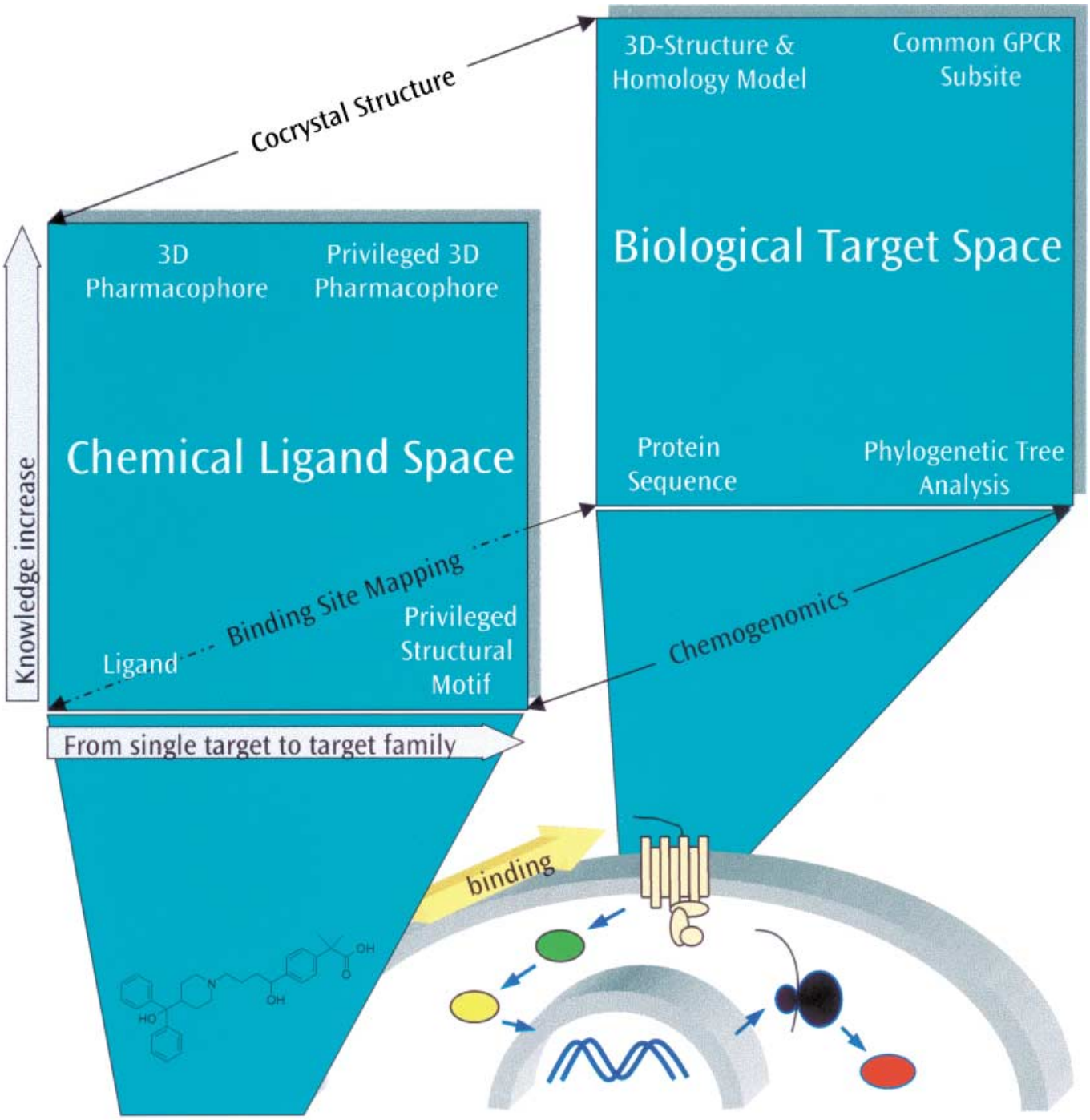


Knowledge for GPCR Lead Finding & Optimization

1439-4227/02/03/10 $ 20.00+.50/0 ChemBioChem **2002**, 3, 928 – 944

Drug Design Strategies for Targeting G-Protein-Coupled Receptors**

Thomas Klabunde* and Gerhard Hessler[a]

G-protein-coupled receptors (GPCRs) form a large protein family that plays an important role in many physiological and pathophysiological processes. Since the sequencing of the human genome has revealed several hundred new members of this receptor family, many new opportunities for developing novel therapeutics have emerged. The increasing knowledge of GPCRs (biological target space) and their ligands (chemical ligand space) enables novel drug design strategies to accelerate the finding and optimization of GPCR leads: The crystal structure of rhodopsin provides the first three-dimensional GPCR information, which now supports homology modeling studies and structure-based drug design approaches within the GPCR target family. On the other hand, the classical ligand-based design approaches (for example, virtual screening, pharmacophore modeling, quantitative structure – activity relationship (QSAR)) are still powerful methods for lead finding and optimization. In addition, the cross-target analysis of GPCR ligands has revealed more and more common structural motifs and three-dimensional pharmacophores. Such GPCR privileged structural motifs have been successfully used by many pharmaceutical companies to design and synthesize combinatorial libraries, which are subsequently tested against novel GPCR targets for lead finding. In the near future structural biology and chemogenomics might allow the mapping of the ligand binding to the receptor. The linking of chemical and biological spaces will aid in generating lead-finding libraries, which are tailor-made for their respective receptor.

KEYWORDS:

drug design · focused libraries · G-protein-coupled receptors · molecular modeling · target family

1. Introduction

G-protein-coupled receptors (GPCRs) comprise a large protein superfamily sharing a conserved structure composed of seven transmembrane (TM) helices. GPCRs are located at the cell surface and are responsible for the transduction of an endogenous signal into an intracellular response (Figure 1). The natural ligands of this receptor superfamily are extremely diverse, comprising peptide and protein hormones (for example, angiotensin, bradykinin, endothelin, melanocortin), biogenic amines (such as adrenaline, dopamine, histamine, serotonin), nucleosides and nucleotides (adenosine, adenosine triphosphate (ATP), uridine triphosphate (UTP), adenosine diphosphate (ADP)), lipids and eicosanoids (for example, cannabinoids, leukotrienes, prostaglandins, thromboxanes), and others (such as glutamate, Ca^{2+} ions). Binding of these specific ligands to the extracellular or transmembrane regions causes conformational changes of the receptor that act as a switch transferring the signal to the trimeric guanine nucleotide binding regulatory proteins (G proteins), thus inhibiting or stimulating the production of intracellular secondary messengers (for example, cyclic adenosine monophosphate (cAMP), Ca^{2+} ions).

Historically, the discovery of drugs acting at GPCRs has been extremely successful. Today 50% of all recently launched drugs are targeted against GPCRs with annual worldwide sales exceeding $30 billion in 2001.[1] Among the 100 top-selling drugs 25% are targeted at members of this protein family (Table 1 and Scheme 1). The human genome project has revealed several hundred members of the GPCR family (excluding olfactory receptors),[2] of which only approximately 30 represent targets of currently marketed drugs. Whereas for an additional 210 receptors the natural ligand is known, another 160, so-called 'orphan receptors', have been identified within the human genome, for which the ligand and the (patho)physiological function is unknown.[3] Due to the excellent potential for drug discovery, the GPCR target family represents up to 30% of the portfolio of many pharmaceutical companies. This review will attempt to show how the current knowledge of the receptors and their ligands along with the understanding of receptor – ligand interactions can aid the design of novel drugs for this target family. For a broader understanding of the subject the reader is referred to other excellent reviews.[4–6]

2. GPCR Structural Information

Although GPCRs share a common membrane topology they are remarkably diverse in sequence and vary especially in size of the extracellular amino-terminal tails, cytoplasmic loops, and carboxy-terminal tails. Based on these structural differences mam-

[a] *Dr. T. Klabunde, Dr. G. Hessler*
Aventis Pharma Deutschland GmbH
DI&A LG Chemistry, Computational Chemistry
Industriepark Hoechst, Building G878
65926 Frankfurt am Main (Germany)
Fax : (+49) 69-331399
E-mail: Thomas.Klabunde@aventis.com

[**] *A list of abbreviations can be found at the end of the text.*

Thomas Klabunde studied chemistry at the University of Münster, Germany, and obtained his PhD in 1996 for biochemical and crystallographic studies on the metalloenzyme purple acid phosphatase. In 1998, after a postdoctoral fellowship with J. C. Sacchettini at the Texas A&M University with a grant from the Deutsche Forschungsgemeinschaft, he was appointed to become research assistant professor in the Albert Alkek Institute of Bioscience and Technology (IBT), Texas A&M University, Houston. Dr. Klabunde pursued his postdoctoral studies on transthyretin – drug complexes for the structure-based design of drugs against human amyloid diseases. In addition, he studied surface proteins from the tick-born spirochete Borrelia burgdorferi *as potential targets for drugs and vaccines against Lyme diseases by X-ray crystallography. In 1999 he joined the department for computational chemistry at Aventis, where the development of drug design strategies for G-protein-coupled receptors became his primary research interest. Currently Dr. Klabunde is heading a global crossfunctional team that is engaged in the design and synthesis of small-molecule libraries directed against G-protein-coupled receptors.*

Dr. Gerhard Hessler studied chemistry at the Technical University of Darmstadt (Germany), at the University of East Anglia, Norwich (UK), and at the Technical University of Munich (Germany), where he graduated in 1995. He obtained his PhD in Munich in the research group of Professor Horst Kessler where he investigated the structure of oligonucleotides and biologically active cyclic peptides with NMR spectroscopy and molecular modeling. In 1998 he joined the computational chemistry department of Central Research at Bayer AG in Leverkusen (Germany). Since 2001 he has been researching in the computational chemistry department at Aventis in Frankfurt (Germany). During his time in industry the main focus of his work has been the application of ligand- and structure-based design techniques for the development of drugs.

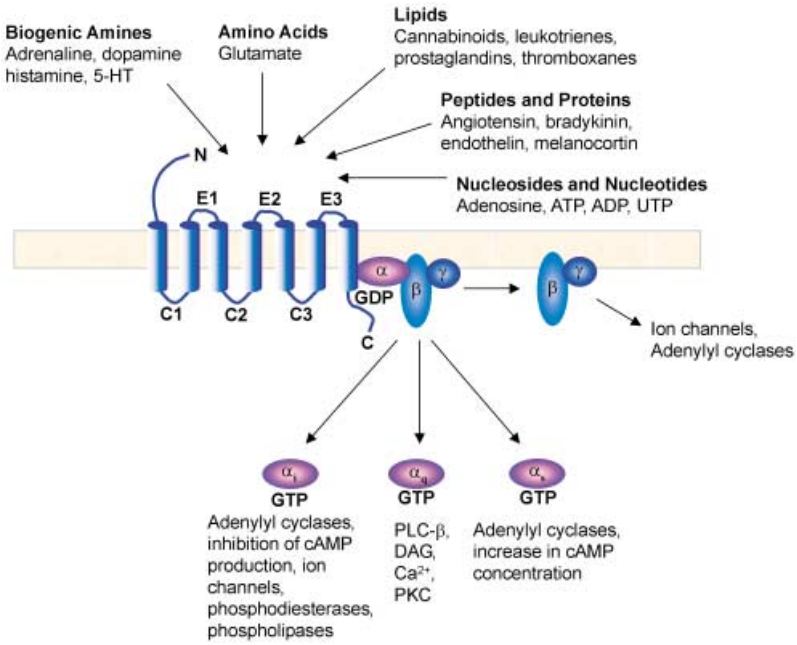


Figure 1. *Signaling pathways of GPCRs. A wide variety of ligands, including biogenic amines, lipids, peptides, proteins, nucleosides, nucleotides, and amino acids use GPCRs to stimulate cytoplasmic targets. Upon activation of the receptor the exchange of GTP for GDP bound to the Gα unit is induced within the cell, followed by the dissociation of the Gα-GTP unit from Gβγ and coupling to effector enzymes. Thus the production of secondary messengers like, for example, cAMP, cyclic guanosine monophosphate (cGMP), diacylglycerol (DAG), and Ca^{2+} ions is induced or inhibited. PKC = protein kinase C, PLC = phospholipase C.*

malian GPCRs are grouped into three major families: family A (rhodopsin-like or adrenergic-receptor-like family), family B (glucagon-receptor-like or secretin-receptor-like family), and family C (metabotropic glutamate receptors). Within each family a certain sequence pattern (so-called fingerprint) and several structural features beyond the generally shared membrane topology are conserved.[7] Family A is by far the largest, displaying short amino-terminal tails and having highly conserved amino acid residues within each transmembrane helix. Family B receptors display longer amino-terminal tails with a set of six conserved cysteine residues. Family C receptors generally have longer amino acid tails (500 – 600 residues) folded as a separate ligand binding domain.

Recently Palczewski et al. reported the 3D structure of bovine rhodopsin determined by X-ray crystallography as the first GPCR structure at atomic resolution.[8, 9] Rhodopsin is unique among the GPCRs because its ligand retinal is covalently bound by formation of a Schiff's base linked to the amino group of Lys296 in TM7. Absorption of a photon by the 11-*cis*-retinal causes its isomerization to all-*trans*-retinal. This isomerization causes a conformational change in the receptor that leads to its activation.

The structure of rhodopsin in its inactivated form at 2.8 Å resolution, illustrated as a ribbon diagram in Figure 2 A, confirms the presence of seven TM helices as implied before by cryoelectron microscopy studies.[10] The TM segments have been found to be tilted to varying degrees with respect to the putative plane of the membrane layers, thus forming the binding site of

Table 1. *Annual worldwide sales of drugs acting at GPCRs in the top 100 best selling prescription drugs in 2000.[1] Compound numbers refer to structures given in Scheme 1.*

Trademark	Generic name	Structure	Company	Disease	Target receptor	million $
Claritin	loratadine	**1**	Schering-Plough	allergies	H_1 antagonist	3011
Zyprexa	olanzapine	**2**	Eli Lilly	schizophrenia	mixed D_2/D_1/5-HT_2	2350
Cozaar	losartan	**3**	Merck & Co	hypertension	AT_1 antagonist	1715
Risperdal	risperidone	**4**	Johnson & Johnson	psychosis	mixed D_2/5-HT_{2A}	1603
Leuplin/Lupron	leuprolide	**5**	Takeda	cancer	LH-RH agonist	1394
Neurontin	gabapentin	**6**	Pfizer	neurogenic pain	GABA B agonist	1334
Allegra/Telfast	fexofenadine	**7**	Aventis	allergies	H_1 antagonist	1070
Imigran/Imitex	sumatriptan	**8**	GlaxoSmithKline	migrane	5-HT_1 agonist	1068
Serevent	salmeterol	**9**	GlaxoSmithKline	asthma	β_2 agonist	942
Plavix	clapidogrel	**10**	Bristol-Myers Squibb	stroke	$P2Y_{12}$ antagonist	903
Zantac	ranitidine	**11**	GlaxoSmithKline	ulcers	H_2 antagonist	871
Singulair	montelukast	**12**	Merck & Co	asthma	LTD4 antagonist	860
Pepcidine	famotidine	**13**	Merck & Co	ulcers	H_2 antagonist	850
Cardura	doxazosin	**14**	Pfizer	hypertension	α_1 antagonist	795
Gaster	famotidine	**13**	Vamanouchi	ulcers	H_2 antagonist	763
Zofran	ondansetron	**15**	GlaxoSmithKline	antiemetic	5-HT_3 antagonist	744
Zoladex	goserelin	**16**	AstraZeneca	cancer	LH-RH agonist	734
Diovan	valsartan	**17**	Novartis	hypertension	AT_1 antagonist	727
BuSpar	buspirone	**18**	Bristol-Myers	depression	5-HT_1 agonist	709
Zyrtec/Reactine	cetirizine	**19**	Pfizer	allergies	H_1 antagonist	699
Duragesic	fentanyl	**20**	Johnson & Johnson	pain	opioid agonist	656
Atrovent	ipratropium	**21**	Boehringer Ingelheim	asthma	anticholinergic	598
Seloken	metoprolol	**22**	AstraZeneca	hypertension	β_1 antagonist	577

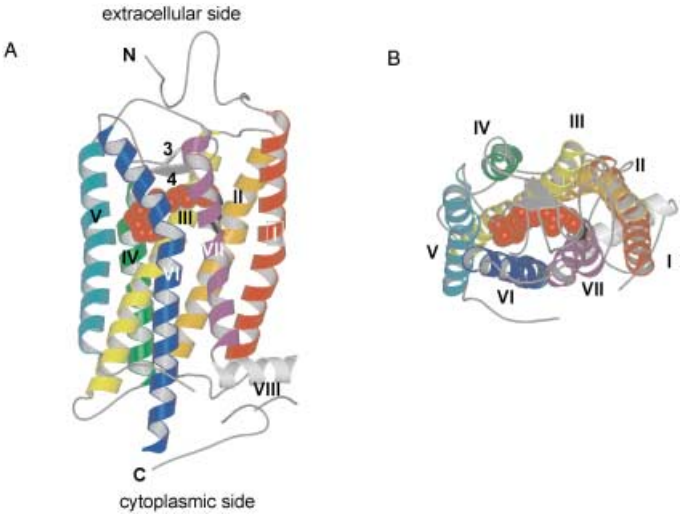


Figure 2. *Ribbon drawing of rhodopsin (A) viewed parallel to the plane of the membrane and (B) seen from the extracellular (intradiscal) side of the membrane revealing a counterclockwise arrangement of the 7-TM helices.[8] The 11-*cis*-retinal molecule is shown in red. The figures have been prepared using the programs Molscript[102] and Raster3d.[103]*

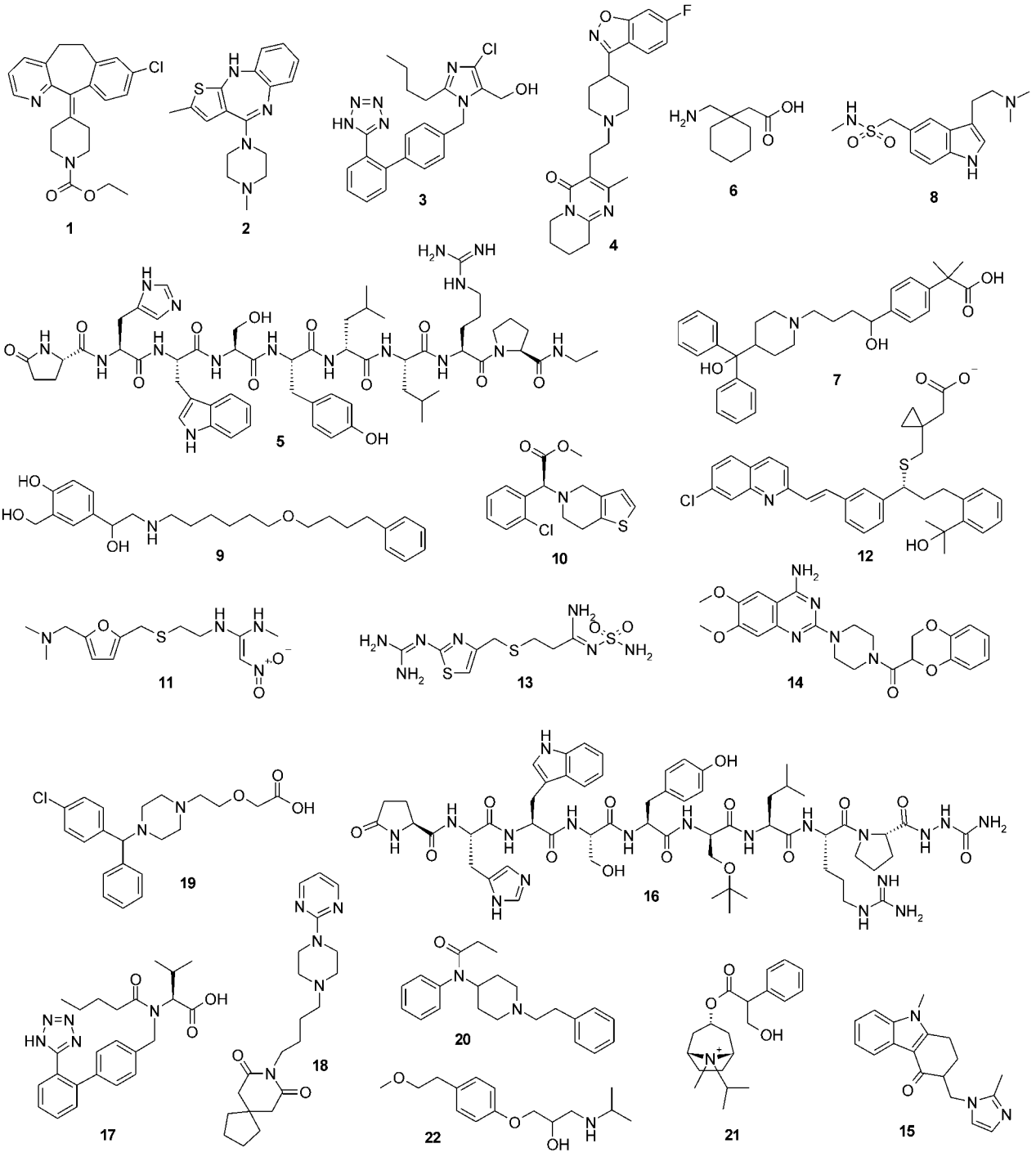
the ground-state chromophore, 11-*cis*-retinal (Figure 2B). The disposition of the seven TM helices and the location of the *cis*-retinal binding site has been analyzed and reviewed recently.[11, 12]

The TM helices are linked together sequentially by extracellular (E1, E2, E3) and cytoplasmic loops (C1, C2, C3; see Figure 1). Within the currently available rhodopsin structure the cytoplasmic surface loops C2 and C3 and the carboxy-terminal tail are not well resolved. The most striking feature is the presence of an amphipathic helix TM8 (VIII in Figure 2) lying almost parallel to the plane of the membrane, which may modulate the rhodopsin – transducin interaction. The extracellular loops of rhodopsin are folded around two twisted β hairpins, thus forming a compact and well-resolved domain structure. The innermost β strand (4 in Figure 2) is folded via the carboxy-terminal residues of the E2 loop and is attached to the TM3 helix by a disulfide bridge, a highly conserved motif among all family A receptors. Together with the TM helices this strand provides contacts with the chromophore; it runs almost parallel along the length of the retinal polyene chain. It is thus tempting to speculate that this strand provides a lid for the ligand in its TM binding pocket, a lid that might also have homologues in nearly all family A seven-transmembrane-helix (7-TM) receptors.

Further structural information is available on the extracellular amino-terminal ligand binding domain of a prototypical family C receptor, the metabotropic glutamate receptor.[13] In this context the recently determined crystal structure of the amino-terminal cysteine-rich domain of the Frizzled 8 protein needs to be mentioned, as these proteins might constitute a novel family of 7-TM receptors most closely related to the secretin family.[14]

3. Binding Site Analysis and Homology Modeling

Although bovine rhodopsin reveals only a low sequence similarity to other GPCRs, the specific arrangement of the 7-TM helices stabilized by a series of intramolecular interactions mediated by several backbone and side-chain atoms seems to be conserved among the family A receptors. Rhodopsin thus represents an improved structural template for the understanding of experimental data available for related 7-TM receptors and for generating improved molecular models of other family A receptors. Before the structure of bovine rhodopsin became



Scheme 1.

available, homology models of GPCRs were based on bacteriorhodopsin as a structural template.[15, 16] However, bacteriorhodopsin, even though it belongs to the family of 7-TM receptors, does not couple through G proteins and thus is not a member of the GPCR family. With the determination of the bovine rhodopsin structure it became evident that the positions and tilts of the seven helices are different between GPCRs and bacterial retinal-binding proteins. Despite these structural differences previous bacteriorhodopsin-based models aided the generation of hypotheses regarding the ligand binding and the signaling functions for experimental testing. Experimental findings from structure – activity, mutagenesis and affinity labeling studies have, in turn, been used to revise and refine the models. By using this combined approach molecular models of GPCRs and receptor – ligand complexes have been generated, which have been recently reviewed.[17]

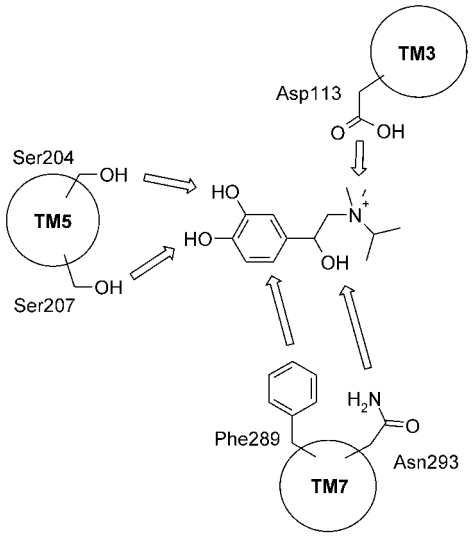
Recently, a new technology, named PREDICT, was described for modeling the 3D structure of any GPCR from its amino acid

sequence (without the use of a structural template).[18] It was demonstrated that PREDICT was capable of reproducing the known experimental structure of rhodopsin, and it is now awaiting successful applications for structure-based drug discovery towards GPCR targets.

Within this chapter we highlight some outstanding homology modeling studies, which are supported by mutagenesis and ligand binding studies. Examples are given for receptors binding biogenic amines (for example, the β_2-adrenergic receptor), purines (such as the $P2Y_1$ receptor), small hormone peptides (like the angiotensin II AT_1 receptor) and large proteins (for example, the chemokine receptor CCR2). These examples demonstrate how the understanding of receptor – ligand interactions derived from 3D structural models can be applied for the design of GPCR ligands.

3.1. β_2-Adrenergic receptor

Except for some allosteric modulators of GPCRs, general evidence has been presented that most small-molecule (ant)agonists bind wholly or partly within the transmembrane region of the receptor, the most reliably modeled part of a GPCR. Ligand binding within the transmembrane region of the receptor seems to occur mainly in a region flanked by helices 3, 5, 6, and 7. The molecular understanding of catecholamine agonist binding to the β_2-adrenergic receptor has been established by a series of pioneering site-directed mutagenesis studies and has been reviewed many times.[19, 20] Scheme 2 displays these interactions schematically with isoproterenol as a prototypical agonist: 1) Ser204 and Ser207 on the TM5 helix interact with the *meta* and *para* aromatic hydroxy groups, respectively; 2) Asp113 on helix TM3, conserved among the biogenic amine binding GPCRs and among several peptide-binding GPCRs (for example, the urotensin II, melanocyte concentrating hormone (MCH), and somatostatin (sst) receptors), forms a salt-bridge with the basic amine nitrogen; 3) Phe289 and Phe290 on the TM7 helix are probably involved in π-stacking interactions with the catechol ring; 4) Asn293 most likely interacts with the chiral β-hydroxy group common to many biogenic amines and could therefore be responsible for the stereoselectivity.[21] Longer-acting agonists, such as salmeterol (trade name Serevent, see Table 1 and Scheme 1), have a lipophilic tail in addition to features common to other agonists. The work of Green et al. suggests that residues on helix TM4 may be involved in binding of the phenalkyl tail of such ligands.[22] The interaction of this 'exosite binding' with the lipophilic tail is thought to be responsible for the enhanced duration of action of these compounds.[23]



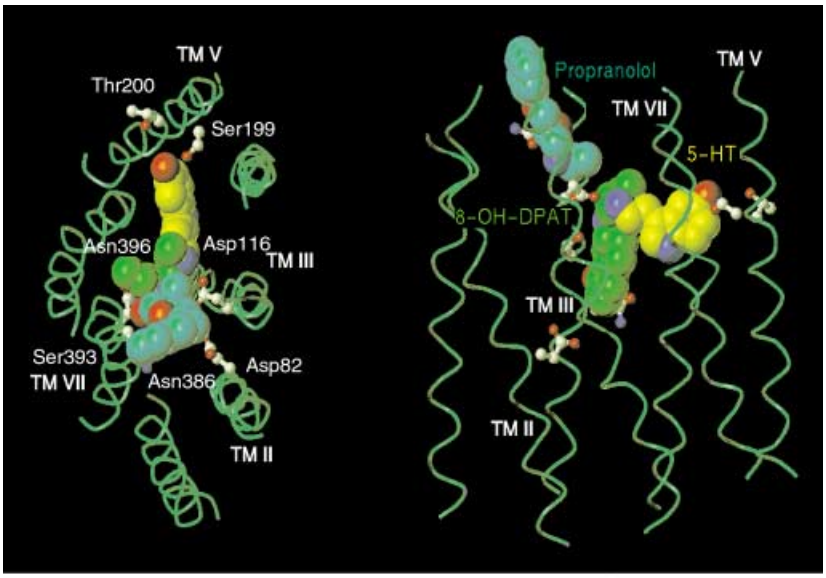
Scheme 2. *Schematic drawing of the interaction of isoproterenol with the β_2-adrenoceptor as derived from mutagenesis and modeling studies. For reasons of clarity the interaction with Phe290 is not shown. References are given within the text.*

For the biogenic amine binding GPCRs a three-binding-site hypothesis was proposed.[24] For the type-1A 5-hydroxytryptamine (seratonin) receptor ($5\text{-}HT_{1A}$ receptor), mutagenesis studies map three spatially distinct binding regions which correspond to the binding sites of the "small, one-site-filling" ligands 5-hydroxytryptamine, propranolol, and 8-hydroxy-*N*,*N*-dipropylaminotetralin (8-OH-DPAT; Figure 3). All three binding sites are located within the highly conserved 7-TM domain of the biogenic amine receptor and overlap at the aspartate residue in helix TM3, which constitutes the key anchor site for basic ligands of the biogenic amine receptors. This key anchor site corresponds to Asp113 in the β_2-adrenergic receptor and to Asp116 in the $5\text{-}HT_{1A}$ receptor. The three distinct binding sites are also reflected by the architectures of known high-affinity ligands which crosslink two or three "one-site-filling" fragments around a basic amino group.

3.2. Angiotensin II receptor type-1 (AT_1)

The type-1 angiotensin II (AngII) receptors (AT_1 receptors) have been a key target for the pharmaceutical industry. Research has led to the discovery of the non-peptide AT_1 antagonist losartan (trade name Cozaar, see Table 1 and Scheme 1) as an orally active antihypertensive. The small molecule binds to the mammalian rat AT_{1b} receptor with high affinity ($IC_{50} = 2$ nM). Interestingly, the type-a frog AT_1 receptor is unresponsive towards losartan binding, while peptide ligands such as AngII and saralasin bind with similar affinities.[25, 26] Site-directed mutagenesis studies revealed several residues within the rat receptor, that when mutated to the corresponding amino acid of the frog receptor, descreased losartan binding dramatically. On the other hand, gain-of-function mutagenesis studies, whereby 13 amino acids of the frog receptor are replaced by the corresponding rat receptor residues, confer the ability to bind losartan with the same affinity as the rat AT_{1b} receptor onto the previously unresponsive amphibian receptor.[26]

As for the biogenic amine ligands, the binding domain of the non-peptide antagonist losartan has been mapped within the 7-TM region. Analysis of single-point and combinatorial mutations of the rat AT_{1b} receptor identified the residues involved in binding of the small-molecule antagonist losartan.[25, 26] The most marked attenuation of losartan binding (greater than a factor of ten as measured in a radioligand-displacement assay with ^{125}I-labeled [Sar^1, Ile^8]AngII) was observed upon replacement of Val108 (TM3), Lys199 (TM5), and Asn295 (TM7) by alanine residues.



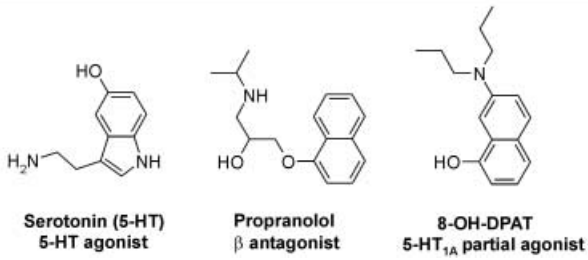


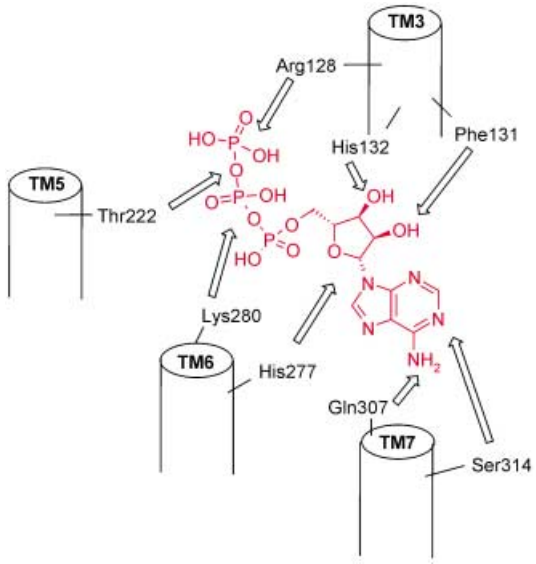
Figure 3. *Three-ligand binding sites of the 5-HT_{1A} receptor in a rhodopsin-based 3D model according to Jacoby and co-workers.[24] Left: extracellular view; right: side view with extracellular side at the top. The three ligands are serotonin or 5-HT (yellow), propranolol (cyan), and 8-OH-DPAT (green). Residues identified by mutagenesis data are indicated. The "5-HT" site is located between TM3 (TM III; with Asp116 as the key recognition site) and TM5 (TM V; providing Ser199 and Thr200 to interact with the 5-OH group of serotonin). A second site, the "propranolol" binding site, is located between TM3 and TM7 (contributing, for example, Asn386 to hydrogen bond the oxygen atoms of the oxypropanolamine fragment in β-blockers). The third binding site, the "8-OH-DPAT" binding site is also located between TM3 and TM7. Ligands addressing this site, like 8-OH-DPAT, are thought to be oriented parallel to the helices (interactions by 8-OH to Ser393 and Asn396 and by amino group to Asp116).*

Underwood and co-workers generated a 3D model of the human AT_1 receptor based on bacteriorhodopsin and were able to dock the losartan-type antagonist L-158,282 (MK-966) into the biogenic amine binding site.[27] The structural model is consistent with the existing knowledge on structure – activity relationships for small-molecule antagonists. The key feature of the model is an ionic interaction between the acid moiety of the ligand and Lys199. The authors suggest that AngII also binds its receptor by an ionic interaction between its C terminus and Lys199. Ji and co-workers provided experimental validation for this hypothesis by site-directed mutagenesis on the rat AT_1 receptor; they revealed a 42-fold reduction of the binding of the AngII analogue saralasin to the Lys199Ala mutant.[26] The mutagenesis data presented in this study indicate that peptide and nonpeptide ligand binding sites on the AT receptor are distinct. However, the reduced binding affinity of the Lys199Ala and also the Asn285Ser mutants for both saralasin and losartan indicates that the binding sites overlap to some extent.

3.3. Purinergic GPCRs

The adenosine A_3 and the purine $P2Y_1$ receptors belong to the family of purinergic receptors comprising the P1 family (for example, A_1, A_2, A_3) and the P2Y family (for example, $P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$, $P2Y_{11}$). While the A_3 receptor (P1 family) is activated by adenosine, the $P2Y_1$ receptor (P2Y family) is stimulated by extracellular ADP and ATP. For both receptors site-directed mutagenesis and molecular modeling studies have provided hypotheses for the binding site of the physiological ligand.[28–31] According to these studies, adenosine and ATP bind to the transmembrane cleft of their respective receptors.

For the $P2Y_1$ receptor the negatively charged triphosphate moiety of ATP and other nucleotide derivatives is likely to be coordinated to Arg128 (TM3), Thr222 (TM5), Lys280 (TM6), and Arg310 (TM7). Supported by structure – activity studies on N^6-alkylated ATP derivatives and mutagenesis studies, the model of the $P2Y_1$ – ATP complex (displayed schematically in Scheme 3) furthermore reveals a potential interaction of the adenine moiety with the receptor by the exocyclic N^6 atom (hydrogen bond to Gln307 (TM7)) and the N1 atom of the adenine ring (hydrogen bond to Ser314 (TM7)). The ribose moiety of ATP appears to be coordinated by Phe131 (TM3), His132 (TM3), and His277 (TM6). These residues are modulatory for agonist action and might interact with the 2′ or 3′ hydroxy groups of the ATP ribose unit.



Scheme 3. *Schematic drawing of the interaction of ATP with the $P2Y_1$ receptor as derived from mutagenesis and modeling studies. For reasons of clarity the interaction with Arg310(TM7) is not shown. References are given within the text.*

3.4. Monocyte chemoattractant-1 (MCP-1) receptor, CCR2

The chemokine receptors (CCRs) represent another subfamily of family A GPCRs stimulated by small proteins; these proteins mediate attraction of leukocytes to inflammatory sites and are known as chemokines.[32] Astonishingly, for many of the chemokine receptors, small-molecule compounds could also be identified antagonizing the action of these small-protein ligands at their respective receptor. For the monocyte chemoattracant-1 (MCP-1) receptor (CCR2), researchers at Roche identified spiropiperidine compounds (for example, RS-102895) binding to CCR2 with a dissociation constant K_D of 60 nM.[33] Mutagenesis studies showed that an acidic glutamate residue in the TM7 helix (found in most chemokine receptors) is critical for binding of the spiropiperidine series. It was hypothesized that the basic nitrogen atom present in the spiropiperidine compounds may be the interaction partner for Glu291 and a model of RS-102895 bound to the CCR2 receptor was generated by using the bacteriorhodopsin structure as template (Figure 4). The structural model suggests that the acid–base pair anchors the spiropiperidine compound within the transmembrane ovoid bundle and that the binding site may overlap with the space required by MCP-1 binding. Interestingly, the Glu291 residue from helix TM7 (VII in Figure 4) is in a similar spatial position to the acidic aspartate residue contributed from helix TM3 of the biogenic amine receptors, which may account for the shared affinity of spiropiperidines for these two receptor classes (see also Section 5.2).

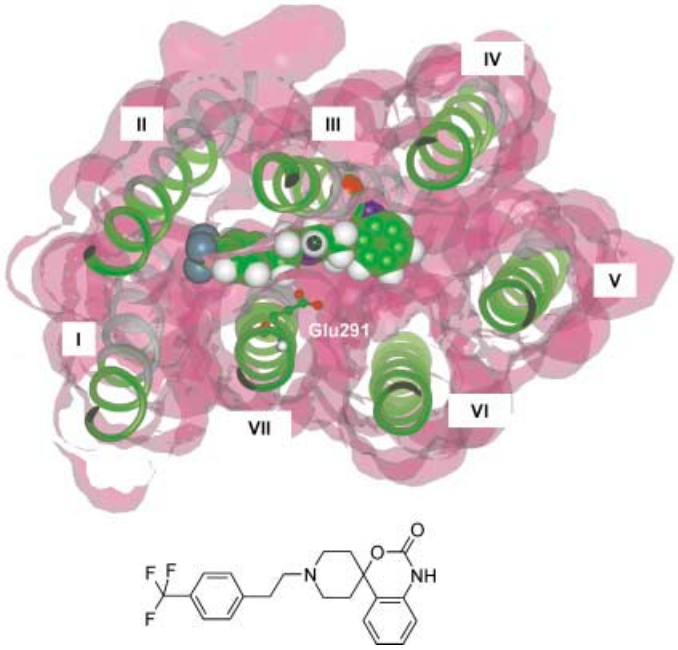


Figure 4. *Model of RS-102895 bound to the MCP-1 receptor. RS-102895 is the space-filling molecule in the center of the bundle of helices, indicated by the ribbons.[33] Receptor residue Glu291 is shown in the ball-and-stick presentation; it is shown interacting with the basic nitrogen atom of the spiropiperidine structure. Please note that in contrast to Figure 2B the view from the intracellular side is displayed.*

Recently, a study has been presented which applies structural models of chemokine and biogenic amine receptors within the chemical optimization program of a CCR5 lead.[34] The initial lead compound showed high affinity towards the CCR5 receptor but was unselective for several biogenic amine receptors and especially for the muscarinic M_2 receptor. Structural differences in the binding pockets of computational models from the biogenic amine and the CCR5 receptor explain how the initially unselective chemokine antagonist could be chemically modified to yield a selective derivative of the initial lead structure.

3.5. A common binding pocket within the 7-TM region?

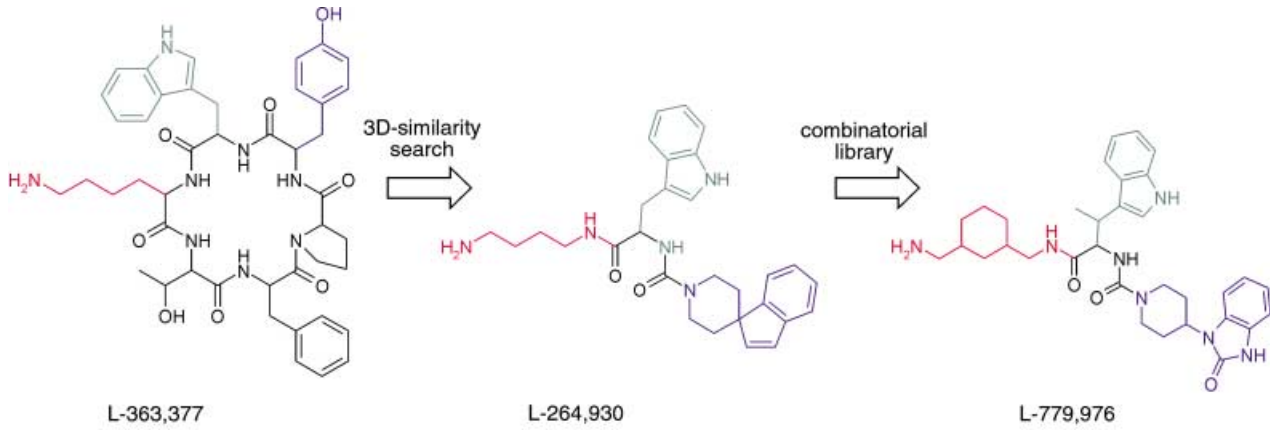
Based on the molecular modeling and mutagenesis studies performed for several GPCRs that bind diverse types of ligands (biogenic amines, peptides, nucleotides, and small proteins), it is tempting to suggest that all family A receptors share a binding pocket located deeply within the 7-TM region, that can serve as an interaction site, not only for monoamines, but for all (ant)agonists of this receptor family.[35, 36] However, this might be a simplified view. The ligands do not necessarily have to bind at this site to stabilize the receptor conformation that is recognized by the G protein as being 'active'.[37] Thus, even certain antibodies directed towards extracellular loops of 7-TM receptors can mimic the action of endogenous ligands and activate the receptor.[38] In addition, allosteric modulation of GPCRs by small molecules enhancing or diminishing the effects of endogenous agonists or antagonists has long been recognized and has recently been reviewed.[39, 40] Most interestingly, the active 7-TM receptor conformation is not only stabilized by an agonist, but the receptor can also convert into the active conformation even in the absence of an agonist, as demonstrated by the frequently observed constitutive signaling.[41, 42]

4. Ligand-Based Drug Design

4.1. Ligand-based lead finding

Due to the limited availability of structural data on GPCRs the design of ligands for this receptor family still heavily relies on ligand-based drug design techniques. For many GPCRs the natural ligand can provide a good starting point in the lead finding process. Structure–activity relationships (SAR) can be directly derived from the natural ligand and its analogues. The resulting pharmacophore models can then be employed for virtual screening to identify lead structures with novel scaffolds. Especially for peptide-binding GPCRs the identification of a non-peptidic ligand is crucial for drug discovery to avoid the inherent pharmacokinetic problems associated with peptide lead structures like poor oral bioavailability or metabolic instability.

The application of this technique to identify novel non-peptidic lead structures was successfully demonstrated for the sst receptor by researchers at Merck. Optimization of somatostatin-derived peptides has resulted in the cyclic hexapeptide c[-Pro-Tyr-D-Trp-Lys-Thr-Phe-] (L-363,377) as a somatostatin agonist.[43] It could be shown that the Tyr-D-Trp-Lys motif is required for the biological activity. The side chains of this motif were used



Scheme 4. *A three-dimensional representation of the Tyr-D-Trp-Lys motif of the cyclic hexapeptide L-363,377 was used to virtually screen the Merck compound collection.[44] L-264,930 was identified as a hit compound, which was further optimized by using combinatorial follow-up libraries to yield L-779,976. The color coding indicates how the different parts of the non-peptidic compounds map to the Tyr-D-Trp-Lys motif of the peptidic lead; this motif is necessary for the biological efficacy.*

as a probe to search Merck's compound collection with a 3D search method (Scheme 4).[44, 45] Biological testing of only 75 compounds yielded compound L-264,930 with an apparent inhibition constant of 100 nM for the human sst2 receptor. Its tripartite structure stimulated the synthesis of several combinatorial follow-up libraries, which contained highly active agonists for the human sst2 receptor as well as other subtype selective somatostatin agonists. Thus, virtual screening in combination with combinatorial chemistry has resulted in numerous subtype selective somatostatin agonists.

Recently, researchers at Aventis have employed a similar virtual screening strategy to identify non-peptidic antagonists for the urotensin II receptor.[46] Various truncated peptide derivatives of the cyclic 11 amino acid peptide urotensin II were synthesized to determine the minimal sequence required for its biological activity. Further on, an alanine scan identified the Trp-Lys-Tyr motif in the cyclic part of human urotensin II as the important pharmacophoric pattern. The spatial arrangement of the Trp-Lys-Tyr motif was deduced from the NMR spectroscopy solution structure of human urotensin II and of the disulfide-bridged analogue Ac-Cys-Phe-D-Trp-Lys-Tyr-Cys-NH_2 (Figure 5). Both results were translated to respective three-dimensional pharmacophore models, which were used to virtually screen the Aventis compound collection. By using the pharmacophore model of the urotensin II analogue as a query, only 1 of the 418 virtual hits showed biological activity. The obtained hit rate is similar to the hit rate observed for high-throughput screens for GPCR antagonists. Interestingly, the pharmacophore model,

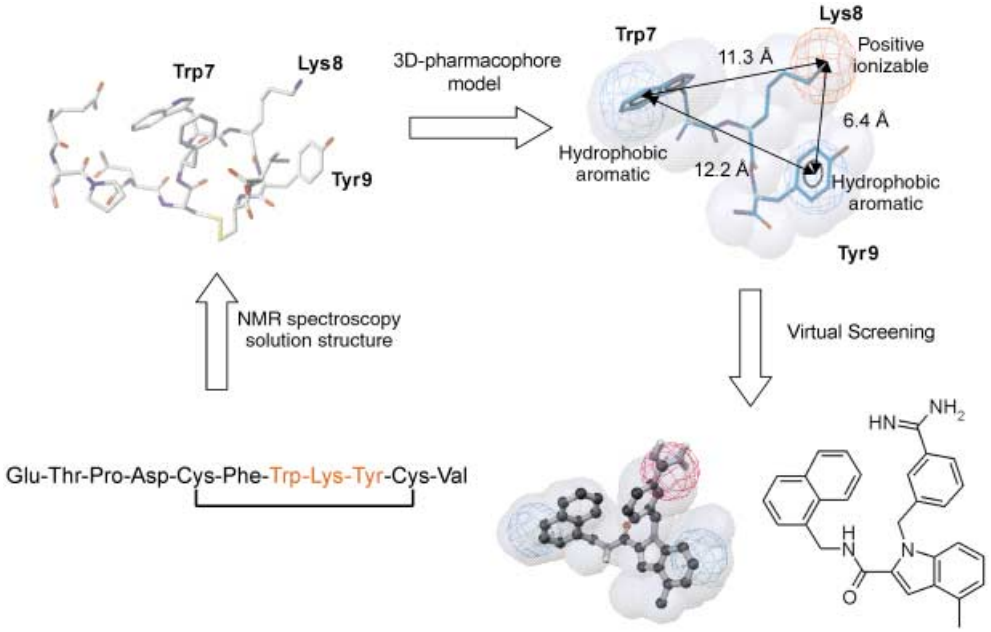


Figure 5. *Virtual screening approach to identify non-peptide urotensin II antagonists.[46] For the cyclic undecapeptide urotensin II (bottom left), the NMR spectroscopy solution structure was determined (top left). The spatial arrangement of the Trp-Lys-Tyr motif as found in the solution structure was used to derive a pharmacophore model (top right). The mapping of a discovered hit structure to the pharmacophore model is shown (bottom right).*

obtained from the human urotensin II structure, showed a hit rate of 2% (10 actives out of 500 compounds tested) with an IC_{50} value of 400 nM for the best compound as derived in a functional in vitro assay. The virtual screening approach has identified numerous novel scaffolds with reasonable antagonist activity providing promising starting points for subsequent chemical optimization programs.

The high hit rate indicates that the solution structure of urotensin II, which was used as structural template for the generation of the pharmacophore model, resembles its receptor-bound conformation. In general, for a high-affinity ligand the conformational energy difference between bound and unbound state is small, thus avoiding the reduction of the free binding energy. However, in several cases conformational changes of the ligand have been observed upon receptor binding.[47, 48] In those cases, the solution structure of the natural ligand does not appear to be a valid template for the search for new lead structures by virtual screening.

The ambiguity resulting from conformationally flexible ligands can be avoided by using fairly rigid molecules to derive a pharmacophore model, as was demonstrated by Flower and co-workers in the search for antagonists of the muscarinic M_3 receptor.[49] M_3 receptor antagonists from three lead series were available. A representative compound was chosen from each series to derive two different pharmacophore models. With these three-dimensional pharmacophore models 172 virtual hits could be identified by searching the Astra compound bank, of which three compounds with a novel chemical scaffold showed a significant biological activity at the muscarine receptor.

The examples given demonstrate the potential of virtual screening for lead identification in ligand-based design programs. Agonists as well as antagonists can be derived from natural or surrogate ligands. The generation of a valid pharmacophore model and thus the hit rate of virtual screening approaches benefits from the availability of potent, structurally diverse, and conformationally restricted receptor ligands as starting points.

Further examples for the successful use of peptide-derived structure – activity relationships to design non-peptidic GPCR ligands are described for the opiate receptor,[50] the thrombin receptor,[51] and the growth hormone secretagogues[52] and somatostatin[53] receptors.

However, especially for peptide-binding GPCRs, screening of diverse or focused compound sets still remains a successful lead finding approach, which has yielded the discovery of several potent, non-peptidic GPCRs ligands.[54–56] Such compounds have been classified as functional mimetics as they elicit agonist or antagonist activity, but do not necessarily mimic the structure of the native ligand.

4.2 Ligand-based lead optimization

Ligand-based three-dimensional quantitative structure – activity relationship (3D-QSAR) methods, like the comparative molecular field analysis (CoMFA),[57, 58] have supported the chemical optimization of numerous GPCR lead compounds. CoMFA correlates the steric and electronic field environment of a set of ligands with their biological activity. Thus, a CoMFA study allows chemical modifications that are benefical or detrimental for the biological activity to be recognized. The interested reader is referred to successful case studies of the CoMFA method in optimizing GPCR-directed ligands as described, for example, for the dopamine receptors,[59–61] the serotonin receptors,[62–64] the endothelin receptor,[65] and the adenosine receptors.[66, 67]

CoMFA models can also be used to recognize molecular features that are responsible for selectivity of the ligands. López-Rodriquez and co-workers observed side affinities for the α_1-adrenergic receptor in a series of aryl piperazines that were active against the 5-HT_{1A} receptor.[68] For each receptor a separate CoMFA model was derived. The comparison of the models indicated that bulky substituents at the *meta* position of the aryl moiety would increase selectivity for the 5-HT_{1A} receptor, since the α1 receptor, in contrast to the 5-HT_{1A} receptor, does not tolerate large residues at this position (Figure 6). Further on,

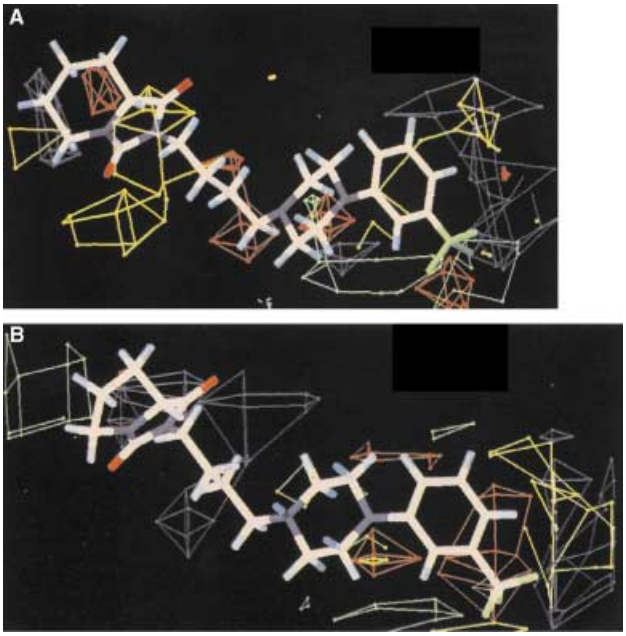


Figure 6. *A) CoMFA model of aryl piperazines for the $5HT_{1A}$ receptor. At the meta position of the aryl piperazine moiety green polyhedra indicate that larger subtituents are beneficial for the affinity towards the $5HT_{1A}$ receptor. In contrast, the CoMFA for the α1 receptor (B) shows that larger substiuents at the meta position of the aryl piperazine moiety would reach a yellow polyhedra, which indicates that steric bulk is detrimental for affinity. Reproduced from ref. [68] with permission. Copyright (1997) American Chemical Society.*

increasing the length of the alkyl chain linking the aryl piperazine with a hydantoin moiety seems to be beneficial for the desired selectivity. Therefore, a bulky NHCO*i*Pr substituent at the corresponding position of the lead series was combined with a long alkyl linker chain. The resulting compound showed the desired high selectivity for the 5-HT_{1A} receptor.[68] Instead of building separate models, the selectivity issue can also be addressed by developing CoMFA models against the ratio of the affinities of each receptor. Such an approach was chosen by Raviña and co-workers for the analysis of selectivities of

conformationally restricted butyrophenones against the dopamine and the serotonin receptors.[69]

The combination of a ligand-derived CoMFA model with a three-dimensional receptor model provides information about the potential binding mode of the investigated ligands and can give additional ideas for compound optimization. Recently, researchers at SmithKline Beecham used such an approach on an indoline urea lead series directed against the 5-$HT_{2C/2B}$ receptors; the series showed only minor side affinity towards the 5-HT_{2A} receptor.[70, 71] The docking of an indoline urea compound into a 3D receptor model placed the indolinyl moiety into a receptor pocket for which key differences between the 5-HT_{2A} and the 5-HT_{2C} receptors could be recognized from the sequence comparison (Figure 7). Due to the presence of two leucine residues in the 5-HT_{2A} receptor instead of two valine residues in the 5-HT_{2C} receptor the binding site of the 5-HT_{2A} receptor appeared to be more constricted than the binding site of the 5-HT_{2C} receptor. In the optimization of the series this key structural difference was employed to increase the selectivity by introducing larger substituents. A series of 55 compounds was synthesized and analyzed with a CoMFA study. The molecules were aligned by docking the ligands into the receptor model. The resulting CoMFA model was able to predict the biological activity of a test set reasonably well. In addition, the presence and location of sterically disfavored regions recognized by the CoMFA model agrees well with the 3D receptor model.

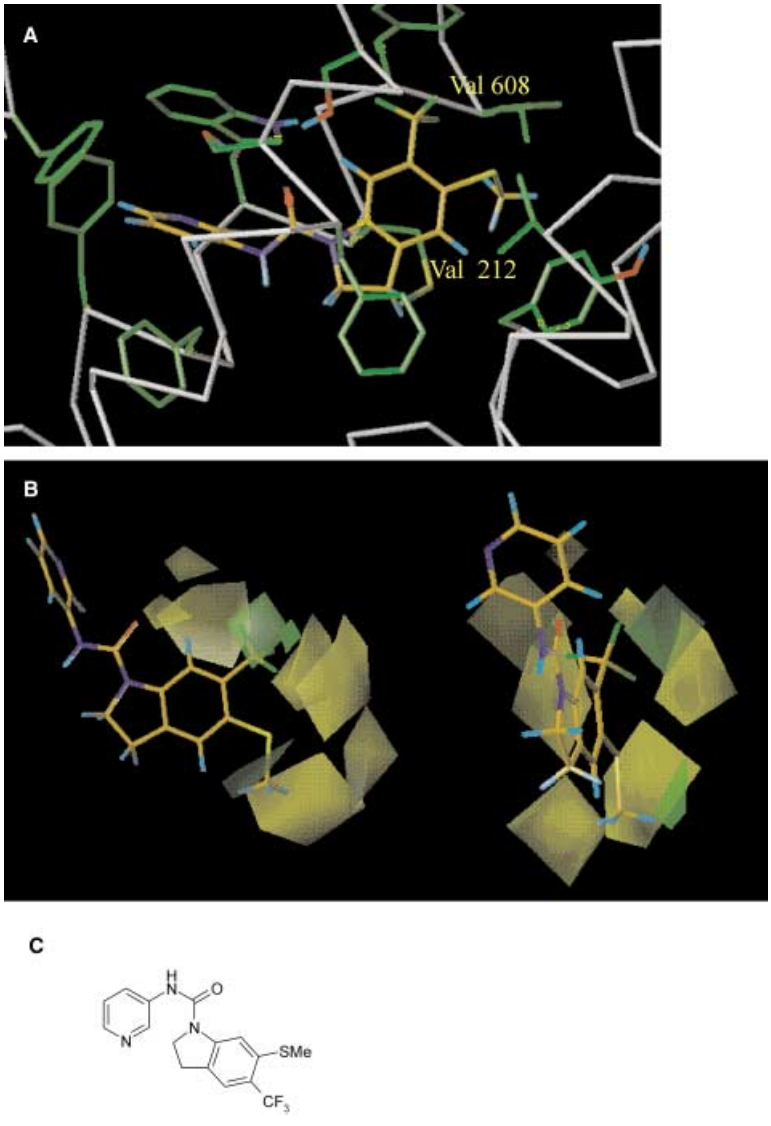
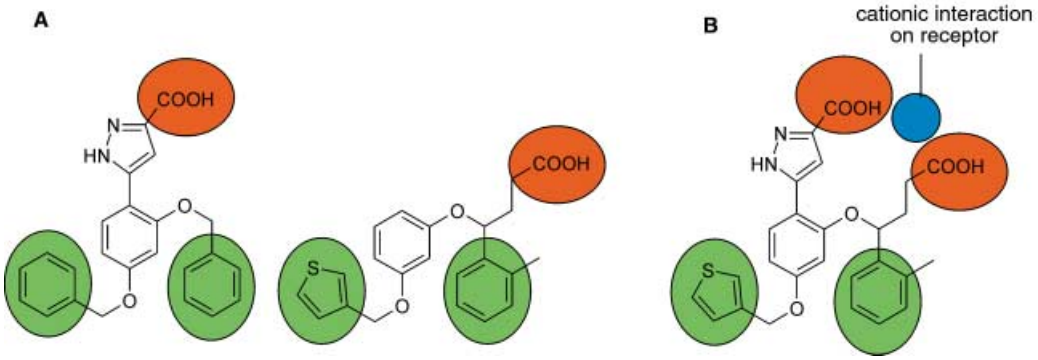


Figure 7. *A) Docking mode of an indoline urea compound, depicted in (C), into the 5-HT_{2C} receptor model. The thiomethyl substituent points towards the "specifity pocket" lined by residues Val212 and Val608. For 5-HT_{2A} these valine residues are replaced by bulkier leucine residues to result in a narrower binding pocket.[70, 71] B) Orthogonal views of the CoMFA steric fields around the indoline urea. The yellow regions identify positions where steric bulk is detrimental for activity. Around the thiomethyl substituent, the yellow regions show a sterically confined region in agreement with the three-dimensional receptor model depicted in (A). Reproduced from ref. [71] with permission. Copyright (1998) American Chemical Society.*

It is not always evident whether two chemically different series targeting at the same receptor share a common binding region. Consequently, often both lead series are independently optimized. Alignment of molecules from different chemical classes can give first hints about a common binding mode, especially if the SAR of corresponding parts of the molecules show parallel trends. In such cases, elements of the first series can be used to optimize the second series at the next step. An illustrative example of such an optimization strategy has been described for the endothelin A receptor (Scheme 5). Astles and co-workers had developed two compound series which were originally derived from the same 3D database query.[72] Thus, it was conceived that these series might share a common binding mode, but no molecular alignment could be found that matched all relevant pharmacophores of the corresponding compounds. The two compound series could be combined based on the assumption that the acidic functions of the molecules might interact with the same receptor site but from different positions. Consequently, a chimeric compound was made, carrying elements of both series. This effort resulted in a compound with tenfold-increased affinity. As it already was known from one series that the thienyl moiety would be suitable for optimizing pharmacokinetic properties, only a few modifications were necessary to receive a highly active compound with an acceptable pharmacokinetic profile.

5. Focused Libraries Directed against GPCR Targets

In the beginning of the last decade parallel synthesis and combinatorial chemistry together with automated high-throughput screening methodologies were expected to speed up the lead discovery process. The rapid and cheap testing of large generalized libraries of diverse compounds was a common strategy within many pharmaceutical companies. However, as indicated by the number of new chemical entities, the desired increase of productivity has not been achieved.[73, 74] Nowadays a paradigm shift has occurred towards the synthesis and screening of small focused



Scheme 5. *A) Comparison of two lead structures for the endothelin A receptor. The hydrophobic groups are shown by the green circles, while the acid functions are shown by the red circles.[72] B) The two compounds could not be aligned in a way that all important pharmacophoric groups were matched. The hypothesis of a cationic interaction center on the receptor (blue circle) allowed the two series to be merged by assuming that the acidic groups interact with this center from different directions. The merged compound revealed a tenfold increase in affinity.*

compound collections that are often designed and directed against target families. The knowledge of the structural requirements for activity of a compound at a certain target family can be derived from structural information on the target class or from the analysis of potent ligands acting on the respective receptor family. The design of combinatorial libraries by using molecular properties and 3D pharmacophore fingerprints has been reviewed recently.[75]

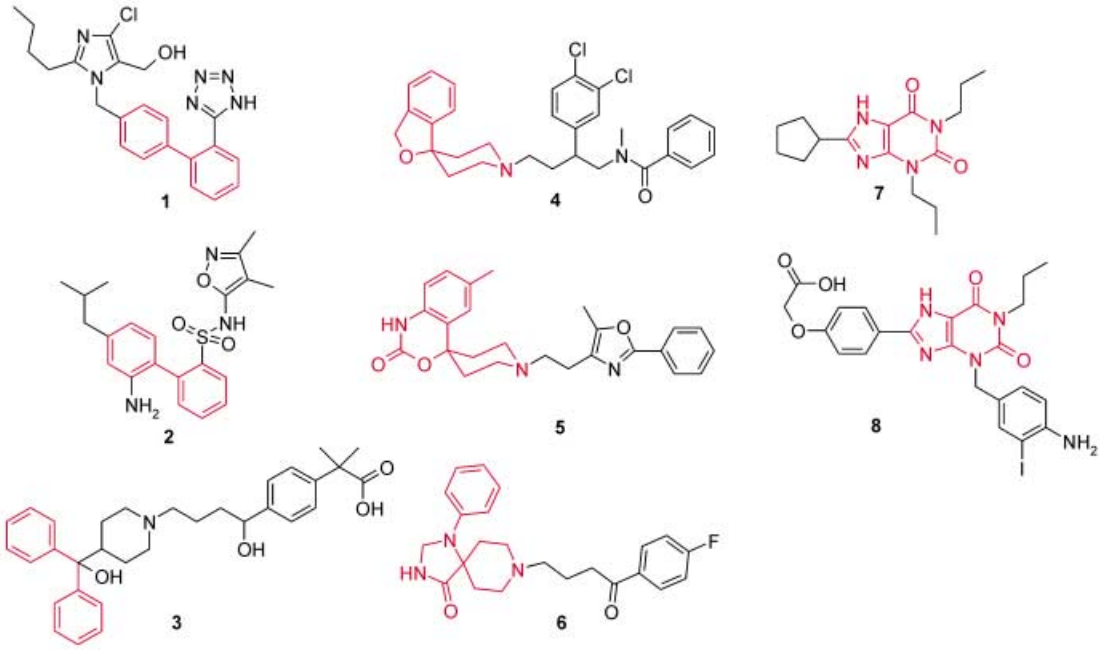
Due to the distinguished past of GPCR research there is a large number of commercially successful compounds acting on GPCRs (see Table 1 and Scheme 1). In addition, for each of these marketed drugs there are hundreds of compounds that, while being potent ligands of GPCRs, never made it to the market. This wealth of medicinal chemistry knowledge covering structural features present within small-molecule GPCR ligands is an excellent source to enhance the discovery of new lead compounds against novel GPCR targets. Many pharmaceutical companies use structural information about internal and public-domain GPCR ligands as a probe to virtually screen compound collections or even virtual compound libraries. Compounds showing similarity to known GPCR ligands are identified and then compiled into proprietary GPCR-directed screening collections or (in cases where a virtual library was analyzed) prioritized for chemical synthesis.

In the past many molecular and pharmacophore descriptors[76] have been successfully applied to define the molecular similarity (or dissimilarity) and to guide the design of GPCR-directed libraries. Mason and co-workers have described a four-point pharmacophore method for molecular similarity, which calculates all potential pharmacophores for a given molecule.[77] The authors were able to identify GPCR-specific pharmacophores and to apply the pharmacophore key for the design of GPCR-focused libraries. A similar approach was used by Bradley et al. to generate a 3D pharmacophore model from 105 published heterocyclic α_1-adrenergic receptor ligands encompassing multiple chemical classes.[78] The pharmacophore descriptor was then applied as a computational filter to assist lead evolution efforts.

Pearlman and his colleagues have described the molecular BCUT descriptor, which combines physicochemical properties relevant to ligand–receptor binding with topological or distance information.[79] The Pearlman descriptors have been utilized for a variety of diversity-related purposes including reagent selection for diversity libraries. At Neurocrine, first examples of the descriptors' application for the design of compound libraries focused against peptide-binding GPCRs have recently been presented.[80] Researchers at Aventis have used 2D descriptors comparing molecules based on the presence (or absence) of functional fragments to compile a proprietary GPCR-screening collection from internal and external compound libraries. Initial results from screening the GPCR-focused compound collection against novel GPCR targets indicate significantly higher hit rates compared to screening of an unbiased library.[81]

5.1. Privileged substructures within GPCR ligands

Accumulated evidence suggests that there are many common features apparent in the small-molecule binding sites of different GPCRs (for example, the presence of an Asp residue within helix TM3 of biogenic amine and several peptide-binding GPCRs). This finding is reflected in the existence of many structural motifs common among the small-molecule ligands of diverse receptors. Originally, Evans et al. introduced the term "privileged structure" for benzodiazepines, which are found in several types of central nervous system agents and in ligands of ion channels and GPCRs.[82] According to this definition, a privileged structure, such as a benzodiazepine, "is a single molecular framework able to provide ligands for diverse receptors"; it concludes that "judicious modification of such structures could be a viable alternative in the search for new receptor agonists and antagonists". Further examples of GPCR privileged substructures like biphenyl, 1,1-diphenylmethane, xanthines, 4-arylpiperidines, 4-arylpiperazines, and spiro versions of the latter are given in Scheme 6. Interestingly, some of the privileged substructures are not restricted to one GPCR subfamily. The spiropiperidine moiety can be found within ligands of biogenic amine receptors as well as within compounds acting on chemokine and peptide-binding GPCRs. This finding provides further evidence that these GPCR



Scheme 6. *GPCR privileged substructures: Examples of GPCR ligands sharing the biphenyl or diphenylmethyl (**1**–**3**), spiropiperidine (**4**–**6**), or xanthine (**7**–**8**) moiety are shown. **1**, Angiotensin II type-1 antagonist (losartan); **2**, endothelin-A antagonist (preclinical);[104] **3**, histamine-H1 antagonist (fexofenadine); **4**, neurokinin NK_2 antagonist (preclinical); **5**, monocyte chemoattractant-1 (MCP-1)/CCR2 antagonist; **6**, dopamine antagonist (spiperone); **7**, adenosine (A_1) antagonist (phase II); **8**, adenosine (A_3) antagonist.*

subfamilies share a common small-molecule binding site despite their diverse physiological ligands.

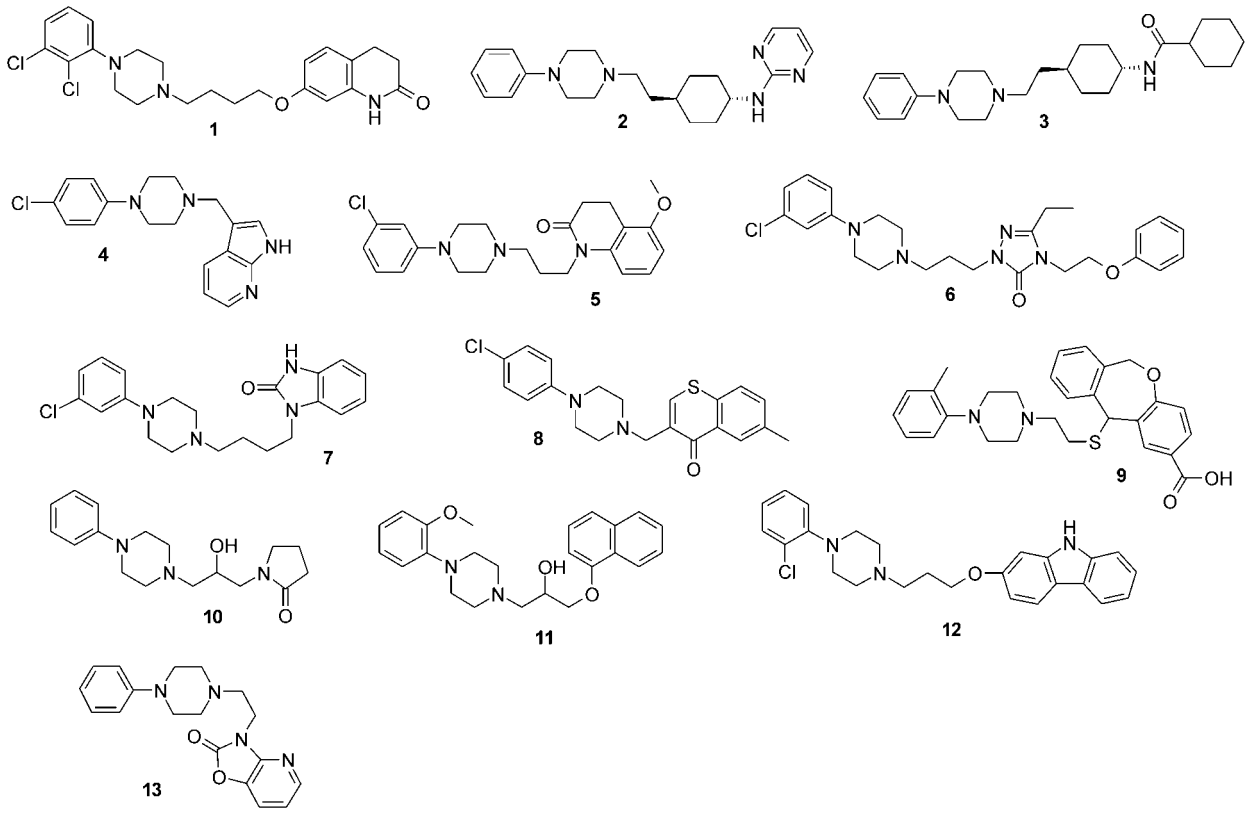
Computational methods are available to recognize biologically active fragments within known GPCR ligands. In a recent study more than 20000 compounds known to act on GPCRs have been virtually cleaved into their fragments with the RECAP (retrosynthetic combinatorial analysis procedure) algorithm,[83] with the assumption of 11 default bond cleavage types.[84] The resulting fragments have then been ranked to identify those which occur in a set of ligands modulating several different GPCRs. The most frequently occurring fragment found was the 4-phenyl-piperazine, present in a set of 32 compounds which act at 13 different GPCRs (Table 2 and Scheme 7).

Table 2. *GPCR ligands sharing the GPCR privileged substructure 4-phenyl-piperazine. Compound numbers refer to structures given in Scheme 7.*

Structure	Company	Status	Activity
1	Otsuka	phase III	D_2 antagonist
2	Pfizer	preclinical	D_2 agonist
3	Pfizer	preclinical	D_3 antagonist
4	MSD	phase II	D_4 antagonist
5	Otsuka	phase II	5-HT_{1A} agonist
6	BMS	launched	5-HT_{2A} antagonist
7	Boehringer Ingelheim	biological testing	5-HT_2 antagonist
8	Taisho	biological testing	5-HT_3 antagonist
9	Kyowa Hakko	biological testing	antihistaminic
10	Med.Acad.	preclinical	α antagonist
11	Roche	launched	α_1 antagonist
12	Taisho	biological testing	sigma antagonist
13	Servier	preclinical	substance P antagonist

Bemis and Murcko analyzed the molecular framework, which is defined as the union of ring systems and linkers within a molecule, of all commercially available drugs in order to identify common drug shapes.[85] In a similar manner, molecular frameworks of 20000 known GPCR ligands can be calculated to identify GPCR privileged frameworks. An example of a molecular framework as common motif of several GPCR binding molecules is shown in Table 3 and Scheme 8.

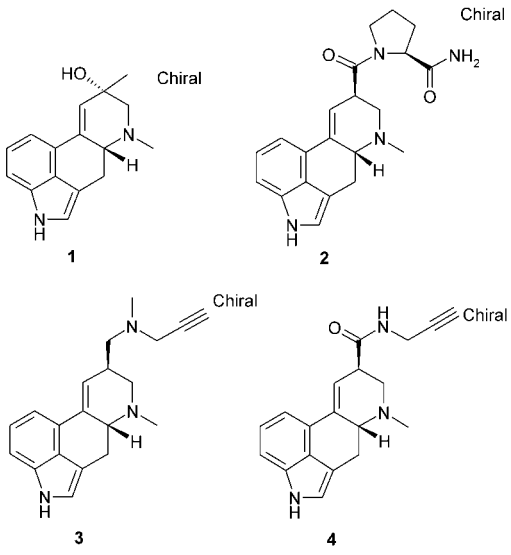
Several libraries based on privileged scaffolds or the incorporation of privileged fragments as building blocks have been synthesized as one strategy for the generation of focused libraries; examples for benzodiazepines, dihydropyridines, 4-phenylpiperidine and 1,1-diphenyl units are described in refs. [86, 87] and have also been reviewed.[88–90] Researches at Merck have made extensive use of privileged structures as building blocks to generate a somatostatin-directed combinatorial library; they chose a capped amino acid scaffold as the molecular framework.[91] After application of initial computational methods leading to the discovery of a 200 nM murine somatostatin subtype-2 (sst2) active compound (see Section 4.1.) a combinatorial library was designed and synthesized to yield selective and highly active ($K_i = 50$ pM to 200 nM) non-peptide small-molecule ligands for somatostatin subtypes-1, -2, -4, and -5. The use of a capped amino acid scaffold decorated with GPCR privileged 4-aryl piperidines/piperazines or their corresponding spiro derivatives has been proven to be a very rewarding approach for the identification of potent antagonists and even agonists for many other peptide-binding GPCRs. In addition to the somatostatin ligands, compounds sharing this GPCR privileged scaffold were found to act on the



Scheme 7.

Table 3. *GPCR ligands sharing the GPCR privileged framework hexahydroindolo[4,3-fg]quinoline. Compound numbers refer to structures given in Scheme 8.*

Structure	Company	Status	Activity
1	Takeda	biological testing	substance P antagonist
2	Poli Industria Chimica	biological testing	dopamine agonist
3	LEK	preclinical	5-HT_{2A}/D_2 antagonist
4	LEK	biological testing	5-HT_{1A}/sigma antagonist



Scheme 8.

melanocortin, substance P, cholecystokinin, endothelin, gastrin-releasing peptide, and growth hormone secretagogue receptors (Scheme 9).

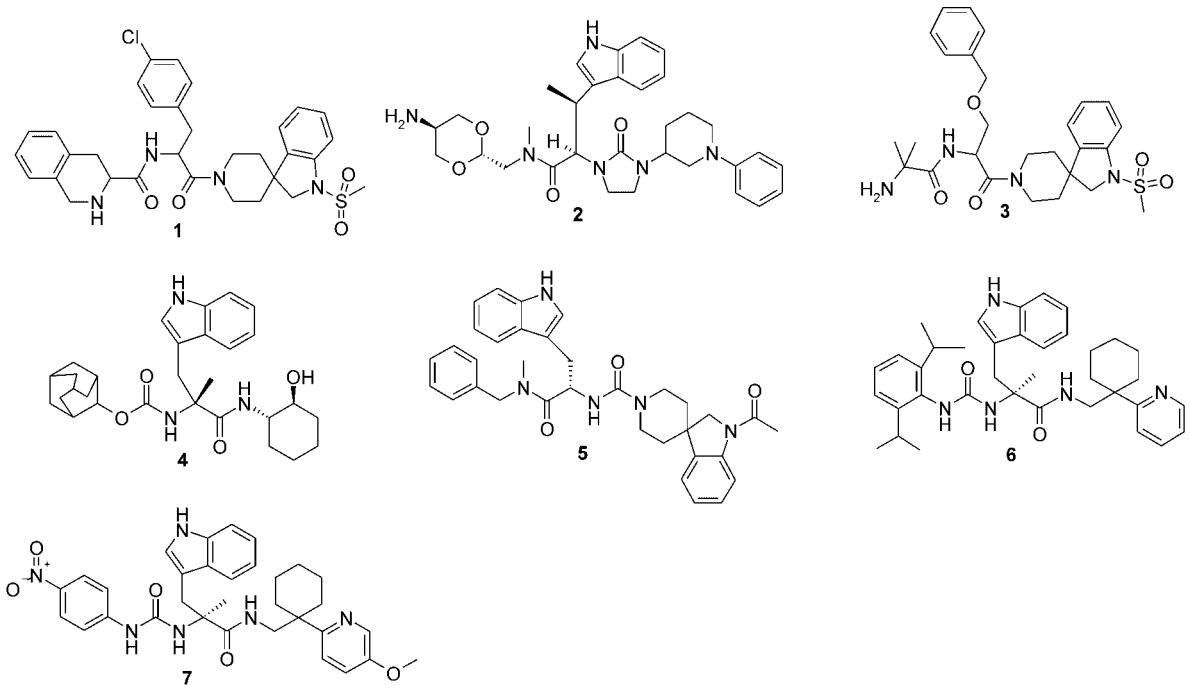
The various examples given above indicate that the usage of privileged substructures for lead finding offers the chance to quickly identify new lead compounds against novel GPCR targets. However, it is important to note that compounds sharing a privileged substructure often reveal activity against many GPCRs. Thus, receptor selectivity of these lead series has to be addressed within the lead optimization process.

6 Future Trends and Challenges

6.1. Optimizing pharmacokinetic properties

Apart from optimizing the pharmacodynamic properties (for example, binding affinity towards receptor, in vitro efficacy), the pharmacokinetic profile of a drug candidates is of paramount importance. About 40 % of drug candidates in clinical trials failed due to poor pharmacokinetics.[92] Therefore, the pharmaceutical industry addresses the issues of absorption, distribution, metabolism, and excretion (ADME) in an early phase of the drug discovery process.[93] Nowadays, compound optimization is truly a multidimensional task, comprised of the joint optimization of affinity to the receptor and ADME-related properties.

Early ADME considerations have resulted in the establishment of different experimental model systems to determine ADME-related parameters by in vitro experiments. Examples are Caco-2



Scheme 9. *GPCR ligands of the capped amino acid type.* ***1****, melanocortin-4 (MC4) receptor agonist;*[105] ***2****, somatostatin (sst2) agonist (clinical candidate);*[106] ***3****, growth hormone (GH) secretagogue antagonist (phase II);*[107] ***4****, cholecystokinin (CCK2) antagonist (clinical candidate);*[108] ***5****, neurokinin (NK_1/NK_2) antagonist;*[109] ***6****, neuromedin-B (NMB)/bombesin (BB_1) antagonist (clinical candidate);*[110] ***7****, bombesin BB_1/BB_2 antagonist.*[111]

cell permeation as a model for intestinal absorption or S9 assays for drug metabolism. Further on, huge efforts have been undertaken to develop in silico models to describe and predict ADME properties based on the chemical structure of the compound. These tools are applied to support the generation of screening libraries as well as the chemical optimization of lead compounds.[94] Such models comprise fairly simple counting rules like the rule of five.[95] It predicts a compound to be a poorly absorbed drug candidate, if it violates two of the following rules: molecular weight < 500 Da, number of hydrogen-bond donors < 5, number of hydrogen-bond acceptors < 10, and logP < 5. More advanced computer models can be derived from experimental in vivo or in vitro ADME-parameters. Recently, various computational techniques for the prediction of ADME properties have been extensively reviewed.[96–99]

6.2. Chemogenomics and structural biology—linking chemical and biological space

In recent years structural biology, homology modeling, and mutagenesis studies have provided valuable insights into GPCR structure and receptor–ligand interactions (see Sections 2 and 3). Typical receptor sites or fingerprints on the primary sequence level for binding of the natural and surrogate ligands have been identified, thereby defining the "biological space" of this target family. On the other hand, knowledge of privileged structural motifs of GPCR ligands (see Section 5.1.) is steadily increasing and more accurately defining the receptor's "chemical space". Today the link between chemical space and biological space has been established to some extent. It is, for example, well accepted that for the biogenic amine receptors the aspartate residue in helix TM3 constitutes the key anchor site of the privileged basic (spiro)piperidine moiety. In the near future it might be possible to identify further "privileged substructure-binding subsites" or "privileged substructure-binding fingerprints", thus bridging chemical and biological space. Nowadays a receptor entering the drug discovery process is still classified based on the natural ligand (for example, peptides, chemokines, biogenic amines) or upon its overall sequence homology with known GPCRs (based on phylogenetic tree analysis). In the future a novel receptor (orphan or liganded) could be classified according to the presence of "privileged substructure-binding fingerprints". This information will aid the design and synthesis of lead-finding libraries tailor-made for the respective receptor.

Recently, a first study on this chemogenomic knowledge-based ligand design strategy has been published for the monoamine-related GPCRs.[100] Based on the sequence analysis of residues contributing to each of the three subsites known for the biogenic amine binding site (see Section 3.1.), 50 receptors were classified. Based on this analysis ligand predictions were made for several orphan receptors.

Although powerful bio- and chemoinformatic tools might be developed in the future to provide the desired linkage between chemical and biological spaces, the success of this approach will clearly depend on the availability of further valid structural information on GPCRs in complex with their natural or surrogate ligands. The elucidation of 3D structures of several membrane proteins within the recent years (for example, cytochrome C oxidase, bacteriorhodopsin, potassium channel, bovine rhodopsin) and the recent formation of the first global network on

structural genomics on GPCRs (MePNet)[101] provide some realistic expectations that this challenging task will be fulfilled in the near future.

Abbreviations

5-HT	5-hydroxytryptamine, serotonin
7-TM	seven-transmembrane-helix (receptor)
8-OH-DPAT	8-hydroxy-N,N,-dipropylaminotetralin
A	adenine
ADME	absorbtion, distribution, metabolism, excretion
ADP	adenosine diphosphate
AngII	angiotensin II
AT	angiotensin
ATP	adenosine triphosphate
BB	bombesin
cAMP	cyclic adenosine monophosphate
CCR	chemokine receptor
CCK	cholecystokinin
cGMP	cyclic guanosine monophosphate
CoMFA	comparitive molecular field analysis
D	dopamine
DAG	diacylglycerol
GH	growth hormone
GABA	γ-aminobutyric acid
GDP	guanosine diphosphate
GPCR	G-protein-coupled receptor
GTP	guanosine triphosphate
H	histamine
Ile	isoleucine
LH-RH	luteinizing hormone-releasing hormone
LTD4	leukotriene D4
M	muscarine
MCH	melanocyte concentrating hormone
MCP	monocyte chemoattractant
NK	neurokinin
P	purine
PKC	protein kinase C
PLC	phospholipase C
QSAR	quantitative structure – activity relationship
SAR	structure – activity relationship
sst	somatostatin
Sar	sarcosine
TM	transmembrane
UTP	uridine triphosphate

The authors would like to thank Evi Kostenis, Holger Heitsch (Aventis) and Danka Elez (Max-Planck-Institut für Biophysik, Frankfurt) for many valuable discussions on this exciting receptor family and for critical reading of the manuscript. We are grateful to Ronald Stenkamp, Tara Mirzadegan, and Ian Forbes for providing illustrations of their research work cited within this article.

[1] Woodmac PharmaView2001 at http://www.woodmacresearch.com/phview
[2] J. C. Venter et al., *Science* **2001**, *291*, 1304 – 1350.
[3] A. Wise, K. Gearing, S. Rees, *Drug Discovery Today* **2002**, *7*, 235 – 246.
[4] D. R. Flower, *Biochim. Biophys. Acta* **1999**, *1422*, 207 – 234.
[5] U. Gether, *Endocr. Rev.* **2000**, *21*, 90 – 113.
[6] J. Wess, *Pharmacol. Ther.* **1998**, *80*, 231 – 264.
[7] T. K. Attwood, *Trends Pharmacol. Sci.* **2001**, *22*, 162 – 165.
[8] K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto, M. Miyano, *Science* **2000**, *289*, 739 – 745.
[9] D. C. Teller, T. Okada, C. A. Behnke, K. Palczewski, R. E. Stenkamp, *Biochemistry* **2001**, *40*, 7761 – 7772.
[10] V. M. Unger, P. A. Hargrave, J. M. Baldwin, G. F. Schertler, *Nature* **1997**, *389*, 203 – 206.
[11] Z.-L. Lu, J. W. Saldanha, E. C. Hulme, *Trends Pharmacol. Sci*. **2002**, *23*, 140 – 146.
[12] T. P. Sakma, *Curr. Opin. Cell Biol.* **2002**, *14*, 189 – 195.
[13] N. Kunishima, Y. Shimada, Y. Tsuji, T. Sato, M. Yamamoto, T. Kumasaka, S. Nakanishi, H. Jingami, K. Morikawa, *Nature* **2000**, *407*, 971 – 977.
[14] C. E. Dann, J.-C. Hsieh, A. Rattner, D. Sharma, J. Nathans, D. J. Leahy, *Nature* **2001**, *412*, 86 – 90.
[15] D. Röper, E. Jacoby, P. Krüger, M. Engels, J. Grötzinger, A. Wollmer, W. Strassburger, *J. Recept. Res*. **1994**, *14*, 167 – 186.
[16] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, J. K. Lanyi, *J. Mol. Biol*. **1999**, *291*, 899 – 911.
[17] M. C. Gershengorn, R. Osman, *Endocrinology* **2001**, *142*, 2 – 10.
[18] S. Shacham, M. Topf, N. Avisar, F. Glaser, Y. Marantz, S. Bar-Haim, S. Noiman, Z. Naor, O. M. Becker, *Med. Res. Rev.* **2001**, *21*, 472 – 483.
[19] M. R. Tota, M. R. Candelore, R. A. Dixon, C. D. Strader, *Trends Pharmacol. Sci*. **1991**, *12*, 4 – 6.
[20] T. M. Savarese, C. M. Fraser, *Biochem. J*. **1992**, *283*, 1 – 19.
[21] K. Wieland, H. M. Zuurmond, C. Krasel, A. P. Ijzerman, M. J. Lohse, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9276 – 9281.
[22] S. A. Green, A. P. Spasoff, R. A. Coleman, M. Johnson, S. B. Liggett, *J. Biol. Chem.* **1996**, *271*, 24 029 – 24 035.
[23] R. A. Coleman, M. Johnson, A. T. Nials, C. J. Vardey, *Trends Pharmacol. Sci*. **1996**, *17*, 324 – 330.
[24] E. Jacoby, J.-L. Fauchère, E. Raimbaud, S. Ollivier, A. Michel, M. Spedding, *Quant. Struct.-Act. Relat*. **1999**, *18*, 561 – 571.
[25] H. Ji, M. Leung, Y. Zhang, K. J. Catt, K. Sandberg, *J. Biol. Chem*. **1994**, *269*, 16 533 – 16 536.
[26] H. Ji, W. Zheng, Y. Zhang, K. J. Catt, K. Sandberg, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9240 – 9244.
[27] D. J. Underwood, C. D. Strader, R. Rivero, A. A. Patchett, W. Greenlee, K. Prendergast, *Chem. Biol.* **1994**, *1*, 211 – 221.
[28] S. Moro, D. P. Guo, E. Camainoni, J. L. Boyer, T. K. Harden, K. A. Jacobsen, *J. Med. Chem.* **1998**, *41*, 1456 – 1466.
[29] S. Moro, C. Hoffmann, K. A. Jacobsen, *Biochemistry* **1999**, *38*, 3498 – 3507.
[30] Q. Jiang, D. Guo, B. X. Lee, A. M. Van Rhee, Y. C. Kim, R. A. Nicholas, J. B. Schachter, T. K. Harden, K. A. Jacobson, *Mol. Pharmacol*. **1997**, *52*, 499 – 507.
[31] S. Moro, A. H. Li, K. A. Jacobson, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 1239 – 1248.
[32] M. K. Schwarz, T. N. C. Wells, *Nat. Rev. Drug Discovery* **2002**, *1*, 347 – 358.
[33] T. Mirzadegan, F. Diehl, B. Ebi, S. Bhakta, I. Polsky, D. McCarley, M. Mulkins, G. S. Weatherhead, J. M. Lapierre, J. Dankwardt, D. Morgans, Jr., R. Wilhelm, K. Jarnagin, *J. Biol. Chem.* **2000**, *275*, 25 562 – 25 571.
[34] T. Mirzadegan, unpublished results.
[35] S. Trumpp-Kallmeyer, J. Joflack, A. Bruinvels, M. Hibert, *J. Med. Chem.* **1992**, *35*, 3448 – 3462.
[36] T. W. Schwartz, M. M. Rosenkilde, *Trends Pharmacol. Sci.* **1996**, *17*, 213 – 216.
[37] T. W. Schwartz, U. Gether, H. T. Schambye, S. A. Hjorth, *Curr. Pharm. Des.* **1995**, *1*, 325 – 342.
[38] S. A. Alla, U. Quitterer, S. Grigoriev, A. Maidhof, M. Haasemann, K. Jarnagin, W. Müller-Esterl, *J. Biol. Chem*. **1996**, *271*, 1748 – 1755.
[39] W. Soudijn, I. Van Wijngaarden, A. P. Ijzerman, *Expert Opin. Ther. Pat.* **2001**, *11*, 1889 – 1904.
[40] A. Christopoulos, *Nat. Rev. Drug Discovery* **2002**, *1*, 198 – 210.
[41] T. P. Kenakin, *Trends Pharmacol. Sci.* **1995**, *16*, 188 – 192.
[42] T. P. Kenakin, *Nat. Rev. Drug Discovery* **2002**, *1*, 103 – 110.
[43] D. F. Veber in *Peptides, Chemistry, and Biology: Proceedings of the 12th American Peptide Symposium* (Eds.: J. A. Smiith, J. E. Rivier) ESCOM, Leiden, the Netherlands, pp. 3 – 14.

[44] S. P. Rohrer, E. T. Birzin, R. T. Mosley, S. C. Berk, S. M. Hutchins, D.-M. Shen, Y. Xiong, E. C. Hayes, R. M. Parmar, F. Foor, S. W. Mitra, S. J. Degrado, M. Shu, J. M. Klopp, S.-J. Cai, A. Blake, W. W. S. Chan, A. Pasternak, L. Yang, A. A. Patchett, R. G. Smith, K. T. Chapman, J. M. Schaeffer, *Science* **1998**, *282*, 737 – 740.
[45] L. Yang, S. C. Berk, S. P. Rohrer, R. T. Mosley, L. Guo, D. J. Underwood, B. H. Arison, E. T. Birzin, E. C. Hayes, S. W. Mitra, R. M. Parmar, K. Cheng, T. J. Wu, B. S. Butler, F. Foor, A. Pasternak, Y. Pan, M. Silva, R. M. Freidinger, R. G. Smith, K. Chapman, J. M. Schaeffer, A. A. Patchett, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 10836 – 10841.
[46] S. Flohr, M. Kurz, E. Kostenis, A. Brkovich, A. Fournier, T. Klabunde, *J. Med. Chem*. **2002**, *45*, 1799 – 1805.
[47] W. L. Jorgensen, *Science* **1991**, *254*, 954 – 955.
[48] R. E. Babine, S. L. Bender, *Chem. Rev.* **1997**, *97*, 1359 – 1472.
[49] D. P. Marriott, I. G. Dougall, P. Meghani, Y.-J. Liu, D. R. Flower, *J. Med. Chem.* **1999**, *42*, 3210 – 3216.
[50] J. Alfaro-Lopez, T. Okayama, K. Hosohata, P. Davis, F. Porreca, H. I. Yamamura, V. J. Hruby, *J. Med. Chem.* **1999**, *42*, 5359 – 5368.
[51] K. Alexopoulos, D. Panagiotopoulus, T. Mavromoustakos, P. Fatseas, M. C. Paredes-Carbajal, D. Mascher, S. Mihailescu, J. Matsoukas, *J. Med. Chem*. **2001**, *44*, 328 – 339.
[52] P. Huang, G. H. Loew, H. Funamizu, M. Mimura, N. Ishiyama, M. Hayashida, T. Okuno, O. Shimada, A. Okuyama, S. Ikegami, J. Nakano, K. Inoguchi, *J. Med. Chem*. **2001**, *44*, 4082 – 4091.
[53] R. Hirschmann, K. C. Nicolaou, S. Pietranico, E. M. Leahy, J. Salvino, B. Arison, *J. Am. Chem. Soc*. **1993**, *115*, 12550 – 12568.
[54] R. M. Freidinger, *Curr. Opin. Chem. Biol*. **1999**, *3*, 395 – 406.
[55] N. R. A. Beeley, *Drug Discovery Today* **2000**, *5*, 354 – 363.
[56] M. Gurrath, *Curr. Med. Chem*. **2001**, *5*, 1605 – 1648.
[57] R. D. Cramer, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc*. **1988**, *110*, 5959 – 5967.
[58] K. H. Kim, G. Greco, E. Novellino, *Perspect. Drug Discovery Des.* **1998**, *12/13/14*, 257 – 315.
[59] H. Laning, W. Utz, P. Gmeiner, *J. Med. Chem*. **2001**, *44*, 1151 – 1157.
[60] B. Hoffmann, S. J. Cho, W. Zheng, S. Wyrick, D. E. Nichols, R. B. Mailman, A. Tropsha, *J. Med. Chem*. **1999**, *42*, 3217 – 3326.
[61] G. B. McGaughey, R. E. Mewshaw, *Bioorg. Med. Chem*. **1999**, *7*, 2453 – 2456.
[62] P. Gaillard, P.-A. Carrupt, B. Testa, P. Schambel, *J. Med. Chem.* **1996**, *39*, 126 – 134.
[63] M. L. Lopez-Rodriguez, B. Benhamu, A. Viso, M. J. Morcillo, M. Murcia, L. Orensanz, M. J. Alfaro, M. I. Martin, *Biorg. Med. Chem.* **1999**, *7*, 2271 – 2281.
[64] M. Lopez-Rodriguez, M. Murcia, B. Benhamu, A. Viso, M. Campillo, L. Pardo, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2807 – 2811.
[65] Q. Chen, C. Wu, D. Maxwell, G. A. Krudy, R. A. F. Dixon, T. J. You, *Quant. Struct.-Act. Relat.* **1999**, *18*, 124 – 133.
[66] S. M. Siddiqi, R. A. Pearlstein, L. H. Sanders, K. A. Jacobsen, *Bioorg. Med. Chem*. **1995**, *3*, 1331 – 1343.
[67] J. M. Rieger, M. L. Brown, G. W. Sullivan, J. Linden, T. L. Macdonald, *J. Med. Chem.* **2001**, *44*, 531 – 539.
[68] M. L. López-Rodriguez, M. L. Rosado, B. Benhamú, M. J. Morcillo, E. Fernández, K.-J. Schaper, *J. Med. Chem*. **1997**, *40*, 1648 – 1656.
[69] E. Ravina, J. Negreira, J. Cid, C. F. Masaguer, E. Rosa, M. E. Rivas, J. A. Fontenla, M. I. Loza, H. Tristan, M. I. Cadavid, F. Sanz, E. Lozoya, A. Carotti, A. Carrieri, *J. Med. Chem*. **1999**, *42*, 2774 – 2797.
[70] I. T. Forbes, S. Dabbs, D. M. Duckworth, P. Ham, G. E. Jones, F. D. King, D. V. Saunders, F. E. Blaney, C. B. Naylor, G. S. Baxter, T. P. Blackburn, G. A. Kennett, M. D. Wood, *J. Med. Chem*. **1996**, *39*, 4966 – 4977.
[71] 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.
[72] P. C. Astles, C. Brealey, T. J. Brown, V. Facchini, C. Handscombe, N. V. Harris, C. McCarthy, I. M. McLay, B. Porter, A. G. Roach, C. Sargent, C. Smith, R. J. Walsh, *J. Med. Chem*. **1998**, *41*, 2732 – 2744.
[73] J. Drews, *Drug Discovery Today* **1998**, *3*, 491 – 494.
[74] G. Wess, M. Urmann, B. Sickenberger, *Angew. Chem.* **2001**, *113*, 3443 – 3453; *Angew. Chem. Int. Ed*. **2001**, *40*, 3341 – 3350.
[75] B. R. Beno, J. S. Mason, *Drug Discovery Today* **2001**, *6*, 251 – 258.
[76] J. Bajorath, *Curr. Opin. Drug Discovery Dev.* **2002**, *3*, 24 – 28.
[77] J. S. Mason, I. Morize, P. R. Menard, D. L. Cheney, C. Hulme, R. F. Labaudiniere, *J. Med. Chem.* **1999**, *42*, 3251 – 3264.
[78] E. K. Bradley, P. Beroza, J. E. Penzotti, P. D. J. Grootenhuis, D. C. Spellmeyer, J. L. Miller, *J. Med. Chem.* **2000**, *43*, 2770 – 2774.
[79] R. S. Pearlman, K. M. Smith, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 28 – 35.
[80] J. Saunders, unpublished results.
[81] T. Klabunde, unpublished results.
[82] 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. L. Chang, V. J. Lotti, D. J. Cerino, T. B. Chen, P. J. Kling, K. A. Kunkel, J. P. Springer, J. Hirshfield, *J. Med. Chem*. **1988**, *31*, 2235 – 2246.
[83] X. Q. Lewell, D. B. Judd, S. P. Watson, M. M. Hann, *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 511 – 522.
[84] T. Klabunde, unpublished results.
[85] G. W. Bemis, M. A. Murcko, *J. Med. Chem*. **1996**, *39*, 2887 – 2893.
[86] B. R. Neustadt, E. M. Smith, N. Lindo, T. Nechuta, A. Bronnenkant, A. Wu, L. Armstrong, C. Kumar, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2395 – 2398.
[87] M. F. Gordeev, D. V. Patel, B. P. England, S. Jonnalagadda, J. D. Combs, E. M. Gordon, *Bioorg. Med. Chem*. **1998**, *6*, 883 – 889.
[88] A. A. Patchett, R. P. Nargund, *Annu. Rep. Med. Chem*. **2000**, *35*, 289 – 298.
[89] L. A. Thompson, J. A. Ellman, *Chem. Rev.* **1996**, *96*, 555 – 600.
[90] R. E. Dolle, *Mol. Diversity* **1998**, *3*, 199 – 233.
[91] S. C. Berk, S. P. Rohrer, S. J. Degrado, E. T. Birzin, R. T. Mosley, S. M. Hutchins, A. Pasternak, J. M. Schaeffer, D. J. Underwood, K. T. Chapman, *J. Comb. Chem*. **1999**, *1*, 388 – 396.
[92] T. Kennedy, *Drug Discovery Today* **1997**, *2*, 436 – 444.
[93] J. Hodgson, *Nat. Biotechnol.* **2001**, *19*, 722 – 726.
[94] H. Matter, K.-H. Baringhaus, T. Naumann, T. Klabunde, B. Pirard, *Comb. Chem. High Throughput Screening* **2001**, *4*, 453 – 475.
[95] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **1997**, *23*, 3 – 25.
[96] W. P. Walters, M. A. Murcko, *Adv. Drug Delivery. Rev.* **2002**, *54*, 255 – 271.
[97] W. J. Egan, G. Lauri, *Adv. Drug Delivery. Rev.* **2002**, *54*, 273 – 289.
[98] U. Norinder, M. Haeberlein, *Adv. Drug Delivery. Rev.* **2002**, *54*, 291 – 313.
[99] D. E. Clark, S. D. Pickett, *Drug Discovery Today* **2000**, *5*, 49 – 58.
[100] E. Jacoby, *Quant. Struct.-Act. Relat*. **2001**, *20*, 115 – 123.
[101] K. Lundstrom, *Current Drug Discovery* **2002**, 5, 29 – 33.
[102] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.
[103] E. A. Merritt, D. J. Bacon, *Methods Enzymol.* **1997**, *277*, 505 – 524.
[104] N. Murugesan, Z. Gu, P. D. Stein, S. Bisaha, S. Spergel, R. Girotra, V. G. Lee, J. Lloyd, R. N. Misra, J. Schmidt, A. Mathur, L. Stratton, Y. F. Kelly, E. Bird, T. Waldron, E. C.-K. Liu, R. Zhang, H. Lee, R. Serafino, B. Abboa-Offei, P. Mathers, M. Giancarli, A. A. Seymour, M. L. Webb, S. Moreland, J. C. Barrish, J. T. Hunt, *J. Med. Chem.* **1998**, *41*, 5198 – 5218.
[105] WO 9964002
[106] A. Pasternack, Y. Pan, D. Marino, P. E. Sanderson, R. Mosley, S. P. Rohrer, E. T. Birzin, S.-E. W. Huskey, T. Jacks, K. D. Schleim, K. Cheng, J. M. Schaeffer, A. A. Patchett, L. Yang, *Bioorg. Med. Chem. Lett*. **1999**, *9*, 491 – 496.
[107] A. A. Patchett, R. P. Nargund, J. R. Tata, M.-H. Chen, K. J. Barakat, D. B. R. Johnston, K. Cheng, W. W.-S. Chan, B. Butler, G. Hickey, T. Jacks, K. Schleim, S.-S. Pong, L.-Y. P. Chuang, H. Y. Chen, E. Frazier, K. H. Leung, S.-H. L. Chiu, R. G. Smith, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7001 – 7005.
[108] B. K. Trivedi, J. K. Padia, A. Holmes, S. Rose, D. S. Wright, J. P. Hinton, M. C. Pritchard, J. M. Eden, C. Kneen, L. Webdale, N. Suman-Chauhan, P. Boden, L. Singh, M. J. Field, D. Hill, *J. Med. Chem.* **1998**, *41*, 38 – 45.
[109] H. Qi, S. K. Shah, M. A. Cascieri, S. J. Sadowski, M. MacCoss, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2259 – 2262.
[110] J. M. Eden, M. D. Hall, M. Higginbottom, D. C. Horwell, W. Howson, J. Hughes, R. E. Jordan, R. A. Lewthwaite, K. Martin, A. T. McKnight, J. C. O'Toole, R. D. Pinnock, M. C. Pritchard, N. Suman-Chauhan, S. C. Williams, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2617 – 2622.
[111] V. Ashwood, V. Brownhill, M. Higginbottom, D. C. Horwell, J. Hughes, R. A. Lewthwaite, A. T. McKnight, R. D. Pinnock, M. C. Pritchard, N. Suman-Chauhan, C. Webb, S. C. Williams, *Bioorg. Med. Chem. Lett*. **1998**, *8*, 2589 – 2594.

Received: May 21, 2002 [A425]