same CEM-T cells. This duplex readout provides clues on compound selectivity in a cost-effective fashion already in primary screening. The approach is made possible by the noninvasive nature of FLIPR calcium assays and enables the prioritization of compounds acting at the receptor level and the exclusion of compounds that interfere with cellular components common to the two GPCRs.

Furthermore, GPCR triplex assays are routinely used at Novartis and rely on three successive readouts obtained from the same well. As shown in Figure 7, the triplex $GABA_B$ heterodimeric receptor calcium assay enables the detection of agonist, modulation, or antagonist properties of screening compounds in a single run. The presence of an agonist is not only revealed by its own activity (Figure 7A), but also through receptor desensitization during the antagonist assay phase. Modulators are detected by using a low concentration of agonist in the second phase (Figure 7B) and may be devoid of agonist properties. Antagonists clearly appear in the third phase following an injection of a higher GABA concentration and are characterized by a lack of intrinsic activity in the first phase (Figure 7C).

Multiplex assays do not achieve the compound throughputs possible with single-measurement assays. However, they produce much richer information already in primary screening, which is invaluable for compound categorization and prioritization by the medicinal chemists. A further advantage of such assays is the possibility to exploit fluorescence kinetic traces to exclude nonspecific compound interference with the readout of the affecting cells. The lower compound throughput per unit time can be largely compensated by a careful assay design and by using assay automation to ensure overnight operations.

Although the information from cell-based, mainly heterologous systems is quite valuable, caution is necessary for the interpretation of its physiological relevance in vivo. For instance, in stably transfected CHO cells, cevimeline (AF102B) behaves as a classical M_1 antagonist as measured by adenylate cyclase activity, fully blocking the activation by carbachol. However, measurement of IP_3 activation via PLC_β or PLA_2 in the same cell line shows the compound to behave as a partial agonist. More amazingly, as monitored through intracellular Ca^{+2} mobilization with confocal microscopy, cevimeline behaves as a superagonist, with a stronger response than carbachol.[118] Therefore, advanced HTS data analysis plays an important role for decision support in drug discovery.[119]

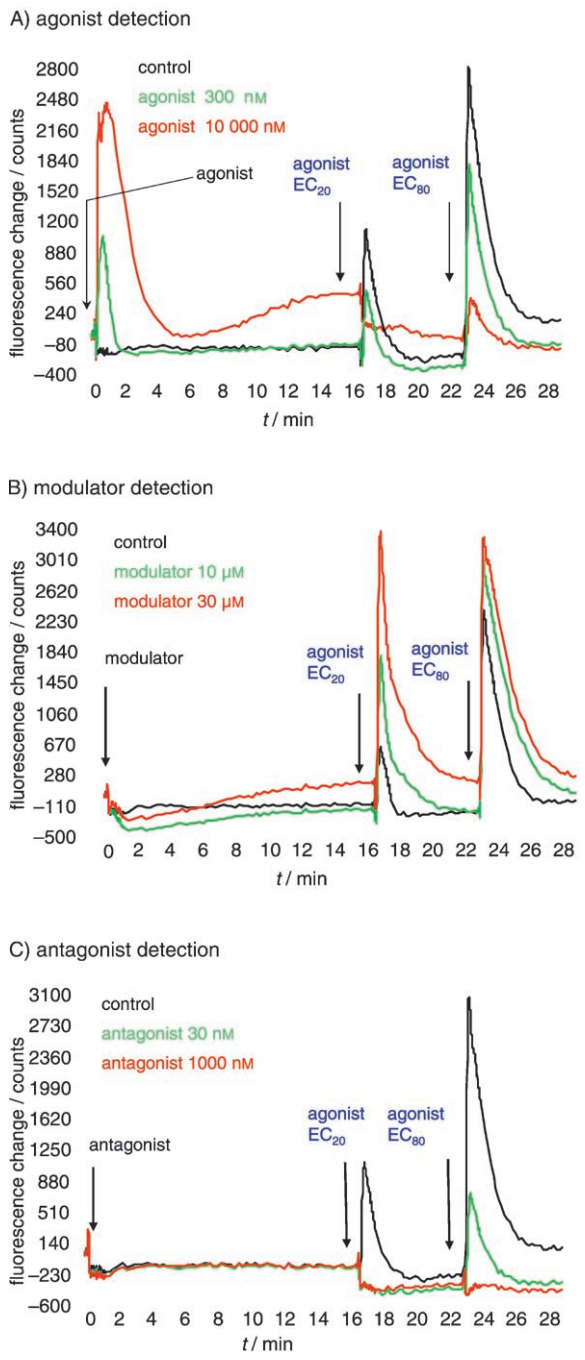


Figure 7. FLIPR calcium traces from a $GABA_B$ receptor triplex assay in a 384-well format. Experiments are performed with Chinese hamster ovarian (CHOK1) cells stably expressing the $GABA_{BR1}$ and $GABA_{BR2}$ receptor subunits. The cells are loaded with the calcium-sensitive dye Fluo-4. Three successive injections are performed during the course of the experiment. The first injection is with screening compound in general at 10 μM, followed by two injections of GABA at concentrations corresponding to its EC_{20} (0.15 μM) and EC_{80} (10 μM) values. Different FLIPR traces are obtained depending on the nature of the screening compound. A) The agonist GABA; B) L-baclofen, a $GABA_B$-receptor modulator; C) CGP56999, a competitive $GABA_B$ antagonist.[117] The signals shown are expressed as non-normalized fluorescence changes.

4.3 Designing a focused combinatorial library for monoamine-related GPCRs

Based on the central chemogenomics principle that similar ligands bind to similar targets and that ligands of close homologous receptors are generally considered as putative starting points in lead-finding programs for receptors for which no specific ligands are yet known, we proposed a chemogenomics knowledge-based combinatorial library design strategy for lead finding.[53] The strategy is founded on the integration of both the deconvolution of known modular ligands of homologous receptors into their component fragments and the structural bioinformatics comparison of the binding sites for the individual ligand fragments. In essence, in the ligand space, by the analysis of both the ligand architectures and the structures of the component "one-site filling" fragments of known ligands, it should be possible, by referring to the locally most directly related and characterized receptors, to identify those component ligand fragments which are potentially best suited for the design of ligands tailored to the new target receptor based on the binding site similarities. The strategy was presented in the context of designing the tertiary amine (TAM) combinatorial library directed toward monoamine-related GPCRs for which the conserved aspartate residue D3.32 in TM3 was demonstrated by two-dimensional mutation experiments to be responsible for the recognition of the charged amino group of monoamine ligands by their GPCRs (Figure 8). Focusing on the central importance of the D3.32 residue and using the D3.32X_{16}(DE)R-(YFH) motif in TM3 as a sequence signature defining relatedness to the monoamine GPCR subfamily, we identified 50 human GPCRs through database searches, including seven orphan GPCRs (two of which are now known to correspond to pseudogenes) which constituted the originally aimed target repertoire of the library. It was later recognized that trace amine receptors, in which the D3.32 residue is conserved, as well as chemokine receptors, which lack the D3.32 residue but have a glutamate residue (E7.39) in TM7 that is responsible for the recognition of the tertiary amine chemotype, have to be considered as monoamine-like GPCRs based on molecular recognition principles. This extends the target repertoire to about 80 GPCRs—a significant portion of all class A GPCRs.[11]

Databases of site-specific ligand fragments, which should be recombined on an appropriate scaffold to yield ligands, are the keystones of such a knowledge-based system. The generation of site-specific ligands through such fragments is possible in principle through the deconvolution of the known ligands guided by SAR and by molecular similarity consideration. Given the promiscuity of some fragments (such as symmetric ligands), one has to be cautious before drawing definitive conclusions about the actual positioning of the fragments. Pragmatically, these limitations to the generation of site-specific ligand fragment databases were approached by collecting fragments into multiple pools and by designing generic combinatorial libraries of known privileged active fragments around appropriate scaffolds. The TAM library was screened in a number of GPCR campaigns, and high hit rates were especially observed for the monoamine and chemokine GPCRs. No-

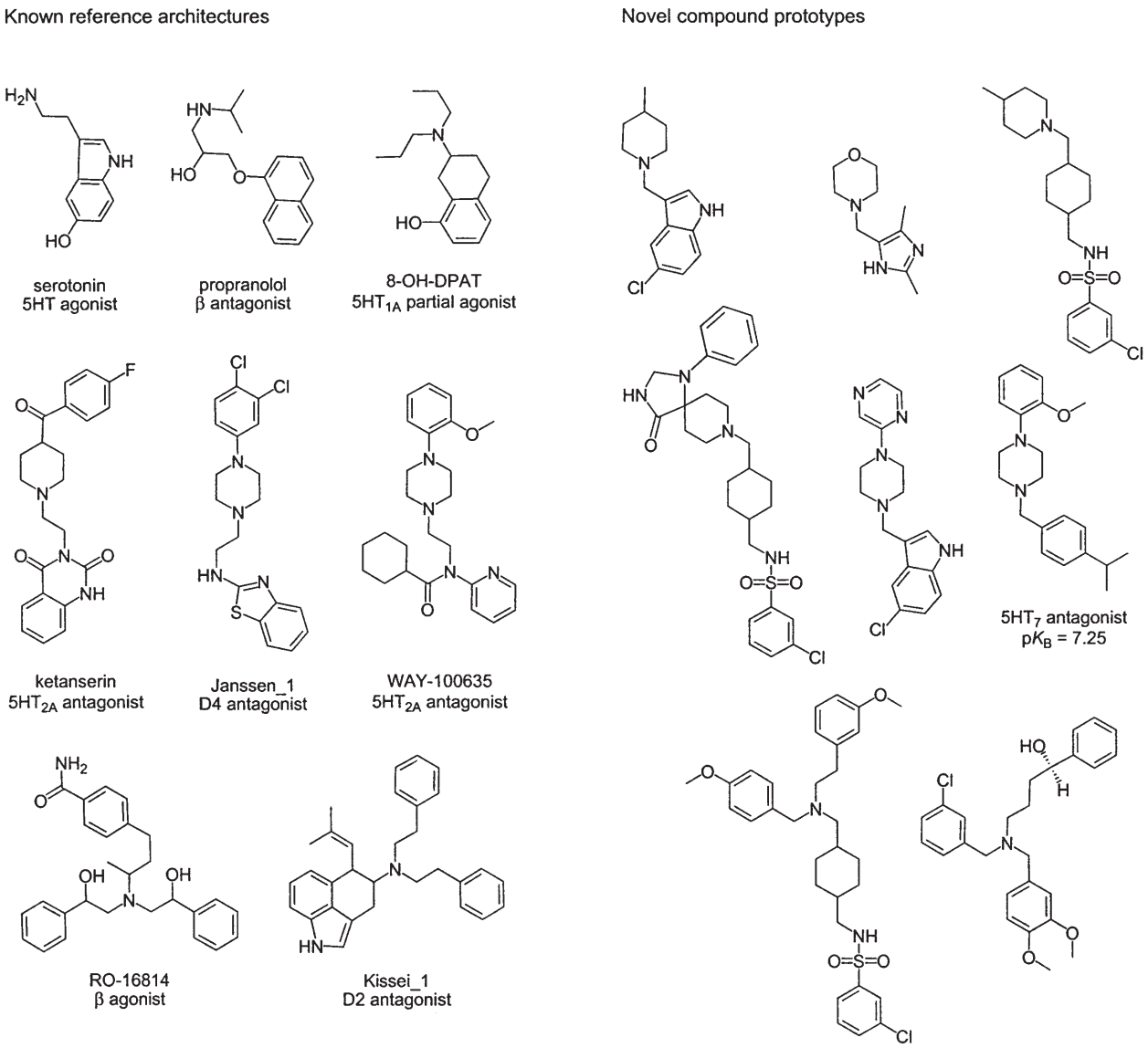


Figure 8. Prototype structures of the Novartis TAM combinatorial libraries generated through reductive amination of selected aldehydes and secondary amines. The new structures for which examples are shown on the right-hand side were designed to be similar to known monoamine GPCR ligands for which examples are shown on the left. New ligands that are the same size as the endogenous ligands are called herein "simple one-site filling" ligands. In addition to this natural architecture, there are ligands in which two or three such "simple" ligand fragments are linked around a basic positively charged group. These ligands are correspondingly termed "double" and "triple" ligands. All three architectures—"simple", "double" and "triple"—of known monoamine GPCR ligands are represented in the TAM library.

tably, the hit rates of the designed TAM library are higher than those observed for a corresponding library without specific design input. The TAM library includes many new combinations of known active fragments and privileged GPCR motifs. In addition to addressing new receptors, this should allow the discovery of interesting multi-receptor profiles of potential pharmacological interest. The search for an antagonist for the $5HT_7$ GPCR, which has the $5HT_{1A}$ receptors as neighbors in the sequence dendrogram, illustrates the successful use of the TAM library. In a search within the TAM library using the Similog method with $5HT_{1A}$ reference compounds, we were able to identify a 10% hit rate ($pK_B < 5\ \mu M$) when only a biological assay with limited throughput capacity was available. The hits were arylpiperazines, which in follow-up studies were also active toward other monoamine GPCRs.

5. Future Development

The molecular knowledge of GPCRs as information-processing units continues to progress at an impressive pace.[4,120] Besides the many efforts and opportunities on orphan receptors, GPCR research focuses on the deeper characterization of GIP networks and receptosomes.

Key questions focus on the physiological and therapeutic relevance of receptor homo- and heterodimerization.[71,73] GPCRs were initially believed to be monomeric entities, but ac-

cumulating evidence from techniques such as immunoblotting and co-immunoprecipitation combined with FRET and BRET experiments in living cells now supports the presence of GPCRs in multimeric forms. The existence of homodimers is established for many class A GPCRs (for example, dopamine D_2 and D_3, β_2-adrenoceptor, muscarinic M_1 and M_2 receptors, NK_1, opiate, and SST_5 receptors) and class C GPCRs (such as mGluR and CaR, which form covalent dimers through a cysteine disulfide bridge that links the N-terminal domains of the two receptors). Proposed roles for heterodimerization include diversifying the pharmacological response, providing a further mechanism for the fine-tuning of hormone signaling and G protein specificity, and the regulation of the receptor ontogenesis and internalization. Differences in the pharmacological properties of heterodimer GPCRs were observed for the δ/κ-opiate receptors, dopamine/somatostatin receptors, and $GABA_{BR1}$/$GABA_{BR2}$ receptors. The $GABA_{BR1}$/$GABA_{BR2}$ heterodimer is particularly illustrative.[121] It is known that $GABA_{BR1}$ is not trafficked effectively to the cell surface in the absence of $GABA_{BR2}$ expression. In addition, $GABA_{BR1}$ binds the agonist ligand, but is not coupled to G proteins, whereas $GABA_{BR2}$ activates G protein signaling, but does not bind the ligand. It was recognized that new compound-screening strategies allowing the detection of ligand binding or function only by a heterodimer pair in the presence of the corresponding homodimers are required to allow rapid and effective identification of ligands with these characteristics. Only with such ligands in hand, will it be possible to tease out the physiological relevance of GPCR heterodimerization.[71] The opioid agonist ligand 6′-guanidinonaltrindole (6′-GNTI) is an initial example of such a ligand. 6′-GNTI has the unique property of selectively activating only δ/κ-opioid receptor heterodimers but not homodimers.[122] Importantly, 6′-GNTI is an analgesic, thereby demonstrating that opioid receptor heterodimers are indeed functionally relevant in vivo. However, 6′-GNTI induces analgesia only when it is administered in the spinal cord but not in the brain, which suggests that the organization of heterodimers is tissue-specific. Other studies are indirect and may reflect cross-talk between the signaling pathways at a level downstream of receptor activation. The ability of β blockers to interfere with angiotensin AT1-mediated signaling, and the ability of the AT1 receptor blocker valsartan to decrease catecholamine-induced elevation in heart rate may indicate functional angiotensin AT1–adrenoceptor interactions in vivo.

The discovery that some GPCRs appear to function in preformed and dynamic complexes with other signal transduction and scaffolding proteins opens many interesting possibilities for drug discovery. For instance, targeting the post-synaptic density (PSD-95) and Homer scaffolding proteins might result in a new way to modulate receptor activity.[123] PSD-95 is known to function in synaptic neurotransmission and plasticity by enhancing or depressing the synaptic strength depending on the frequency of neuronal firing. The protein is a multi-adapter protein which binds specific GPCRs (such as $5HT_{2A}$ and $5HT_{2C}$) and ion channels (such as NMDA) through its PDZ domains. It enables, together with other protein–protein interactions, the spatial organization of complex micro-architectures in conjunction with the cytoskeleton. Similarly, the Homer proteins, which play a role in glutamatergic synaptic transmission, are composed of an N-terminal EVH1 (enabled VASP homology type 1) domain that interacts with GPCRs (such as $mGluR_1$ and $mGluR_5$), ion channels (such as IP_3 or ryanodine Ca^{+2} receptor channels, $TRPC_1$, and $TRPC_2$), and other proteins. Homer proteins also have a C-terminal coiled-coil domain that enables dimerization and complex formation. It remains to be seen just how general or specific these intracellular GPCR modulator mechanisms are. In addition, small molecular compounds that are able to disrupt or reinforce these interactions are needed to further understand their physiological importance.

A new trend is also the therapeutic evaluation of monoclonal antibodies against GPCRs. Although small-molecule drugs seem to be the preferred agents, recent successes in targeting the CCR_5 receptors against HIV entry, or the thyroid-stimulating hormone (TSH) receptor in Grave's disease show that this route is also feasible.

More generally, in the age of genomic medicine the pharmacogenetics of GPCRs is becoming increasingly important and plays a role especially in the target validation of new GPCRs and the clinical validation of drugs,[124] as was recently exemplified for adrenoceptors.[125] The study of allelic variations based on single-nucleotide polymorphism (SNP) or other sequence polymorphism data allows the identification of the major allele of the target gene needed for the development of screening and profiling assays. Alternative splicing (e.g., 20 isoforms are reported for the human histamine H_3 receptor), RNA editing (e.g., seven major isoforms that differ in their second intracellular loop are predicted for the $5HT_{2C}$ receptor), and coupling to specific G proteins have all been selected by evolution to modulate the activity of GPCRs, providing multiple regulatory switches to fine-tune basal cellular activities. In addition, genetic linkage studies provide evidence that a mutation in the gene is associated with susceptibility to the appearance and progression of disease indication.

Compared with the emphasis on drug-discovery applications discussed thus far, olfactory receptors play an important role for the perfume and cosmetic industry. The screening and design of new odorants is an economically interesting application. The discovery that the malaria-transmitting mosquito *Anopheles*, which is responsible for the death of more than one million people each year, possesses odorant receptors for particular components of human sweat means that different ligands could be screened for the activation or inhibition of these receptors, potentially leading to new and more effective insect traps and repellents.[126]

6. Conclusion

Chemical biology investigations of GPCRs started with very simple questions to understand how hormones such as adrenaline or glucagon elicit a signal at the intracellular level and how this signaling translates into a physiological response. During the last 25 years of molecular GPCR research, the understanding of the machinery was elucidated in great detail for a few model GPCRs and revealed a fascinating beauty which turned out to be far more complex than initially expected.[4]

During the next several years, our detailed knowledge about many newly deorphanized GPCRs and the organization and regulation of the GIP network that constitutes the receptosomes will certainly continue to grow. The chemical biology approaches mentioned herein will all contribute to the identification of compounds which enable the directed targeting of each of these components. From the perspective of drug discovery, it will be especially interesting to follow how signaling drugs will be discovered further downstream, or whether the GPCR ligand binding sites will remain the preferred entry point for medication. A central question will be how fast these molecular discoveries will translate into new medicines. Especially with the newly discovered and deorphanized receptors, the ultimate challenge resides in the enormous knowledge gap that exists between the new molecular discoveries and their significance for disease processes and medicine. Whereas the classical hormone GPCR targets were "top-down" validated based on pharmacology, physiology, and clinical medicine, the new hormone GPCR systems come "bottom-up": their early validation based on bioinformatics and genetics data are expected to direct clinical research, and the comprehensive understanding of their role in physiology will take time. It will be interesting to see the medical outcome of these activities after another decade of research.

Acknowledgements

Drs. K. Azzaoui, B. Faller, P. Floersheim, P. Fuerst, D. Hoyer, H. Mattes, H.-J. Roth, A. Sailer, G. Scheel, P. Schoeffter, S. Siehler and R. S. Tsai (all NIBR associates) are acknowledged for various support and discussions. We thank Drs. J. Mosbacher and K. Kaupman (also NIBR associates) for CHOK1 cells expressing the heterotrimeric $GABA_B$ receptor and the selective modulators and antagonists. Dr. R. Wolf is acknowledged for providing the frontispiece picture. B. Frisch and M. Brasey, from NIBR Knowledge Center, are acknowledged for support with IMS Knowledge Link. S. C. DiClemente, from NIBR Communications, is gratefully acknowledged for editorial assistance.

Keywords: chemical biology · expression analysis · high-throughput screening · library design · structural biology

[1] A. L. Hopkins, C. R. Groom, *Nat. Rev. Drug Discovery* **2002**, *1*, 727 – 730.
[2] T. Klabunde, G. Hessler, *ChemBioChem* **2002**, *3*, 928 – 944.
[3] T. W. Schwartz, B. Holst in *Textbook of Receptor Pharmacology* (Eds.: J. C. Foreman, T. Johansen), CRC, Washington, DC, **2003**, pp. 81 – 109.
[4] C. Ellis, *Nat. Rev. Drug Discovery* **2004**, *3*, 577 – 626.
[5] H. C. Huang, P. S. Klein, *Adv. Genome Biol.* **2004**, *5*, 234.1 – 234.7.
[6] F. Horn, E. Bettler, L. Oliveira, F. Campagne, F. E. Cohen, G. Vriend, *Nucleic Acids Res.* **2003**, *31*, 294 – 297.
[7] S. R. J. Hoare, *Drug Discovery Today* **2005**, *10*, 417 – 427.
[8] J. P. Pin, T. Galvez, L. Prezeau, *Pharmacol. Ther.* **2003**, *98*, 325 – 354.
[9] R. Fredriksson, M. C. Lagerström, L. G. Lundin, H. B. Schiöth, *Mol. Pharmacol.* **2003**, *63*, 1256 – 1272.
[10] IUPHAR Committee on Receptor Nomenclature and Drug Classification, *The IUPHAR Compendium of Receptor Characterization and Classification*, 2nd ed., **2000**, IUPHAR Media.
[11] E. Jacoby, A. Schuffenhauer, P. Acklin in *Chemogenomics in Drug Discovery—A Medicinal Chemistry Perspective* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 139 – 166.
[12] G. Wess, *Drug Discovery Today* **2002**, 4, 533 – 535.
[13] R. Lefkowitz, *Trends Pharmacol. Sci.* **2004**, *25*, 413 – 422.
[14] R. R. Neubig, M. Spedding, T. Kenakin, A. Christopoulos, *Pharmacol. Rev.* **2003**, **55**, 597 – 606.
[15] M. Rodbell, *Biosci. Rep.* **1995**, *15*, 117 – 133.
[16] A. G. Gilman, *Biosci. Rep.* **1995**, *15*, 65 – 97.
[17] E. W. Sutherland, *Science* **1972**, *177*, 401 – 408.
[18] M. J. Marinissen, J. S. Gutkind, *Trends Pharmacol. Sci.* **2001**, *22*, 368 – 376.
[19] H. Hamm, *J. Biol. Chem.* **1998**, *273*, 669 – 672.
[20] J. Black, *In Vitro Cell. Dev. Biol.* **1989**, *25*, 311 – 320.
[21] M. G. Caron, Y. Srinivasan, J. Pitha, K. Kociolek, R. J. Lefkowitz, *J. Biol. Chem.* **1979**, *254*, 2923 – 2927.
[22] R. A. Cerione, B. Strulovici, J. L. Benovic, C. D. Strader, M. G. Caron, R. J. Lefkowitz, *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 4899 – 4903.
[23] R. A. F. Dixon, B. K. Kobilka, D. J. Strader, J. L. Benovic, H. G. Dohlman, T. Frielle, M. A. Bolanowski, C. D. Bennett, E. Rands, R. E. Diehl, R. A. Mumford, E. E. Slater, I. S. Sigal, M. G. Caron, R. J. Lefkowitz, C. D. Strader, *Nature* **1986**, *321*, 75 – 79.
[24] Y. A. Ovchinnikov, *Photochem. Photobiol.* **1987**, *45*, 909 – 914.
[25] R. Henderson, P. N. Unwin, *Nature* **1975**, *257*, 28 – 32.
[26] N. Robas, M. O'Reilly, S. Katugampola, M. Fidock, *Curr. Opin. Pharmacol.* **2003**, *3*, 121 – 126.
[27] A. Wise, S. C. Jupe, S. Rees, *Annu. Rev. Pharmacol. Toxicol.* **2004**, *44*, 43 – 66.
[28] R. Seifert, K. Wenzel-Seifert, *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2002**, *366*, 381 – 416.
[29] E. Schipani, K. Kruse, H. Juppner, *Science* **1995**, *268*, 98 – 100.
[30] P. Strange, *Trends Pharmacol. Sci.* **2002**, *23*, 89 – 95.
[31] C. D. Strader, I. S. Sigal, R. B. Register, M. R. Candelore, E. Rands, R. A. Dixon, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 4384 – 4388.
[32] A. M. van Rhee, K. A. Jacobson, *Drug Dev. Res.* **1996**, *37*, 1 – 38.
[33] J. Bockaert, P. Marin, A. Dumuis, L. Fagni, *FEBS Lett.* **2003**, *546*, 65 – 72.
[34] J. Bockaert, J. P. Pin, *EMBO J.* **1999**, *18*, 1723 – 1729.
[35] J. Bockaert, A. Dumuis, L. Fagni, P. Marin, *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 649 – 657.
[36] D. K. Vassilatis, J. G. Hohmann, H. Zeng, F. Li, J. E. Ranchalis, M. T. Mortrud, A. Brown, S. S. Rodriguez, J. R. Weller, A. C. Wright, J. E. Bergmann, G. A. Gaitanaris, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4903 – 4908.
[37] L. Buck, R. Axel, *Cell* **1991**, *65*, 175 – 187.
[38] A. Terakita, *Adv. Genome Biol.* **2005**, *6*, 213.1 – 213.9.
[39] R. Fredriksson, H. B. Schiöth, *Mol. Pharmacol.* **2005**, *67*, 1414 – 1425.
[40] A. Sodhi, S. Montaner, J. E. Gutkind, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 998 – 1012.
[41] S. Offermanns, M. I. Simon, *J. Biol. Chem.* **1995**, *270*, 15 175 – 15 180.
[42] B. R. Conklin, Z. Farfel, K. D. Lustig, D. Julius, H. R. Bourne, *Nature* **1993**, *363*, 274 – 276.
[43] D. T. Chalmers, D. P. Behan, *Nat. Drug Discov. Rev.* **2002**, *1*, 599 – 607.
[44] L. M. McLatchie, N. J. Fraser, M. J. Main, A. Wise, J. Brown, N. Thompson, R. Solari, M. G. Lee, S. M. Foord, *Nature* **1998**, *393*, 333 – 339.
[45] G. Milligan, *Drug Discovery Today* **2003**, *8*, 579 – 585.
[46] C. Grånäs, B. K. Lundholt, A. Heydorn, V. Linde, H.-C. Pedersen, C. Krog-Jensen, M. M. Rosenkilde, L. Pagliaro, *Comb. Chem. High Throughput Screening* **2005**, *8*, 301 – 309.
[47] M. G. Ludwig, M. Vanek, D. Guerin, J. A. Gasser, C. E. Jones, U. Junker, H. Hofstetter, R. M. Wolf, K. Seuwen, *Nature* **2003**, *425*, 93 – 98.
[48] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, *Science* **2000**, *289*, 739 – 745.
[49] E. Pebay-Peyroula, G. Rummel, J. P. Rosenbusch, E. M. Landau, *Science* **1997**, *277*, 1676 – 1681.
[50] M. F. Hibert, S. Trumpp-Kallmeyer, J. Hofflack, *Trends Pharmacol. Sci.* **1993**, *14*, 7 – 12.
[51] E. Jacoby, J. L. Fauchère, E. Raimbaud, S. Ollivier, A. Michel, M. Spedding, *Quant. Struct.-Act. Relat.* **1999**, *18*, 561 – 572.
[52] S. Filipek, D. C. Teller, K. Palczewski, R. Stenkamp, *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 375 – 397.
[53] E. Jacoby, *Quant. Struct.-Act. Relat.* **2001**, *20*, 115 – 123.

[54] D. R. Flower, *Biochim. Biophys. Acta* **1999**, *1422*, 207 – 234.
[55] S. M. Bromidge, S. Dabbs, D. T. Davies, D. M. Duckworth, I. T. Forbes, P. Ham, G. E. Jones, F. D. King, D. V. Saunders, S. Starr, K. M. Thewlis, P. A. Wyman, F. E. Blaney, C. B. Naylor, F. Bailey, T. P. Blackburn, V. Holland, G. A. Kennett, G. J. Riley, M. D. Wood, *J. Med. Chem.* **1998**, *41*, 1598 – 1612.
[56] K. Bondensgaard, M. Ankersen, H. Thørgersen, B. S. Hansen, B. S. Wulff, R. P. Bywater, *J. Med. Chem.* **2004**, *47*, 888 – 899.
[57] O. M. Becker, S. Shacham, Y. Marantz, S. Noiman, *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 353 – 361.
[58] C. Bissantz, A. Logean, D. Rognan, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1162 – 1176.
[59] J. M. Klco, C. B. Wiegand, K. Narzinski, T. J. Baranski, *Nat. Struct. Mol. Biol.* **2005**, *12*, 320 – 326.
[60] L. Shi, J. A. Javitch, *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 437 – 467.
[61] A. Evers, G. Klebe, *J. Med. Chem.* **2004**, *47*, 5381 – 5392.
[62] C. Bissantz, P. Bernard, M. Hibert, D. Rognan, *Proteins Struct. Funct. Genet.* **2003**, *50*, 5 – 25.
[63] O. M. Becker, Y. Marantz, S. Shacham, B. Inbal, A. Heifetz, O. Kalid, S. Bar-Haim, D. Warhaviak, M. Fichman, S. Noiman, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11 304 – 11 309.
[64] J. Ballesteros, S. Kitanovic, F. Guarnieri, P. Davies, B. J. Fromme, K. Konvicka, L. Chi, R. P. Millar, J. S. Davidson, H. Weinstein, S. C. Sealfon, *J. Biol. Chem.* **1998**, *273*, 10 445 – 10 453.
[65] T. Kenakin, *Ann. N. Y. Acad. Sci.* **1997**, *812*, 116 – 125.
[66] C. R. Grace, M. H. Perrin, M. R. DiGruccio, C. L. Miller, J. E. Rivier, W. W. Vale, R. Riek, *Natl. Acad. Sci. USA* **2004**, *101*, 12 836 – 12 841.
[67] R. C. Gensure, N. Shimizu, J. Tsang, T. J. Gardella, *Mol. Endocrinol.* **2003**, *17*, 2647 – 2658.
[68] S. U. Miedlich, L. Gama, K. Seuwen, R. M. Wolf, G. E. Breitwieser, *J. Biol. Chem.* **2004**, *279*, 7254 – 7263.
[69] A. Pagano, D. Ruegg, S. Litschig, N. Stoehr, C. Stierlin, M. Heinrich, P. Floersheim, L. Prezeau, F. Carroll, J. P. Pin, A. Cambria, I. Vranesic, P. J. Flor, F. Gasparini, R. Kuhn, *J. Biol. Chem.* **2000**, *275*, 33 750 – 33 758.
[70] S. Angers, A. Salahpour, M. Bouvier, *Annu. Rev. Pharmacol. Toxicol.* **2002**, *42*, 409 – 435.
[71] G. Milligan, *Mol. Pharmacol.* **2004**, *66*, 1 – 7.
[72] S. Filipek, K. A. Krzysko, D. Fotiadis, Y. Liang, D. A. Saperstein, A. Engel, K. Palczewski, *Photochem. Photobiol. Sci.* **2004**, *3*, 628 – 638.
[73] J. L. Baneres, J. Parello, *J. Mol. Biol.* **2003**, *329*, 815 – 829.
[74] K. Lundstrom, *Trends Biotechnol.* **2005**, *23*, 103 – 108.
[75] K. H. Bleicher, L. G. Green, R. E. Martin, M. Rogers-Evans, *Curr. Opin. Chem. Biol.* **2004**, *8*, 287 – 296.
[76] P. Jimonet, R. Jäger, *Curr. Opin. Drug Discovery Dev.* **2004**, *7*, 325 – 333.
[77] R. Crossley, *Curr. Top. Med. Chem.* **2004**, *4*, 581 – 588.
[78] J. Bajorath, *Nat. Rev. Drug Discovery* **2002**, *1*, 882 – 894.
[79] M. M. Hann, T. I. Oprea, *Curr. Opin. Chem. Biol.* **2004**, *8*, 255 – 263.
[80] S. Halazy, *Expert Opin. Ther. Pat.* **1999**, *9*, 431 – 446.
[81] D. J. Daniels, A. Kulkarni, Z. Xie, R. G. Bhushan, P. S. Portoghese, *J. Med. Chem.* **2005**, *48*, 1713 – 1716.
[82] V. J. Hruby, *Nat. Rev. Drug Discovery.* **2002**, *1*, 847 – 858.
[83] M. Eguchi, M. McMillan, C. Nguyen, J. L. Teo, E. Y. Chi, W. R. Henderson, Jr., M. Kahn, *Comb. Chem. High Throughput Screening* **2003**, *6*, 611 – 621.
[84] T. R. Webb in *Chemogenomics in Drug Discovery—A Medicinal Chemistry Perspective* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 313 – 324.
[85] S. L. Garland, P. M. Dean, *J. Comput.-Aided Mol. Des.* **1999**, *13*, 469 – 483.
[86] S. L. Garland, P. M. Dean, *J. Comput.-Aided Mol. Des.* **1999**, *13*, 485 – 498.
[87] B. E. Evans, K. E. Rittle, M. G. Bock, R. M. DiPardo, R. M. Freidinger, W. L. Whitter, G. F. Lundell, D. F. Veber, P. S. Anderson, R. S. Chang, *J. Med. Chem.* **1988**, *31*, 2235 – 2246.
[88] A. A. Patchett, R. P. Nargund in *Annual Reports in Medicinal Chemistry Vol. 35* (Ed: G. L. Trainor), Academic Press, San Diego, **2000**, pp. 289 – 298.
[89] G. Müller in *Chemogenomics in Drug Discovery—A Medicinal Chemistry Perspective* (Eds.: H. Kubinyi, G. Müller), Wiley-VCH, Weinheim, **2004**, pp. 7 – 42.
[90] T. Guo, D. W. Hobbs, *Assay Drug Dev. Technol.* **2003**, *1*, 579 – 592.
[91] C. A. Willoughby, S. M. Hutchins, K. G. Rosauer, M. J. Dhar, K. T. Chapman, G. G. Chicchi, S. Sadowski, D. H. Weinberg, S. Patel, L. Malkowitz, J. Di Salvo, S. G. Pacholok, K. Cheng, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 93 – 96.
[92] R. P. Sheridan, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1037 – 1050.
[93] G. W. Bemis, M. A. Murcko, *J. Med. Chem.* **1996**, *39*, 2887 – 2893.
[94] C. G. Wermuth, *J. Med. Chem.* **2004**, *47*, 1303 – 1314.
[95] R. P. Sheridan, S. K. Kearsley, *Drug Discovery Today* **2002**, *4*, 903 – 911.
[96] J. Hert, P. Willett, D. J. Wilton, P. Acklin, K. Azzaoui, E. Jacoby, A. Schuffenhauer, *Org. Biomol. Chem.* **2004**, *2*, 3256 – 3266.
[97] A. Schuffenhauer, P. Floersheim, P. Acklin, E. Jacoby, *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 391 – 405.
[98] N. P. Savchuck, K. V. Balakin, S. E. Tkachenko, *Curr. Opin. Chem. Biol.* **2004**, *8*, 412 – 417.
[99] M. von Korff, M. Steger, *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1137 – 1147.
[100] G. Schneider, M. L. Lee, M. Stahl, P. Schneider, *J. Comput.-Aided Mol. Des.* **2000**, *14*, 487 – 494.
[101] G. Jenkins, *Curr. Drug Discovery Technol.* **2004**, *3*, 23 – 26.
[102] A. Schuffenhauer, J. Zimmermann, R. Stoop, J. J. van der Vyver, S. Lecchini, E. Jacoby, *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 947 – 955.
[103] C. M. Krejsa, D. Horvath, S. L. Rogalski, J. E. Penzotti, B. Mao, F. Barbosa, J. C. Migeon, *Curr. Opin. Drug Discovery Dev.* **2003**, *6*, 470 – 480.
[104] H. Roter, *Curr. Opin. Drug Discovery Dev.* **2005**, *8*, 309 – 315.
[105] T. Klabunde, A. Evers, *ChemBioChem* **2005**, *6*, 876 – 889.
[106] E. S. Huang, *Drug Discovery Today* **2005**, *10*, 69 – 73.
[107] A. Gaulton, T. K. Attwood, *Curr. Opin. Pharmacol.* **2003**, *3*, 114 – 120.
[108] E. S. Huang, *Protein Sci.* **2003**, *12*, 1360 – 1367.
[109] T. M. Frimurer, T. Ulven, C. E. Elling, L. O. Gerlach, E. Kostenis, T. Hogberg, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3707 – 3712.
[110] C. Y. Lin, M. G. Varma, A. Joubel, S. Madabushi, O. Lichtarge, D. L. Barber, *J. Biol. Chem.* **2003**, *278*, 15 128 – 15 135.
[111] A. Cacace, M. Banks, T. Spicer, F. Civoli, J. Watson, *Drug Discovery Today* **2003**, *8*, 785 – 792.
[112] D. Gabriel, M. Vernier, M. J. Pfeifer, B. Dasen, L. Tenaillon, R. Bouhelal, *Assay Drug Dev. Technol.* **2003**, *1*, 291 – 303.
[113] R. M. Eglen, *Comb. Chem. High Throughput Screening* **2005**, *8*, 311 – 318.
[114] C. Williams, A. Sewing, *Comb. Chem. High Throughput Screening* **2005**, *8*, 285 – 292.
[115] D. Colquhoun, *Br. J. Pharmacol.* **1998**, *125*, 924 – 947.
[116] I. Sabroe, M. J. Peck, B. J. van Keulen, A. Jorritsma, G. Simmons, P. R. Clapham, T. J. Williams, J. E. Pease, *J. Biol. Chem.* **2000**, *275*, 25 985 – 25 992.
[117] M. Marcoli, S. Scarrone, G. Maura, G. Bonanno, M. Raiteri, *J. Pharmacol. Exp. Ther.* **2000**, *293*, 42 – 47.
[118] D. Gurwitz, R. Haring, *Drug Discovery Today* **2003**, *8*, 1108 – 1109.
[119] H. P. Fischer, S. Heyse, *Curr. Opin. Drug Discovery Dev.* **2005**, *8*, 334 – 346.
[120] T. Kenakin, *Nat. Rev. Drug Discovery* **2003**, *2*, 429 – 438.
[121] A. Couve, A. R. Calver, B. Fairfax, S. J. Moss, M. N. Pangalos, *Biochem. Pharmacol.* **2004**, *68*, 1527 – 1536.
[122] M. Waldhoer, J. Fong, R. M. Jones, M. M. Lunzer, S. K. Sharma, E. Kostenis, P. S. Portoghese, J. L. Whistler, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9050 – 9055.
[123] J. Bockaert, L. Fagni, A. Dumuis, P. Marin, *Pharmacol. Ther.* **2004**, *103*, 203 – 221.
[124] W. E. Evans, H. L. McLeod, *N. Engl. J. Med.* **2003**, *348*, 538 – 549.
[125] K. M. Small, D. W. McGraw, S. B. Liggett, *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 381 – 411.
[126] E. A. Hallem, A. Nicole Fox, L. J. Zwiebel, J. R. Carlson, *Nature* **2004**, *427*, 212 – 213.

Received: May 31, 2006

