A Chemogenomic Analysis of the Transmembrane Binding Cavity of Human G-Protein-Coupled Receptors

Jean-Sebastien Surgand, Jordi Rodrigo, Esther Kellenberger, and Didier Rognan*
Bioinformatics of the Drug, CNRS UMR7081, Illkirch, France

ABSTRACT **The amino acid sequences of 369 human nonolfactory G-protein-coupled receptors (GPCRs) have been aligned at the seven transmembrane domain (TM) and used to extract the nature of 30 critical residues supposed—from the X-ray structure of bovine rhodopsin bound to retinal—to line the TM binding cavity of ground-state receptors. Interestingly, the clustering of human GPCRs from these 30 residues mirrors the recently described phylogenetic tree of full-sequence human GPCRs (Fredriksson et al., Mol Pharmacol 2003;63:1256–1272) with few exceptions. A TM cavity could be found for all investigated GPCRs with physicochemical properties matching that of their cognate ligands. The current approach allows a very fast comparison of most human GPCRs from the focused perspective of the predicted TM cavity and permits to easily detect key residues that drive ligand selectivity or promiscuity. Proteins 2006;62:509–538.**

Key words: amino acid sequences; human G-protein-coupled receptors; transmembrane domain; bovine rhodopsin

INTRODUCTION

G protein-coupled receptors (GPCRs) constitute a large superfamily of very heterogeneous membrane receptors characterized by a typical heptahelical membrane-spanning fold usually described as a seven-transmembrane (TM) domain.[1,2] A striking feature of this protein superfamily is the tremendous chemical diversity of possible ligands including light, small molecular-weight ions (e.g., glutamate, Ca^{2+}), biogenic amines (e.g., dopamine, serotonin), nucleosides and nucleotides (e.g., adenosine, adenosine triphosphate), peptide and protein hormones (e.g., chemokines, glucagon), lipids and eicosanoids (e.g., sphingolipids, prostaglandins).[3] Activation of GPCRs upon ligand binding induces a conformational change of the receptor, thereby triggering a specific interaction with intracellular G proteins and subsequent activation/inhibition of secondary messengers.[4] Because of the ubiquitous distribution of GPCRs at the surface of many cells, these receptors are regulating a wide array of physiological and pathological processes. As a consequence, GPCRs are particularly attractive targets for therapeutic intervention. Hence, about 30% of top-selling drugs modulate the activity of this family of receptors.[3] Until now, few GPCRs (ca. 40) have been targeted by existing drugs. Analyzing human genomic sequences predicts the existence of about 400 nonolfactory GPCRs[5] and opens a new avenue for drug discovery, especially with respect to the 100 orphan receptors for which even the endogenous ligand still has not been characterized.[6]

Historically, GPCRs have been classified according to the chemical nature of their ligands, their specificity for known agonists/antagonists and the pharmacology associated with their activation/inhibition. The knowledge of the amino acid sequence of cloned GPCRs widened the classification to more receptors and resulted in a standard classification into three different families of mammalian GPCRs (from A to C, or I to III) depending on the alignability of the corresponding amino acid sequences.[7,8] Because of their typical organization in seven membrane-spanning α-helices, discriminating GPCRs and more generally 7-TM receptors from other genomic targets is rather straightforward.[9] Classifying GPCRs into subfamilies has been addressed by several methods including simple distance-based neighbor joining,[10] support vector machines,[11] hidden Markov models,[12] amino acid fingerprints,[13] covariant-discriminant analysis of amino acid composition,[14] and alignment-independent extraction of principal chemical properties of amino acid sequences.[15] However, previous clustering analyses, despite methodological merits, did not address the full dataset of human GPCRs. Recently, an exhaustive phylogenetic analysis of 342 monofunctional nonolfactory GPCRs[16] led to a revised classification consisting of five main families named glutamate (G family), rhodopsin (R), adhesion (A), frizzled/taste2 (F), and secretin (S). The classification can be used to search for similarities/dissimilarities among selected receptors. This information is crucial in drug discovery for many reasons: (1) it can be used to prioritize binding studies of a given ligand to GPCRs in order to address the ligand

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: the French Ministry of Research and Technology; Grant sponsor: the European Union (Marie Curie fellowship); Grant number: HPMF-CT-2002-02141; Grant sponsor: the Centre National de la Recherche Scientifique (CNRS, Programme Protéomique et génie des protéines).

*Correspondence to: Dr. Didier Rognan, CNRS UMR7081, 74 route du Rhin, B.P.24, F-67400 Illkirch, France. E-mail: didier.rognan@pharma.u-strasbg.fr

Received 14 April 2005; Revised 30 June 2005; Accepted 8 August 2005

Published online 17 November 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20768

selectivity as early as possible. (2) It helps to locate poorly studied or orphan targets in the GPCR universe by predicting the druggability of selected GPCRs in the light of known data on already investigated receptors. However, establishing a direct link between such phylogenetic trees and structure-based drug design is not straightforward. It requires an exact alignment of amino acid sequences, an accurate identification of amino acids lining the binding cavity and a 3-D model for every GPCR. Looking at similarities or differences at the binding sites would provide information that can be easily translated into ligand structures. Up to now, this chemogenomic approach has only been applied to a small subset of GPCRs[17–22] for which enough experimental information is available.

In the current study, we present an easy and straightforward classification scheme based on the alignment of 30 critical GPCR positions supposed to line the TM-binding cavity for inverse agonists or antagonists. The analysis, applied to a set of 369 nonredundant nonolfactory human GPCRs, provides a binding site-driven phylogenetic tree, as well as precise 2D and 3D structural determinants for each cluster that can be translated into structure-based ligand design.

MATERIALS AND METHODS

Input Data

Receptor sequences were extracted from the UniProt[23] release 3.1 and the NCBI Entrez Protein database[24] and parsed in plain text format, and further in XML format. First, only human proteins were selected by the field name "OS" (plain text version) or "organism" (XML version) containing the value "Human." Only GPCRs were retained, by checking the field "KW" or "keyword" with the value "G protein-coupled receptor." Olfactory receptors were then suppressed from the receptors list by checking the same field with the value "Olfaction." Last, receptors fragments were removed from the dataset, by looking whether the case-independent word "fragment" appeared in the description field ("DE" or "protein name"). Throughout this study, the UniProt entry name has been used to describe each receptor. Therefore, the entry name described herein does not necessarily map the gene name coding for the corresponding receptor, especially for orphan receptors whose official gene's name usually begins with GPR (e.g., GPR110). In the UniProt nomenclature, entries registered in the Swiss-Prot database are given a four or five character identifier: GPRx ($x \leq 9$), GPRxx ($10 \leq xx \leq 99$) or GPxxx ($xxx \geq 100$).

Amino Acid Alignment of the 7-TMs

The GPCRmod program was used to align the 7-TMs of selected human GPCRs, as recently described.[25] Briefly, GPCRmod first predicts the rough location of TMs using the TMHMM algorithm[9] and then, in each isolated TM, looks for family and TM-specific amino acid patterns.[13] Upon family detection (rhodopsin, secretin, glutamate, frizzled), the query TM is then aligned with that of the corresponding family template (rhodopsin: bovine rhodopsin, secretin: human calcitonin receptor, glutamate: human calcium-sensing receptor, frizzled: human frizzled type 1) assuming TM lengths similar to that depicted in the X-ray structure of bovine rhodopsin.[26] No attempts were made to align either N- or C-terminal domains or intra- and extracellular loops. A filtering procedure was then applied to suppress duplicate 7-TM sequences, usually keeping the Swiss-Prot entry and removing the TrEMBL duplicates. The final GPCR dataset contained 369 entries when the manuscript was prepared (see Annex 1 in Supplementary Data).

Clustering Human GPCRS From 30 Discontinuous Amino Acids

The selection of the 30 amino acids was performed as follow: in the X-ray structure of the bovine rhodopsin–retinal complex (pdb entry 1f88),[26] we defined the receptor cavity as the collection of 81 residues enclosed within a 10-Å sphere centered on the bound ligand. The solvent-accessible surface was then computed for the 81 selected amino acids. We finally picked out the 30 residues (Fig. 1) that present at least 25% of their surface accessible to a putative ligand and whose side chain is pointing inward the 7-TM bundle. Each GPCR was then described by an ungapped sequence of 30 residues. The resulting 369 30-amino-acid-long sequences were hierarchically clustered using the UPGMA method[27] to yield a rooted phylogenetic tree. Protein pair-wise distances were measured by computing sequence identity. The statistical significance of the obtained binary tree was assessed by bootstrapping. The starting dataset (369 sequences of 30 residues) was bootstrapped 1,000 times by randomly selecting the sequences input order and allowing substituting not more than 10% of the starting dataset by randomly selected entries. The 1,000 bootstrapped datasets were clustered as previously described to yield 1,000 binary trees, which after normalization were used to derive a consensus tree using the program CONSENSE, from the PHYLIP package (http://cmgm.stanford.edu/phylip/). Trees were visualized with the NJplot program.[28]

3D Model Building and Cavity Detection

A 3D model of each GPCR in its ground state was obtained using the GPCRmod program as recently described.[25] Briefly, the GPCRmod alignment of the 7 TMs was converted into a 3D model using eight backbone templates originating from previously described homology models and two rotamer libraries for side chain positioning.[25] For each model, all possible channels and cavities[29] were detected using the MOLCAD module (Multi channel function) of the SYBYL package.[30] Connolly surfaces and channels were calculated using a 1.4-Å radius probe and a dot density of six points/area. The biggest cavity was finally selected and its surface and volume computed.

RESULTS AND DISCUSSION

Classifying GPCRs According to the Discontinuous Sequence of Amino Acids that Form the TM Ligand Binding Site

The current study proposes a straightforward classification of most druggable human GPCRs from the point of

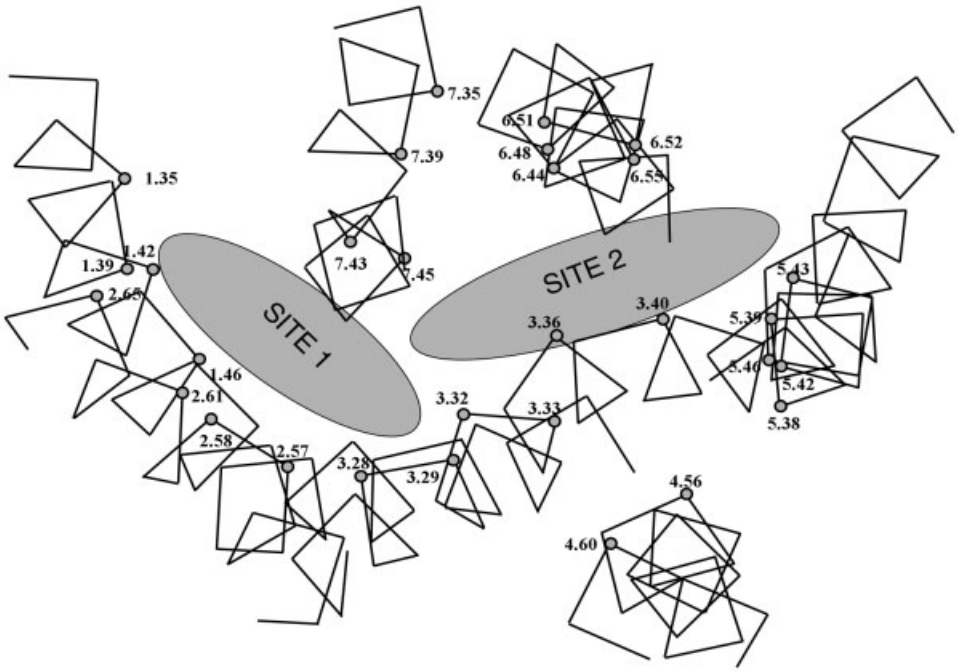


Fig. 1. Top view of 30 critical positions, labeled at the Cα position according to the Ballesteros numbering,[32] supposed to possibly line the generic TM binding cavity of GPCRs. Areas defining prototypical subsites 1 and 2 (see text) are circled.

view of a putative ligand. To achieve this task, 30 critical amino acids disseminated on the 7-TM domain were extracted for each receptor (Fig. 1).

Our clustering approach implies two assumptions: (1) the overall fold of the 7-TM domain around the binding crevice has been conserved along evolution; (2) critical hotspots spread over the 7-TM domain repeatedly account for ligand binding. Although very few structural information is available for the three most important GPCR classes (Classes A, B, C), numerous experimental data do provide evidence in favor of strong structural similarities among many GPCRs: (1) most of the residues known to affect small molecular-weight ligand binding to unrelated GPCRs are included in the selected 30 residues, suggesting a common architecture of the TM pocket, (2) many known ligands are promiscuous for even unrelated GPCRs[3] and are usually anchored through so-called privileged structures to common features of different GPCRs.[31] Of course, class B and class C GPCRs exhibit an additional orthosteric site located outside the 7-TM bundle. Therefore, conclusions drawn in the present paper only apply to the 7-TM binding site.

TM-Cavity Based Clustering of 369 Human GPCRs Recalls Classification Based on Full TM Sequences

Almost all nonredundant druggable human GPCRs have been analyzed in the present study. Forty-six putative GPCRs (see Annex 2 in supplementary data) were rejected from our dataset for three main reasons: (1) seven TMs could be detected but no satisfactory alignment with other GPCRs could be found (e.g., Taste 2 receptors), (2) less/more than seven TMs could be detected (e.g., GPR172A), (3) no TM domain at all could be detected (e.g., CRCP_HUMAN).

The TM cavity-derived phylogenetic tree obtained from 369 human nonredundant nonolfactory GPCRs is shown in Figure 2. A total of 22 clusters were defined in order to encompass the maximum number of related entries within a branch characterized by the highest possible bootstrap value. Thirty-four out of 369 entries could not be assigned to any of the existing 22 clusters (Annex 1). Instead of generating very small-sized additional clusters, we prefer to define them as singletons unrelated to any of the current clusters. The resulting tree is very similar to the most complete phylogenetic tree (GRAFS classification) known to date[16] although the latter has been obtained from full TM sequences. In both classifications, GPCRs of the Frizzled, Glutamate, Secretin, and Adhesion families cluster in well separated groups. The large Rhodopsin family is split into 18 different clusters. Remarkably, all known GPCR subfamilies (e.g., receptors for biogenic amines, purines, and chemokines) are reproduced with high bootstrap support. The five main families (Glutamate, Rhodopsin, Adhesion, Frizzled, Secretin) reported in the GRAFS classification[16] are recovered with no overlaps between the corresponding clusters. The single exception is a rhodopsin-like GPCR (GPR88) which clusters with class C GPCRs. Interestingly, receptors for which the orthosteric binding site is not located in the TM domain (Adhesion, Secretin, and Glutamate families) are nevertheless grouped into homogeneous clusters. Furthermore, we have detected a putative binding site in the TM cavity of each of these

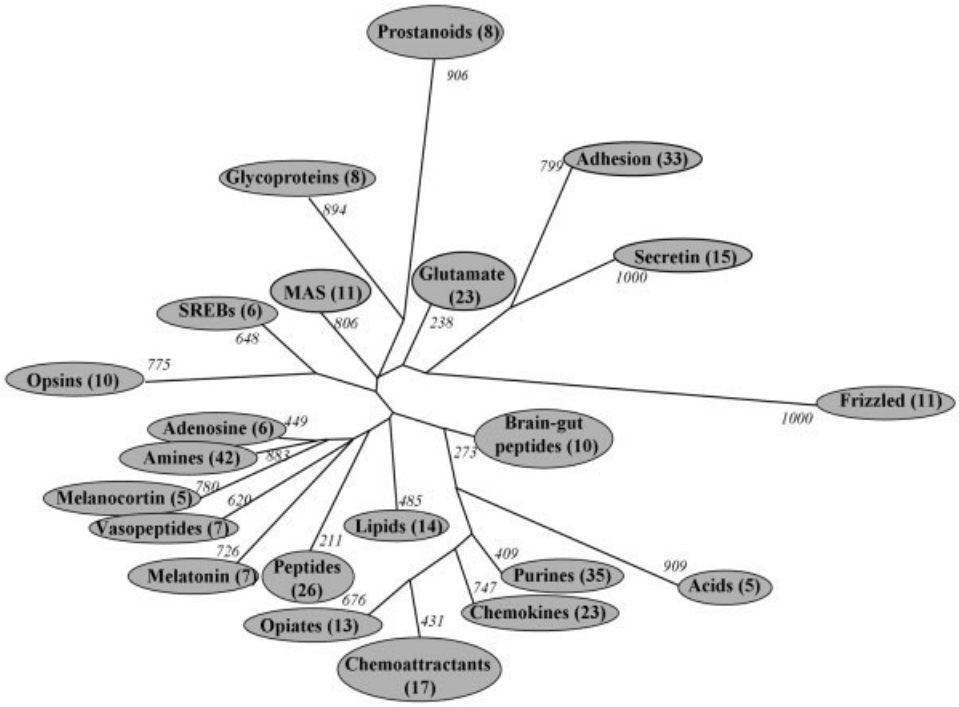


Fig. 2. TM cavity-derived phylogenetic tree for 369 human GPCRs. Numbers in parentheses indicate the number of entries in each cluster. Numbers in italic represent bootstrap values to assess the statistical significance of the tree. Receptors classified as singletons (see text) are not displayed here for sake of clarity.

receptors. We assume that a strong evolutionary pressure has been applied to the 7-TM domain of class B and C GPCRs to maintain a TM cavity whose function, besides triggering G-Protein coupling, could be to define a putative binding site for endogenous inverse agonists or allosteric regulators. To ensure the statistical significance of class B and C GPCRs clustering, a new tree was computed by selecting the same number of amino acids (30), the same proportion of residues per TM, but from amino acids whose side chains do not point inward the TM cavity. No homogeneous clusters could be found in that case suggesting that the herein selected 30 residues do have a clear functional role.

In the next sections, the physicochemical properties of the predicted TM cavity will be described for each cluster. Where possible, relationships were established between the TM cavity and their respective ligands. For sake of clarity, GPCRs are identified by their UniProt entry names (see Annex 1 for an exhaustive description of all entries) and TM residues by the Ballesteros numbering.[32] As expected from the reduced set of templates used for comparative modelling, the modeled TM cavity of most GPCRs share significant similarities, notably the presence of two subsites, a first one delimited by TMs 1, 2, 3, and 7 (site 1, Fig. 1), and a second one delimited by TMs 3, 4, 5, and 6 (site 2, Fig. 1). Throughout this study, subsites 1 and 2 will refer to this definition.

The Frizzled Receptor Cluster (11)

The frizzled receptor cluster regroups 11 receptors (10 frizzled receptors, smoothened) basically involved in cell growth and proliferation [Fig. 3(A)]. Known ligands for the frizzled receptors are Wnt proteins[33] and are supposed to bind to the N-terminal cystein-rich domain of frizzled receptors. The cluster is identical to the frizzled family in the GRAFS classification. The alignment shown in Figure 3(B) indicates six fully conserved residues (Phe$^{3.29}$, Tyr$^{3.33}$, Phe$^{5.38}$, Val$^{5.39}$, Pro$^{5.42}$, and Ile$^{7.45}$). The predicted 3D structure is characterized by a large TM cavity (about 1,500 Å^3) extending very deep toward the intracellular side between TMs 5 and 6. Interestingly, TMs 1–4 greatly contribute to the cavity through hydrophilic and charged residues whereas TMs 5–7 bring small aliphatic amino acids [Fig. 3(B)]. Although TM-binding endogenous ligands of this GPCR family are still elusive, the herein described topology of the TM cavity perfectly matches the chemotype of known smoothened synthetic antagonists[34,35] which consists in bulky hydrophobic molecules with a cationic head.

The Secretin Receptor Cluster (15)

The secretin receptor cluster is made of 15 GPCRs belonging to the class of secretin-like receptors [Fig. 4(A)]. This very homogeneous cluster has recently been proposed to form a homogeneous GPCR family.[16] Ligands of GPCRs belonging to this cluster are peptide hormones supposed to bind to the long N-terminal extracellular tail via their C-terminal residues and to contact the TM cavity through their N-terminal residues.[36,37] However, a typical GPCR from this cluster (e.g., CALCR) presents a narrow and hydrophobic TM cavity (ca. 1,000 Å^3) with a set of residues (Ile/Val$^{1.39}$, Phe/Leu$^{2.58}$, Asp$^{2.61}$, Ile/Leu/Val$^{3.36}$, Phe/

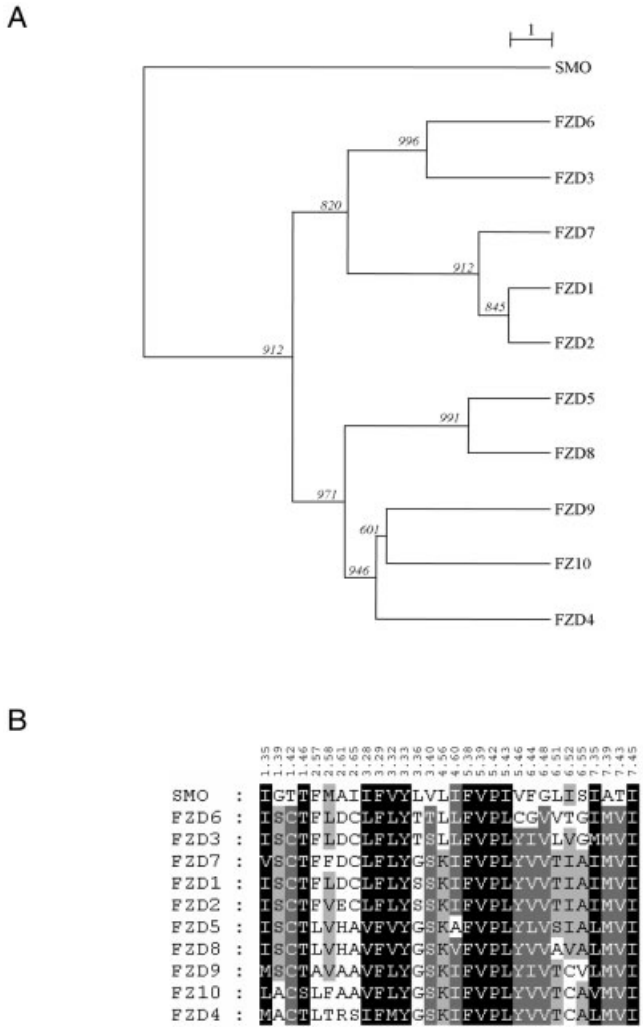


Fig. 3. The frizzled receptor cluster: (**A**) Phylogenetic tree: Branch lengths are proportional to the distance between two receptors. Number in italic represent bootstrap values. (**B**) Cavity-based alignment of 30 discontinuous positions colored according to the degree of similarity (white foreground/black background, 100%; white foreground/grey background, $>$ 80%, black foreground/grey background, $>$ 60%). Receptors are labeled according to their UniProt entry names and amino acid positions identified by the Ballesteros numering scheme.[32] FZ10: Frizzled-10; FZD1: Frizzled-1; FZD2: Frizzled-2; FZD3: Frizzled-3; FZD4: Frizzled 4; FZD5: Frizzled-5; FZD6: Frizzled-6; FZD7: Frizzled-7; FZD8: Frizzled-8; FZD9: Frizzled-9; SMO: Smoothened.

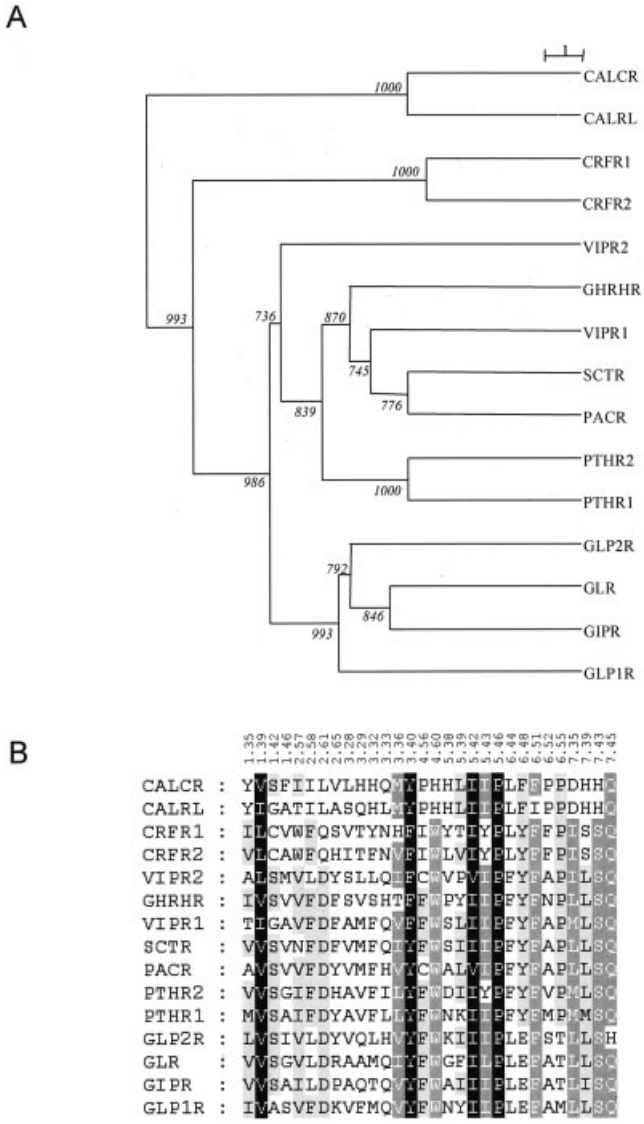
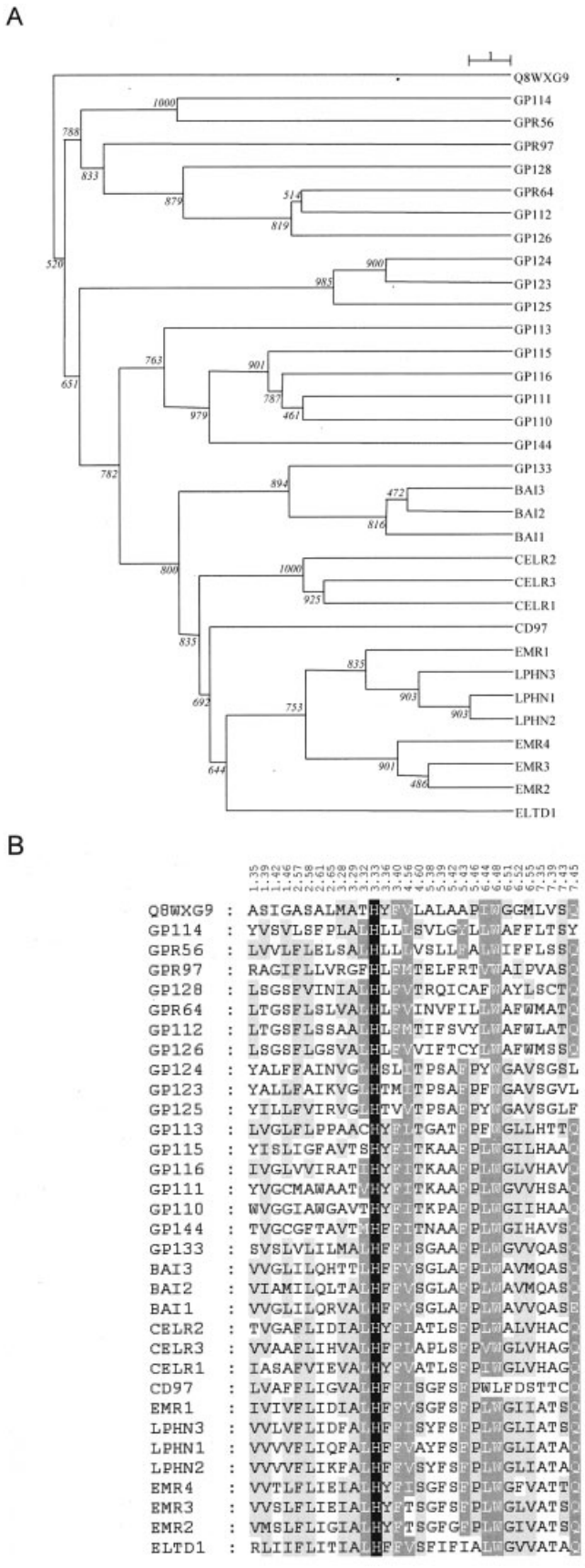


Fig. 4. The secretin receptor cluster: (**A**) Phylogenetic tree. (**B**) Cavity-based alignment. CALCR: Calcitonin receptor; CALRL: Calcitonin gene-related peptide type 1 receptor; CRFR1: Corticotropin releasing factor receptor; CRFR2: Corticotropin releasing factor receptor 2; GHRHR: Growth hormone-releasing hormone receptor; GIPR: Gastric inhibitory polypeptide receptor; GLP1R: Glucagon-like peptide 1 receptor; GLP2R: Glucagon-like peptide 2 receptor; GLR: Glucagon receptor; PACR: Pituitary adenylate cyclase activating polypeptide type I receptor; PTHR1: Parathyroid hormone receptor 1; PTHR2: Parathyroid hormone receptor 2; SCTR: Secretin receptor; VIPR1: Vasoactive intestinal polypeptide receptor 1; VIPR2: Vasoactive intestinal polypeptide receptor 2.

Tyr$^{3.40}$, Trp$^{4.60}$, Ile/Val$^{5.42}$, Ile$^{5.43}$, Pro$^{5.46}$, Phe$^{6.51}$, Ile/Leu/Met$^{7.35}$, Ser$^{7.43}$ and Gln$^{7.45}$) rather conserved throughout the cluster [Fig. 4(B)]. The TM cavity is predicted to be large enough to accommodate small molecular-weight ligands (notably though a salt bridge to Asp$^{2.61}$) that could inhibit the action of the natural hormone at the remote binding site.

It remains difficult to probe 3D models of this GPCR cluster because of the paucity of known nonpeptide antagonists. For the glucagon receptor (GLR) however, there are several small molecular weight antagonists available.[38–40] The GLR cavity presents a rather hydrophilic centre (Gln$^{3.33}$, Ser$^{7.43}$, Gln$^{7.45}$) surrounded by two hydrophobic subsites. Consensus positions between TMs 1 and 2 (Val$^{1.39}$, Ser$^{1.42}$, Gly$^{1.46}$, Val$^{2.57}$, Leu$^{2.58}$) as well as residues not taken into account by our clustering scheme (Phe$^{2.50}$, Ala$^{2.54}$) define a narrow and deep subsite 1. Subsite 2 is a large pocket formed by TMs 4, 5 and 6 (Tyr$^{3.40}$, Trp$^{4.60}$, Phe$^{5.39}$, Ile$^{5.42}$, Pro$^{5.46}$, and Phe$^{6.51}$). The TM cavity of a few class B GPCRs (CRFR1, GLR, CGRP, GLP1) is probably an allosteric binding site for nonpeptide antagonists that precludes for the activation of the receptor by endogeneous peptide agonists.[41] Interestingly, site-directed mutagenesis suggests that Phe184$^{2.50}$ and Tyr239$^{3.40}$ are



involved in the binding of the nonpeptide GLR antagonist L-168049 whose binding site does not overlap that of the endogenous hormone.[39] Point mutation of position 3.36 in the CRF_1 receptor has also been shown to affect the binding of a nonpeptide CRF_1 antagonist.[42] Despite some warnings about the druggability of the allosteric TM binding site, notably regarding its hydrophobicity,[41] it is very likely that the potential of this GPCR cluster has not been fully addressed yet.

The Adhesion Receptor Cluster (33)

Members of the adhesion receptor cluster [Fig. 5(A)] are related to the secretin receptor cluster as being part of the former secretin-like GPCRs, but are unambiguously clustered with high bootstrap values into a separate family, in agreement with the GRAFS classification.[16] These GPCRs exhibit specific repeats at the N-terminal extracellular domain (EGF, cadherin, mucin) likely to be involved in cell adhesion processes.[43,44] Many of these GPCRs are expressed at the cell surface as heterodimers consisting of a large N-terminal domain associated with the 7-TM domain. A unique feature of some of these GPCRs is the intracellular location of their ligands (e.g., CD55, chondroitin sulfate) that at first associate with the extracellular domain.[43]

The binding cavity of adhesion GPCRs resembles that of the former cluster of secretin-like hormone receptors. It is of limited volume (around 1,000 $Å^3$) and centered on a conserved histidine residue [$His^{3.33}$, Fig. 5(B)]. Side chains of conserved aromatic amino acids at 3.36 and 3.40 limit the width of the cavity center. Medium-sized amino acids of both TM1 and TM2 (Ala, Val, Leu, and Ile) participate to the formation of a deep hydrophobic subsite. Depending on position 5.43, the cavity extents towards TMs 5 and 6. In about a half of these receptors, a Phe at 5.43 restricts the depth of the cavity between TMs 5 and 6. Smaller residues are found at position 5.43 in the rest of adhesion-like GPCRs and permit the extension of the TM cavity towards TMs 5 and 6. By analogy to most rhodopsin-like GPCRs, an important Trp is often present at position 6.48. In rhodop-

Fig. 5. The adhesion receptor cluster: (**A**) Phylogenetic tree. (**B**) Cavity-based alignment. BAI1: Brain-specific angiogenesis inhibitor 1; BAI2: Brain-specific angiogenesis inhibitor 2; BAI3: Brain-specific angiogenesis inhibitor 3; CD97: Leucocyte antigen CD97; CELR1: Cadherin EGF LAG seven-pass G-type receptor 1; CELR2: Cadherin EGF LAG seven-pass G-type receptor 2; CELR3: Cadherin EGF LAG seven-pass G-type receptor 3; ELTD1: latrophilin and seven transmembrane domain containing protein 1; EMR1: Cell surface glycoprotein EMR1; EMR2: EGF-like module EMR2; EMR3: EGF-like module containing mucin-like hormone receptor-like 3; EMR4: EGF-like module containing mucin-like hormone receptor-like 4; GP110: G protein-coupled receptor 110; GP111: G protein-coupled receptor 111; GP112: G protein-coupled receptor 112; GP113: G protein-coupled receptor 113; GP114: G protein-coupled receptor 114; GP115: G protein-coupled receptor 115; GP116: G protein-coupled receptor 116; GP123: G protein-coupled receptor 123; GP124: Tumor endothelial marker 5; GP125: G protein-coupled receptor 125; GP126: G protein-coupled receptor 126; GP128: G protein-coupled receptor 128; GP133: G protein-coupled receptor 133; GP144: G protein-coupled receptor 144; GPR56: G protein-coupled receptor 56; GPR64: Epididymis-specific protein 6; GPR97: G protein-coupled receptor 97; LPHN1: Lectomedin-2; LPHN2: Lectomedin-1 beta; LPHN3: Lectomedin-3; Q8WXG9: Very large G protein-coupled receptor 1b.

sin-like GPCRs, the residues at position 6.48 of the TM cavity has been proposed to act as a molecular switch for locking the GPCR either in a ground state or in an activated form.[32] It is tempting to speculate the involvement of the conserved Trp$^{6.48}$ in the conformational changes occurring upon activation of this cluster of receptors as well.

No small molecular-weight ligands have ever been described for this GPCR cluster. Therefore, it is currently impossible to match the proposed cavity description to known ligand chemotypes.

The Glutamate Receptor Cluster (23)

The glutamate receptor cluster regroups members of the formerly known class C GPCRs. Class C RCPGs are characterized by a long N-terminal domain which folds into a bilobed venus-fly trap domain delimiting the orthosteric binding site of the endogenous ligand (GABA, glutamate, Ca^{2+}).[45] Their 7-TM domain is responsible for receptor activation and G protein coupling[46] and contains a cavity that is known to bind positive or negative allosteric modulators.[47,48] Bootstrap values clearly indicate the presence of three separated branches [Fig. 6(A)].

A typical 7-TM cavity of a GPCR from this cluster has a volume of ca. 1,500 Å^3 and can be viewed as two hydrophobic subsites linked by charged and/or hydrophilic residues at positions 3.28 and 3.29. Looking at the eight metabotropic glutamate receptors, the ligand-accessible residue at position 3.28 is a glutamine in MGR1 and MGR5 whereas it is an arginine in the six others receptors. This may account for the functional similarity of MGR1 and MGR5. Conserved hydrophobic residues in TM1 (position 1.42), TM2 (position 2.58), TM3 (position 3.32), and TM7 (positions 7.43 and 7.45) delimit subsite 1. Subsite 2 is lined by conserved aromatic residues in TM3 (position 3.40) and TM6 (positions 6.48, 6.51 and 6.55). Many of these amino acids have been experimentally found to map the binding site of both positive and negative allosteric modulators of GPCRs belonging to this cluster (MGR1, MGR2, MGR5, CASR).[47–50] Moreover, recent findings proved that the MGR5 receptor does not require the N-terminal venus fly-trap module for the recognition of allosteric modulators.[46] Hence, receptors of the current cluster probably comprise at least two binding sites; the extracellular binding site for endogenous ligands and the transmembrane binding site for allosteric modulators. It has been shown that the truncated MGR5 receptor which lacks the N-terminal extracellular domain, exhibits constitutive activity attributable to the 7-TM domain only.[46] Interestingly, Trp$^{6.48}$ is also found in class C GPCRs which provides support to the hypothesis that the overall fold of the 7-TM domain of most GPCRs in their ground states has been rather conserved during evolution.

The known chemotypes of noncompetitive agonists/antagonists for either metabotropic glutamate receptors (mGluR) or calcium-sensing receptor nicely fits the above-described cavity description. mGluR allosteric modulators usually bear two aromatic groups separated by an amide/sulfonamide moiety[39,47,50] that is likely to bind to one of

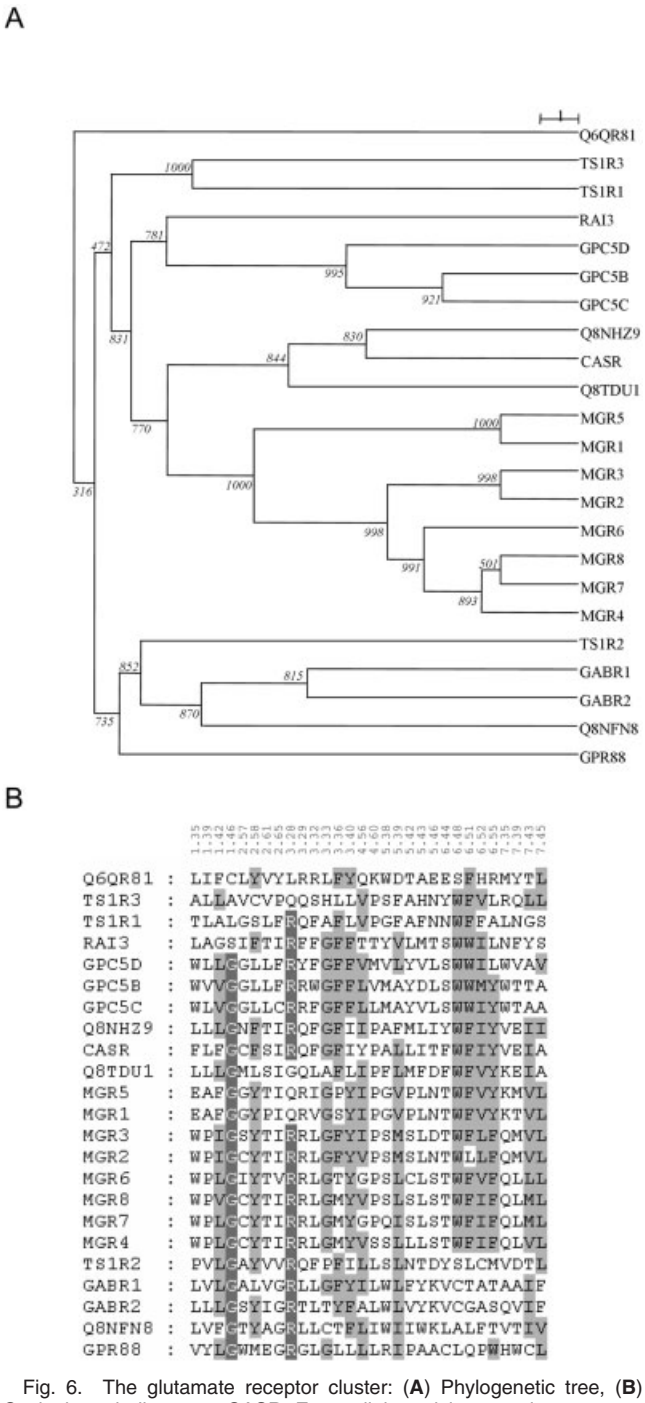


Fig. 6. The glutamate receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. CASR: Extracellular calcium-sensing receptor; GABR1: Gamma-aminobutyric acid type B receptor, subunit 1; GABR2; Gamma-aminobutyric acid type B receptor, subunit 2; GPC5B: A-69G12.1; GPC5C: RAIG-3; GPC5D: G protein-coupled receptor family C group 5 member D; GPR88: Striatum-specific G protein-coupled receptor 88; MGR1: Metabotropic glutamate receptor 1; MGR2: Metabotropic glutamate receptor 2; MGR3: Metabotropic glutamate receptor 3; MGR4: Metabotropic glutamate receptor 4; MGR5: Metabotropic glutamate receptor 5; MGR6: Metabotropic glutamate receptor 6; MGR7: Metabotropic glutamate receptor 7; MGR8: Metabotropic glutamate receptor 8; Q6QR81: G protein-coupled receptor 158; Q8NFN8: GABAB-related G protein-coupled receptor; Q8NHZ9: GPRC6A; Q8TDU1: Putative G protein-coupled receptor Q8TDU1; RAI3: Orphan G protein-coupling receptor PEIG-1; TS1R1: Taste receptor TAS1R1; TS1R2: Taste receptor TAS1R2; TS1R3: Taste Receptor TAS1R3

the three hydrophilic residues at the center of the cavity (positions 3.28, 3.29, and 7.39). In CaSR allosteric ligands, a positively charged secondary amine is anchored to Glu837$^{7.39}$and is surrounded by two aromatic moieties filling the hydrophobic subsites.[48,49]

By contrast to glutamate receptors, the GABA-B receptor is a heterodimer consisting in two subunits; GABR1 is responsible for ligand binding at the venus fly-trap domain whereas GABR2 participates in G protein coupling.[51] The 7-TM domain of GABR2 has recently been shown to bind a positive allosteric regulator of the GABA-B receptor.[52] The 7-TM binding cavity resembles that of the metabotropic glutamate receptors but is more hydrophobic and significantly smaller. Another striking difference is the presence of small hydrophilic amino acids at position 6.48 and 6.51 that could be anchoring points for small-sized allosteric regulators of the GABA-B receptor[53,54] although their binding site has not been precisely mapped up to now.

The above-cited data suggest that the 7-TM cavity of class C receptors has a functional role in triggering receptor activation, and can be viewed as an allosteric site for putative ligands. Whether this role may be linked to the existence of naturally occuring ligands targeting the allosteric subsite for class C GPCRS (as well as for members of the previously defined adhesion and secretin clusters) still has to be demonstrated.

The Prostanoids Receptor Cluster (8)

The prostanoids receptor cluster contains eight receptors classified in two main subgroups [Fig. 7(A)]. Prostanoids are cyclooxygenase metabolites and consist in C-20 unsaturated fatty acids. They comprise prostaglandins, prostacyclins, and thromboxanes, and exert a variety of actions, including the relaxation and contraction of various types of smooth muscles. This branch is separated in two parts according to their G protein coupling.[55,56] Strong homology between the four G_s-coupled prostaglandin receptors (PD2R, PE2R2, PE2R4, and PI2R) and the four G_q/G_i-coupled receptors (PE2R1, PE2R3, PF2R, and TA2R) suggests that these receptors evolved from an ancestral receptor. Receptors for prostanoids present a large TM cavity (1,840 Å^3 for PI2R) clearly composed of two separate hydrophobic subsites. Subsite 1 is delimited by conserved small or medium-sized residues at positions 1.46, 2.58, 2.61, and 7.43 [Fig. 7(B)], thereby extends relatively deep between TMs 1 and 2. Subsite 2 located between TMs 3, 5, and 6 is delimited by positions 3.29, 3.36, 5.38, 5.39, 5.42, 5.43, 6.51, and 6.55. This apolar subsite extends deep toward the intracellular side of the cavity between small and medium side chains at TMs 5 and 6. Main differences between the two prostanoid receptor subgroups occurs at two positions (6.48 and 7.35) of subsite 2, the latter being polar for the first subgroup and hydrophobic for the second one. Experimental evidences suggested that Arg$^{7.40}$, a conserved amino acid throughout prostanoid receptors but not taken into account to define the current cluster, interacts with the carboxylate moiety of prostanoids.[57] According to our 3D model of prostanoid receptor, the side chain of Arg$^{7.40}$ is accessible and takes part in the first

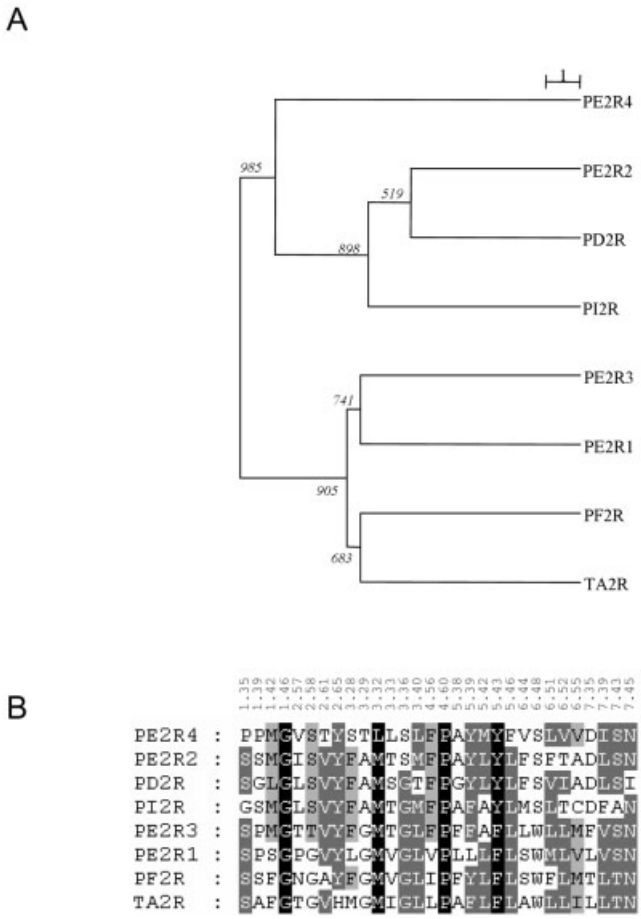


Fig. 7. The prostanoids receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment.PD2R: Prostaglandin D2 receptor; PE2R1: Prostaglandin E2 receptor, EP1 subtype; PE2R2: Prostaglandin E2 receptor, EP2 subtype; PE2R3: Prostaglandin E2 receptor, EP3 subtype; PE2R4: Prostaglandin E2 receptor, EP4 subtype; PF2R: Prostaglandin F2-alpha receptor; PI2R: Prostacyclin receptor; TA2R: Thromboxane A2 receptor.

subsite of the TM cavity. More generally, site-directed mutagenesis mapping of prostanoid binding sites[58–60] agrees well with the herein predicted TM cavities. The cyclopentane ring of prostanoids is proposed to interact with small aliphatic residues of TMs 1,2 and both alkyl chains are directed towards TMs 3 and 7 with an ionic bond between Arg$^{7.40}$ and the acidic moiety of the ligand.[59,60] The recent mapping of the prostacyclin receptor (PI2R) binding pocket[65] revealed four important anchoring residues: Tyr$^{2.65}$, Phe$^{3.28}$, Phe$^{7.39}$, Arg$^{7.40}$. Three of them belong to our subset of 30 positions. Synthetic agonists and antagonists of prostanoid receptors are likely to bind in a similar way because of their high structural similarity to endogenous prostanoids.[61,62]

The Glycoproteins Receptor Cluster (8)

The well-defined glycoproteins receptor cluster regroups eight members of the rhodopsin family [Fig. 8(A)]. All representatives are receptors for glycoprotein hormones. Their structure consists of a N-terminal domain containing several leucine-rich repeats (LRR), followed by the

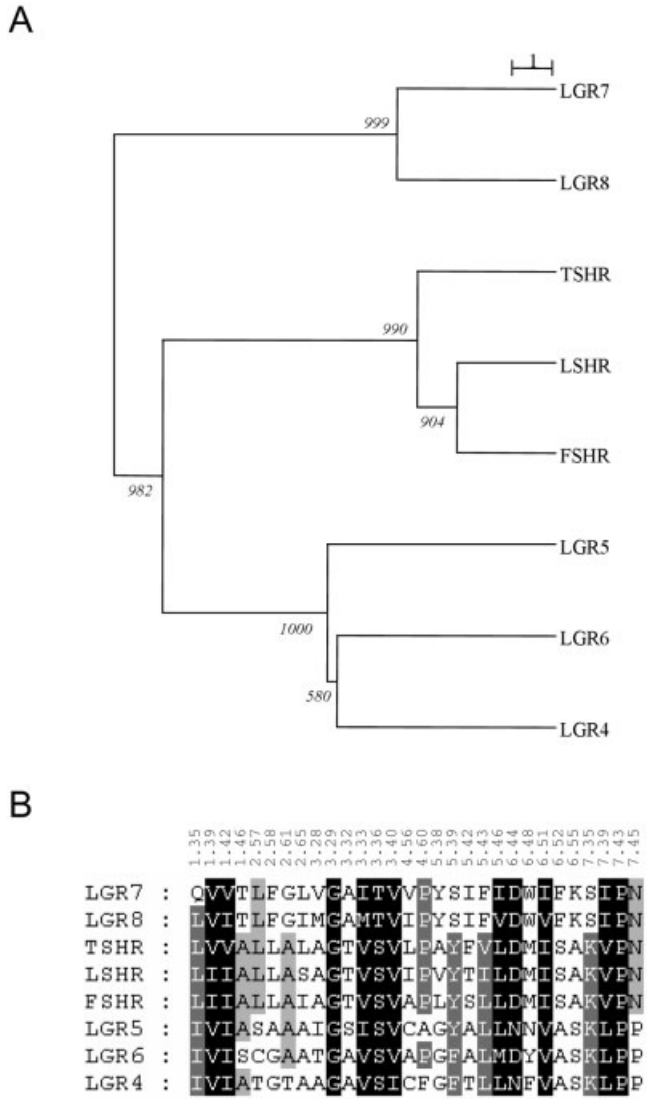


Fig. 8. The glycoproteins receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. FSHR: Follicle stimulating hormone receptor; LGR4: Leucine-rich repeat-containing G protein-coupled receptor 4; LGR5: Leucine-rich repeat-containing G protein-coupled receptor 5; LGR6: Leucine-rich repeat-containing G protein-coupled receptor 6; LGR7: Relaxin receptor 1; LGR8: Relaxin receptor 2; LSHR: Lutropin-choriogonadotropic hormone receptor; TSHR: Thyrotropin receptor.

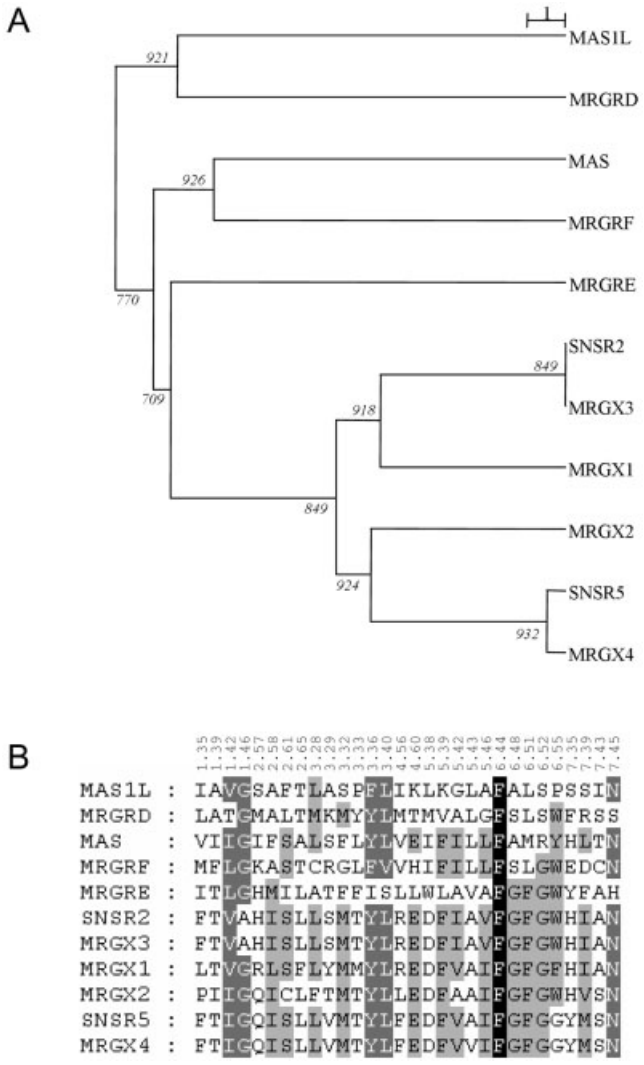


Fig. 9. The MAS-related receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. SNSR5: G protein-coupled receptor SNSR5; MAS: MAS proto-oncogene; MAS1L: Mas-related G protein-coupled receptor MRG; MRGRD: MrgD; MRGRE: MAS-related G protein-coupled receptor member E; MRGRF: Mas-related G protein-coupled receptor MRGF; MRGX1: G protein-coupled receptor MRGX1 (SNRSR3, SNSR4); MRGX2: G protein-coupled receptor MRGX2; MRGX3: MRGX3 G-protein-coupled receptor SNSR1: MRGX4 G protein-coupled receptor SNSR6; SNSR2: G protein-coupled receptor SNSR2.

7-TM domain.[63] The first domain is responsible for the binding of the natural hormone and the second one is involved in triggering receptor activation. Although the natural hormone binds to the N-terminal ectodomain, these GPCRs exhibit a clear TM cavity (ca. 1,230 Å^3 for FSHR) composed by two subsites (hydrophobic site between TMs 1, 2, and 7; hydrophilic site between TMs 3, 5, and 6) linked by an hydrophilic channel [Ser/Thr$^{3.36}$, Asn$^{7.45}$; Fig. 8(B)]. Several conserved polar side chains (Ser/Thr$^{3.36}$; Asp/Asn$^{6.44}$) could constitute excellent anchoring points for any ligand lying in the TM cavity.

The MAS-Related Receptor Clusters (11)

This group includes the MAS proto-oncogen and related receptors [Fig. 9(A)] expressed in specific subpopulations of sensory neurons that detect painful stimuli and may regulate nociceptive function and/or development.[64,65] Few data are available on these receptors. However, several endogenous signaling molecules have recently been reported for MAS-related (MRG) receptors. Angiotensin(1-7) is a ligand for the MAS receptor,[66] thereby counteracting effects of Angiotensin II. β-Alanine is a signaling molecule for MRGRD, a receptor known to module neuropathic pain.[67] Adenine and neuropeptide RF-amide receptors in rodents likely to be involved in nociception are close in sequence to human SNS/MRG receptors.[68,69] Last, MRGX2 has been shown to be a receptor for a new neuropeptide, corticostatin[70], which exerts a sleep regulation effect.

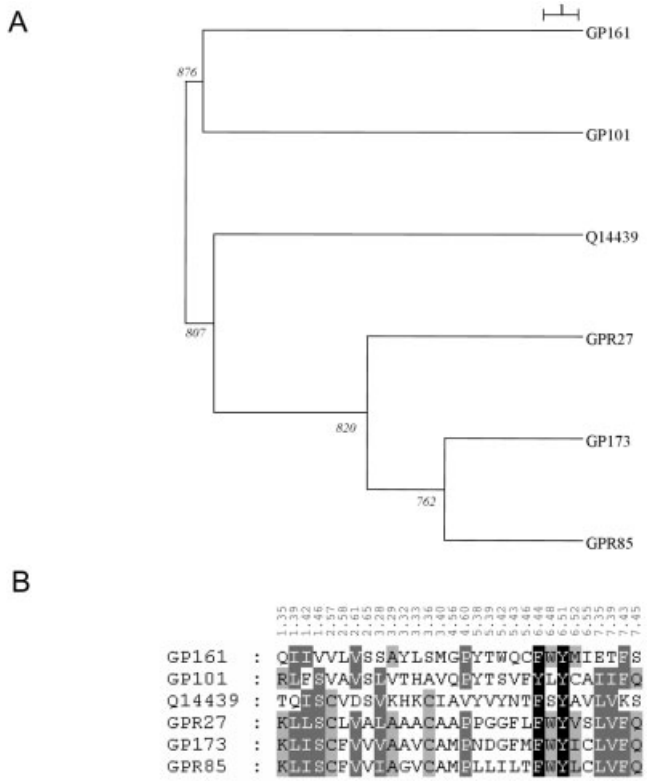


Fig. 10. The SRB receptor clusters: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. GP101: G protein-coupled receptor 101; GP161: G protein-coupled receptor RE2; GP173: G protein-coupled receptor 173; GPR27: G protein-coupled receptor 27; GPR85: G protein-coupled receptor GPR85; Q14439: G protein-coupled receptor Q14439.

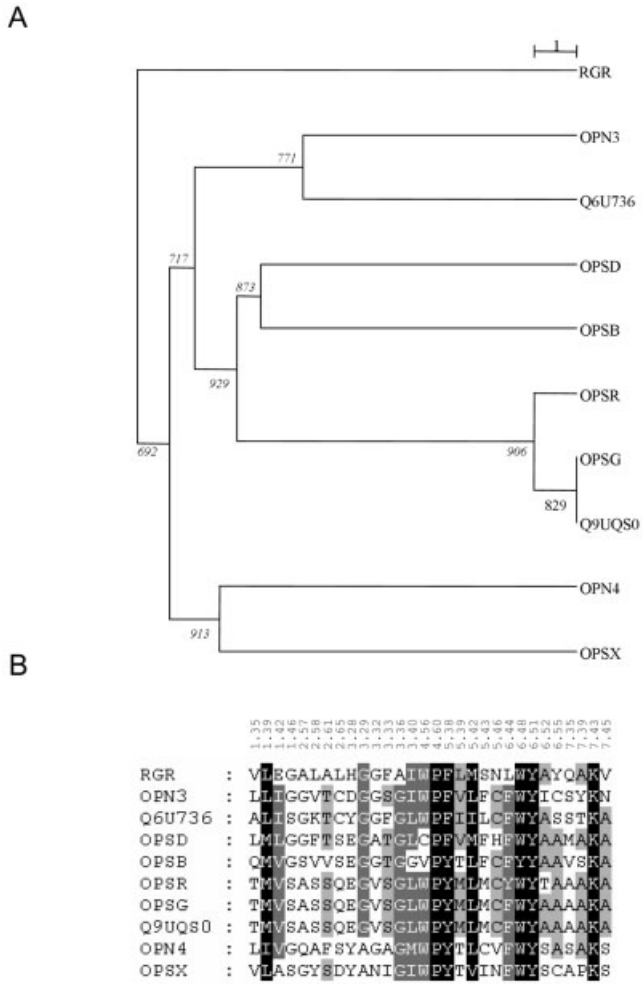


Fig. 11. The opsin receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. OPN3: Opsin 3; OPN4: Opsin 4; OPSB: Blue-sensitive opsin; OPSD: Rhodopsin; OPSG: Green-sensitive opsin; OPSR: Red-sensitive opsin; OPSX: Visual pigment-like receptor peropsin; Q6U736: Neuropsin; Q9UQS0: Photopigment apoprotein; RGR: RPE-retinal G protein-coupled receptor.

The predicted TM cavity of the MAS receptor has a volume of ca. 1,250 Å^3. Subsite 2 is significantly more polar than subsite 1 for most entries of this cluster [Fig. 9(B)]. Interestingly, several receptors of this cluster share a negatively charged spot in subsite 2 (Glu$^{4.60}$, Asp$^{5.38}$) which could be, like in chemokine receptors (see Chemokines cluster), an important anchoring point for ligands of that subfamily.

The SREB Cluster (6)

This cluster encloses six orphan receptors [Fig. 10(A)] out of which three entries (GPR27, GPR85, GP173) are subtypes of super-conserved receptors expressed in the brain (SREBs) and highly conserved in vertebrates.[71] Three additional receptors (GP101, GP161, and Q14439) whose function is currently unknown are proposed to belong to this cluster. Members of the SREB family exhibit a small apolar TM cavity (ca. 950 Å^3) with conserved aliphatic residues at TMs 1, 2 and 6 [Fig. 10(B)]. A charged residue is found for four out of the six GPCRs at position 1.35 (Arg/Lys) and represents a putative target for ligands of these poorly characterized receptors.

The Opsins Receptor Cluster (10)

The opsins receptor cluster regroups ten opsin and related receptors [Fig. 11(A)] activated by light. It includes the only GPCR (OPSD) for which a high-resolution X-ray structure is available.[26] 11-Cis retinal covalently binds to the conserved Lys$^{7.43}$ via the formation of a Schiff base with the aldehyde moiety of retinal. For many opsins, a negatively charged residue at 3.28 helps stabilizing the protonated Schiff base and closes the cavity between TMs 3 and 7. The cavity is about 1,100 Å^3 large and mainly consists in a very hydrophobic subsite 2 (Gly/Ala/Val$^{3.32}$, Val/Leu/Met$^{5.42}$, Phe/Tyr/Leu$^{6.44}$, Trp/Tyr$^{6.48}$) that serves to anchor the β-ionone moiety of retinal. Activation of the GPCR upon light induces an isomerization of 11-cis to all-trans retinal and subsequent translational/rotational TM motions triggering G protein coupling.[4] Strikingly, RGR binds preferentially all-trans retinal[72] despite the large similarity with other opsins in the TM cavity [Fig. 11(B)].

The Lipids Receptor Cluster (14)

This homogeneous cluster [Fig. 12(A)] groups 11 receptors for sphingosine-1-phosphate and lysophosphatidic

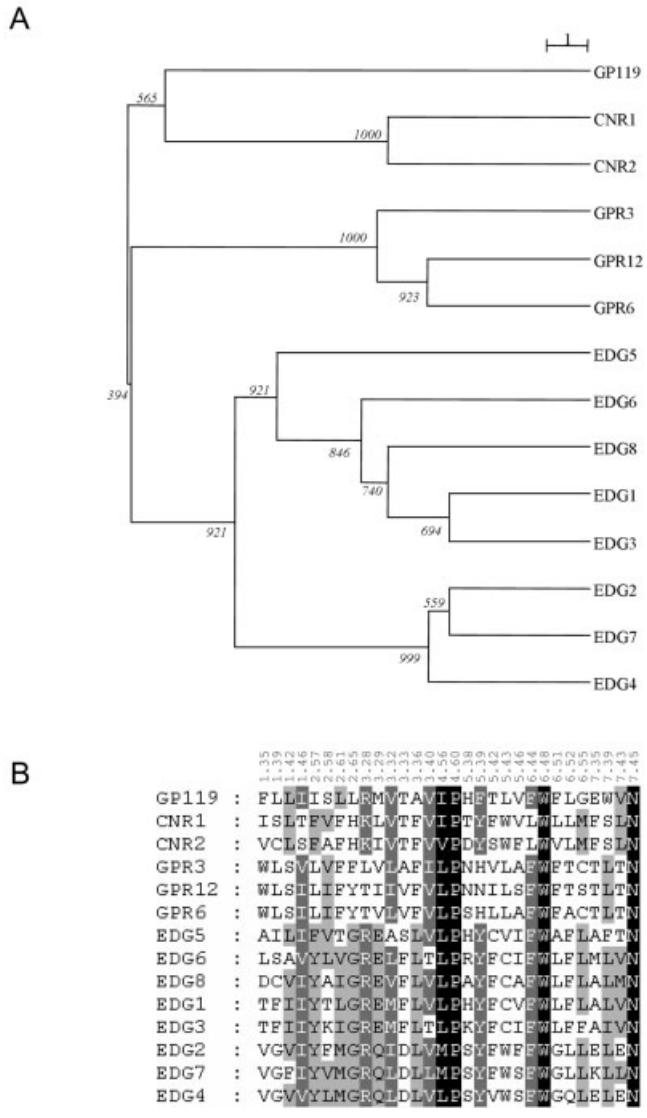


Fig. 12. The lipids receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. CNR1: Cannabinoid receptor 1; CNR2: Cannabinoid receptor 2; EDG1: G protein-coupled receptor EDG-1; EDG2: Lysophosphatidic acid receptor Edg-2; EDG3: Lysosphingolipid receptor EDG-3; EDG4: Lysophosphatidic acid receptor Edg-4; EDG5: Lysosphingolipid receptor Edg5; EDG6: Putative G protein-coupled receptor, EDG6; EDG7: Lysophosphatidic acid receptor Edg-7; EDG8: Sphingosine 1-phosphate receptor Edg-8; GP119: G protein-coupled receptor 119; GPR12: G protein-coupled receptor 12; GPR3: G protein-coupled receptor 3; GPR6: G protein-coupled receptor 6.

acid (EDG1-8, GPR3, GPR6, GPR12), two cannabinoid receptors (CN1R, CN2R), and an orphan receptor (GP119).[73] Our clustering approach manages to unambiguously separate EDG receptor subtypes for sphingosine-1-phosphate or S1P (EDG1, 3, 5, 6, and 8) from receptors for lysophosphatidic acid or LPA (EDG2, 4, and 7). Interestingly, three constitutively active GPCRs predominantly expressed in the brain (GPR3, GPR6, GPR12) known to bind S1P with nanomolar affinities[74] cluster just beneath EDG receptors. Similarly, the clustering of cannabinoid receptors with EDG receptors can be related to the chemically similar nature of their respective endogenous ligands (anandamide, S1P, LPA) consisting of a long alkyl chain and a polar head (ethanolamine, phosphate).[75] Four of the 30 analyzed residues are very well conserved throughout the cluster (Val/Ile/Leu/Met$^{4.56}$, Pro$^{4.60}$, Trp$^{6.48}$, and Asn$^{7.45}$). If GP119, GPR3, GPR6, and GPR12 are discarded from the analysis, three additional positions are well conserved (Phe/Tyr$^{2.57}$, Lys/Arg$^{3.28}$, Tyr$^{5.39}$). Looking at the eight EDG receptors, receptors for LPA present a unique combination of six conserved residues (Gly$^{1.39}$, Gln$^{3.29}$, Asp$^{3.33}$, Ser$^{5.38}$, Trp$^{5.43}$, Gly$^{6.51}$) that differentiate them from S1P receptors [Fig. 12(B)]. Notably, physicochemical properties of position 3.29 (Gln for EDG2, 4, 7; Glu for EDG1, 3, 5, 6, 8) is used as a discriminant to select the most appropriate endogenous ligand through the establishment of specific hydrogen bonds with either a hydroxyl group of LPA or a protonable amino group of S1P.[76]

The cavity of the 14 GPCRs can be clearly divided into a hydrophilic subsite 1 (positions 1.35, 2.65, 3.28, 3.29) and a long hydrophobic channel extending from the polar subsite towards the center of the TM bundle between TMs 4 and 5 (positions 3.33, 4.56, 4.60, 5.38, 5.39, 5.42). Residues not taken into account in the present clustering also contribute to the TM binding cavity (e.g., position 7.36 for the hydrophilic subsite 1 and 4.59 for the hydrophobic channel).

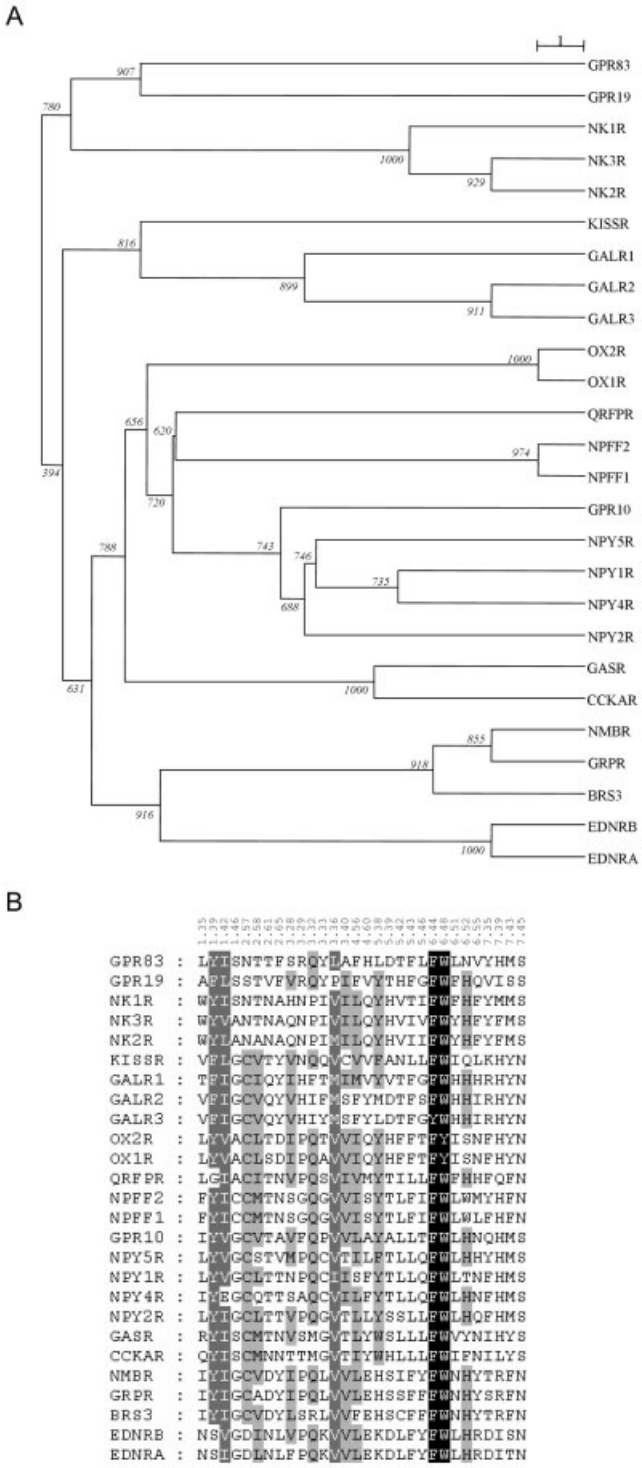
Our modeled cavities are in good agreement with site-directed mutagenesis experiments on EDG1 and EDG2 receptors[77] as well as on the cannabinoid receptors.[78,79] As an example, the positively charged residue at 3.28 represents the main anchoring atom for both receptor agonists and antagonists. Although GPR3, GPR6, and GPR12 also bind to S1P, it is very unlikely that their binding modes resemble that proposed for S1P/EDG receptors. Indeed, the two key residues in EDG receptors (Arg$^{3.28}$ and Glu$^{3.29}$) are both absent in GPR3, GPR6, and GPR12 receptors [Fig. 12(B)].

The Peptides Receptor Cluster (26)

The peptides receptor cluster includes 26 receptors for peptide ligands that can be classified in four main branches [Fig. 13(A)]. The low conservation of amino acids in the 7-TM cavity [Fig. 13(B)] reflects the large heterogeneity of the cognate peptide ligands and the shifted location of the peptide binding site towards the extracellular loops. Besides the quasi invariant location of an aromatic residue at positions 1.39, 6.44, 6.48 and the presence of aliphatic amino acids at positions, 1.42 and 3.36, the TM cavity of peptide ligands receptors show a very diverse pattern. Interestingly, many of the receptors of this cluster (20 out of 26) exhibit a cysteine at position 2.57. Throughout 369 human GPCRs, 40 entries share a Cys at 2.57. Out of these 40 entries, there are 16 orphan receptors. For 23 of the 24 liganded GPCRs, the endogenous ligand is a peptide (the exception refers to the histamine H3 receptor which belongs to the cluster of biogenic amine receptors). Site-directed mutagenesis data on cholecystokin receptors (GASR, CCKAR) have previously identified Cys$^{2.57}$ as important for either nonpeptide ligand binding or G Protein coupling.[80,81]

A typical cavity of a peptide receptor comprises following subsites; site 1 between TMs 1-2, 3, and 7 is either

hydrophobic (e.g., galanine, NPFF, and NPY receptors) or more hydrophilic (endothelin, bombesin, and orexin receptors). Subsite 2 delimited by TMs 4-5-6 is a large aromatic cage for most receptors (aromatic residue often found at positions 4.56, 5.38, 5.43, 5.46, 6.51, 6.53, and 6.55 in addition to conserved Phe$^{6.44}$ and Trp/Tyr$^{6.48}$). Since most peptides bind to a surface delimited by extracellular loops and the upper part of the TM bundle,[82] our selected 30 residues only partially account for peptide recognition. For some peptides (e.g., galanine), the N-terminal tail is buried in the TM region[83] and the C-terminus contacts extracellular loops at the receptor surface. Other peptides (e.g., neuropeptide Y, cholecystokinin, tachykinins) probably bind in a reverse orientation.[81,84] Many studies agree to conclude that peptide and nonpeptide ligand binding sites although sharing some key residues significantly differ.[85–87] Even the same nonpeptide antagonist can use distinct binding sites for two subtypes of the same receptor (e.g., NK1R, NK2R).[88] Due to the very large heterogeneity of possible binding mode to peptide receptors, it is impossible to draw general conclusions about the TM cavity. It is also very difficult to match target and ligand chemical spaces in this cluster. We will therefore analyze few key molecular determinants of peptide receptors on an individual basis.



Receptors for bombesin peptides (GRPR, NMBR, BRS3) are characterized by a conserved triad Asp$^{2.61}$/Glu$^{6.40}$/Arg$^{7.39}$ out of which the first two residues have been shown to be of crucial importance for activation of the gastrin-releasing peptide receptor.[89] Bombesins as well as peptidic antagonists are believed to mainly bind to the third extracellular loop of bombesin receptors[90,91] and also to some TM residues (e.g., 3.32, 6.55, 7.43).[90,92] Furthermore, position 5.42 which is variable within the subfamily [Fig. 13(B)] has been demonstrated to account for the specific recognition of a peptoid NMBR antagonist.[93]

The two receptors for endothelins (EDNRA, EDNRB) share 27 out of 30 identical residues. Thus their TM cavities are very similar. This explains why many endothelin receptor antagonists exhibit a poor selectivity among the two subtypes. Subtype selective antagonists have nevertheless been reported[94] and are likely to take advantage of the variability at position 3.28 (Phe for EDNRA, Val for EDNRB). The remarkable feature of EDNRA and EDNRB TM cavities is their high hydrophilicity: out the 30 amino acids studied, seven are conserved charged residues [Asp$^{2.57}$, Lys$^{3.33}$, Glu$^{4.60}$, Lys$^{5.38}$, Asp$^{5.39}$, Arg$^{6.55}$, Asp$^{7.35}$; Fig. 13(B)]. Accordingly, nonpeptide ligands of these receptors are very polar and usually share an acidic moiety.[95] Distinct binding sites for peptide agonists and nonpeptide antagonists have been located to subsites 1 and 2, respectively.[85] The four charged amino acids Asp$^{2.57}$, Lys$^{3.33}$,

Fig. 13. The peptides receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. BRS3: Bombesin receptor subtype-3; CCKAR: Cholecystokinin type A receptor; EDNRA: Endothelin-1 receptor; EDNRB: Endothelin B receptor; GALR1: Galanin receptor type 1; GALR2: Galanin receptor type 2; GALR3: Galanin receptor type 3; GASR: Gastrin/cholecystokinin type B receptor; GPR10: Prolactin-releasing peptide receptor; GPR19: G protein-coupled receptor 9; GPR83: G protein-coupled receptor 83; GRPR: Gastrin-releasing peptide receptor; KISSR: G protein-coupled receptor 54; NK1R: Substance-P receptor; NK2R: Substance-K receptor; NK3R: Neuromedin K receptor; NMBR: Neuromedin-B receptor; NPFF1: Neuropeptide FF receptor 1; NPFF2: Neuropeptide FF receptor 2; NPY1R: Neuropeptide Y receptor type 1; NPY2R: Neuropeptide Y receptor type 2; NPY4R: Neuropeptide Y receptor type 4; NPY5R: Neuropeptide Y receptor type 5; OX1R: Orexin receptor type 1; OX2R: Orexin receptor type 2; QRFPR: Orexigenic neuropeptide QRFP receptor.

Arg$^{6.55}$, Asp$^{7.35}$ have been found to delimit the nonpeptide antagonist binding site.[85,96,97] Lys$^{3.33}$ has been proposed to form a salt bridge with the carboxylic acid moiety of many endothelin receptor antagonists.[96] However, because of the spatial proximity of some charged residues (e.g., Arg$^{6.55}$/Asp$^{7.35}$, Lys$^{3.33}$/Glu$^{4.60}$/Lys$^{5.38}$), its is difficult to ascertain whether these amino acids contact nonpeptide ligands and/or form intramolecular salt bridges for stabilizing the receptor. It should be noticed that the Endothelin B receptor-like protein-2 (ETBR2) does not cluster with known endothelin receptor subtypes, nor with any other GPCR cluster.

Receptors for tachykinins (NK1R, NK2R, and NK3R) present a much more hydrophobic TM cavity [Fig. 13(B)]. Substance P and Neurokinin A binding sites in their cognate receptors (NK1R, NK2R, respectively) are deeply buried in the TM cavity. The substance P binding site strongly involves conserved residues at TM2 (Asn$^{2.57}$, Asn$^{2.61}$) and TM7 (Tyr$^{7.35}$)[98,99] whereas neurokinin A mainly interacts with TM3 (Gln$^{3.28}$), TM5 (His$^{5.39}$, Ile$^{5.43}$), and TM6 (Tyr$^{6.51}$, His$^{6.52}$, Phe$^{6.55}$).[98–100] Since the peptide binding sites are significantly buried in the TM cavity, binding areas for peptidic and nonpeptide ligands largely overlap, especially at the aromatic cage between TMs 5, 6, and 7.[86,101] Conformational adaptation of the ligand to its receptor is likely to play an important role for tachykinin receptors since mutation of residues conserved among the three subtypes (e.g., His$^{6.52}$, Tyr$^{7.35}$) differentially affects the binding the same nonpeptide antagonist.[88]

Receptors for cholecystokinin (CCKAR, GASR) exhibit a rather apolar TM cavity. The binding site of the biologically important C-terminal tail of the cholecystokinin CCK8 has been mapped to the herein predicted TM cavity.[81] Nine out of the 10 residues shown to line the peptide binding site are included in the 30 critical positions taken into account in our study. Interestingly, the peptide binding site, which is deeply buried in the TM cavity, involves residues from five TMs (2, 3, 5, 6, and 7) and is shared with that of a nonpeptide agonist.[81] Like for neurokinin receptors, it seems that the binding modes of peptide/nonpeptide agonists and nonpeptide antagonists to both CCK receptors (CCKAR, GASR) are different.[87]

Last, site-directed mutagenesis data on NPY receptors have mapped binding site for the C-terminal tail of neuropeptide Y to a hydrophobic pocket between TMs 1, 2, 6, and 7. More especially Tyr$^{2.64}$ (not taken into consideration in our analysis) appears to be a critical anchoring point. And once more, binding sites for peptidic and nonpeptide ligands are distinct although some overlap exists depending on the ligand.[102,103]

The Melatonin Receptor Cluster (7)

The melatonin receptor cluster contains seven receptors predominantly expressed in the brain and divided in two branches [Fig. 14(A)]. Members of the melatonin cluster present in their TM cavity (ca. 1,050 Å^3 for MTR1A) a clear hydrophobic-aromatic subsite 2 (hydrophobic/aromatic side chains frequently observed at position 3.36, 3.40, 5.42, 5.43, 5.46, 6.44, 6.48, and 6.51) with a possible polar

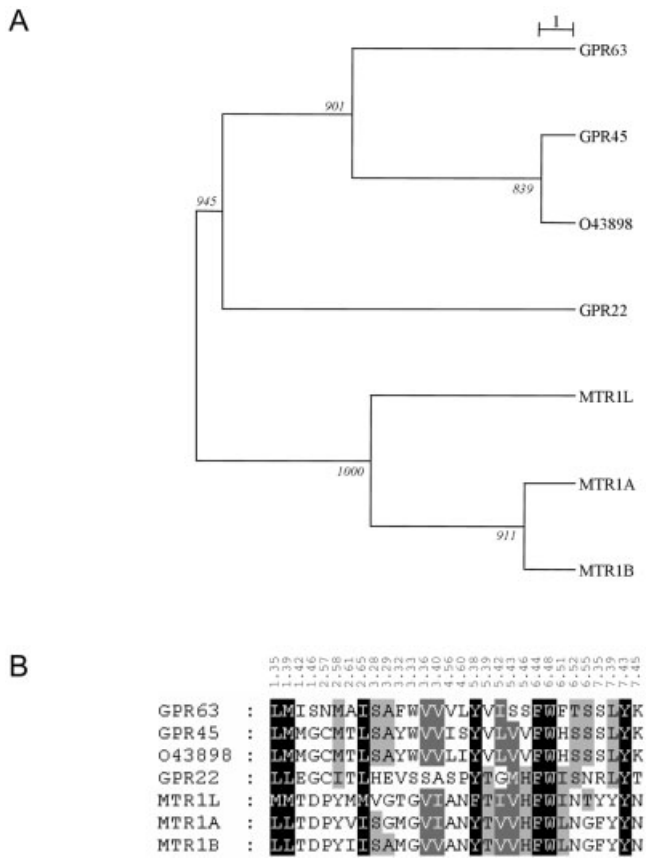


Fig. 14. The melanotonin receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. GPR22: G protein-coupled receptor 22; GPR45: G protein-coupled receptor 45; GPR63: G protein-coupled receptor 63; MTR1A: Melatonin receptor type 1A; MTR1B: Melatonin receptor type 1B; MTR1L: Melatonin-related receptor; O43898: High-affinity lysophosphatidic acid receptor homolog.

contribution of two residues (Asn/Ser$^{4.60}$ and Tyr$^{5.38}$) at one end of the cavity [Fig. 14(B)]. The concomitant presence of bulky side chains at 2.58 (Tyr/Met), 7.39 (Tyr/Leu), and 7.43 (Tyr) restricts the width of the cavity between TMs 2 and 7. At the other side of the cavity, several small side chains at 3.44 and 3.47 (two positions not considered for the phylogenetic tree) allow the formation of a hydrophobic needle between TMs 3 and 5.

The proposed cavity for melatonin receptors (MTR1A, MTR1B) is in perfect agreement with experimental data delineating the crucial role of few residues (Ser$^{3.35}$, Ser$^{3.39}$, Asn$^{4.60}$, His$^{5.46}$, Gly$^{6.55}$) for endogenous and synthetic agonist binding.[104–108] Receptor antagonists (e.g., Luzindole, 2-Phenylmelatonin, and 4-P-ADOT) are structurally and chemically related to synthetic agonists yet are usually bulkier with an extra aromatic ring. Their binding mode to melatonin receptor resembles that of synthetic agonists[109] but also involves additional interactions with Trp$^{6.48}$ and the hydrophobic needle between TMs 3 and 5. The latter area may nicely fit the extra phenyl ring often found in melatonin receptor antagonists.

Interestingly, some residues proved to be essential for ligand binding in several GPCR clusters (e.g., 3.28 and

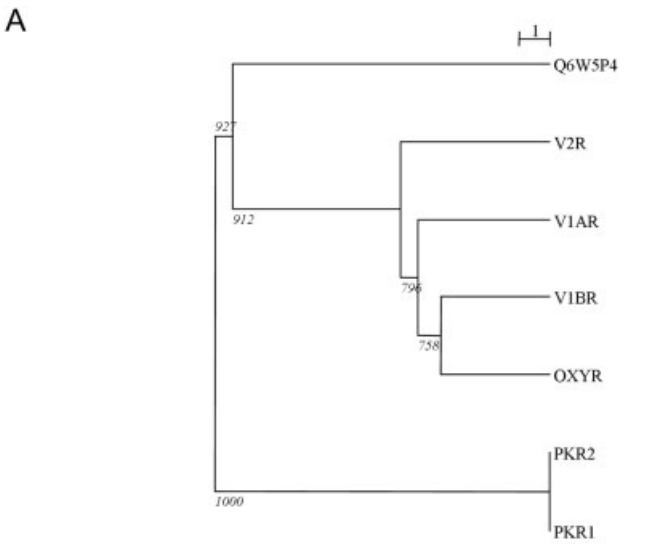


Fig. 15. The vasopeptides receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. OXYR: Oxytocin receptor; PKR1: Prokineticin receptor 1; PKR2: Prokineticin receptor 2; Q6W5P4: GPRA isoform A; V1AR: Vasopressin V1a receptor; V1BR: Vasopressin V1b receptor; V2R: Vasopressin V2 receptor.

3.32) are not anchoring points in melatonin receptors.[109] The orphan MTR1L receptor is closely related to MTR1A and MTR1B receptors with respect to sequence alignment [Fig. 14(B)] but is not a receptor for melatonin. Main differences observed in the TM cavity resides in five positions (2.61, 3.28, 3.32, 5.38, 6.55, 7.35). Site-directed mutagenesis[106,109] suggests that only the Gly/Thr$^{6.55}$ and the Phe/Tyr$^{5.38}$ mutations located between two critical residues (Asn$^{4.60}$ and His$^{5.46}$) are likely to be responsible for the singularity of MTR1L whose endogenous ligand still remains to be identified.

The Vasopeptides Receptor Cluster (7)

The vasopeptides receptor cluster is small group of only seven receptors [Fig. 15(A)]: the orphan Q6W5P4 receptor (GPR154) recently proposed to be involved in the pathogenesis of atopy and asthma,[110] four receptors for vasoconstrictive peptides (oxytocin, vasopressin), and two receptors for prokineticins.[124]

The TM cavity of these GPCRs is about 1,500 Å^3 large and uses overlapping but distinct binding sites for agonists and antagonists. The endogenous peptidic ligands and analogs fit the rather hydrophilic subsite 1 at the top of the 7-TM bundle[111] and nonpeptide antagonists bind to subsite 2 between TMs 3, 5, 6, and 7.[111–114] A clear difference between vasopressin and prokineticin receptors lies in the physiochemical properties of the second subsite which is quite polar in prokineticin receptors and much more hydrophobic in vasopressin receptors [see positions 3.32, 5.46, and 7.39; Fig. 15(B)]. This explains why typical vasopressin receptor antagonists are very hydrophobic.[115,116] Due to the high similarity between vasopressin and oxytocin receptors, it is difficult to find selective nonpeptide agonists for these receptors, especially for the oxytocin receptor (only moderately selective oxytocin receptor agonists have been described yet).[117] Fine subtype selectivity for antagonists is easier to tackle and may be regulated by single amino acid differences at the receptor level (e.g., positions 5.42 and 7.39 for vasopressin V_{1a} and V_{1b} receptors).[114] The close proximity of GPR154 and vasopressin/oxytocin receptors in our tree suggests that vaspopressin/oxytocin receptor antagonists may represent good starting points for identifying GPR154 antagonists.

The Adenosine Receptor Cluster (6)

The adenosine receptor cluster groups six receptors in two well separated branches [Fig. 16(A)]. The first one encloses two receptors for gonadotropin-releasing hormones GnRH and GnRH II[132], the second branch comprises four receptors for adenosine.[133] Clustering GnRH with adenosine receptors is unexpected with respect to the GRAFS classification.[16] However, a conserved pattern in the TM cavity of GnRH and adenosine receptors can be clearly identified [Fig. 16(B)]. Similarities mainly involve hydrophobic residues. As an example, four of the five 100% conserved residues in the TM cavity are hydrophobic. By contrast, polar side chains undergo a clearly different distribution [Fig. 14(B)] and may be responsible for receptor selectivity. Indeed two charged residues (Asp$^{2.61}$, Lys$^{3.32}$) conserved only in GnRH receptors are crucial anchoring amino acids for the GnRH hormone. GnRH binding site has been extensively investigated by site-directed mutagenesis and has been mapped to TMs 2, 5, and 6.[118–120] Eight of the 12 TM amino acids (Asp$^{2.61}$, Asn$^{2.65}$, Lys$^{3.32}$, Tyr$^{5.38}$, Asn$^{5.39}$, Thr$^{5.42}$, Tyr$^{6.51}$, Tyr$^{6.52}$) experimentally shown to line the hormone binding site are comprised in the 30 critical residues analyzed herein. Three of the four other important residues (2.50, 2.53, 2.64, and 5.41) are either accessible to a putative ligand (2.53, 2.64) or involved in the structural integrity of the receptor (Asn$^{2.50}$).

The TM cavity of GnRH receptors (1,500 Å^3 for GNRHR) is surprisingly hydrophilic, notably at subsite 1 (Arg$^{1.35}$, Glu$^{2.53}$, Asp$^{2.61}$, Asn$^{2.65}$, Lys$^{3.32}$). A remote aromatic pocket, delimited between TMs 5 and 6 (Tyr$^{5.38}$, Phe$^{6.44,}$ Trp$^{6.48}$, Tyr$^{6.51}$, Tyr$^{6.52}$) interacts with aromatic amino acids of the GnRH peptide. The chemotypes of known nonpeptide GnRH antagonists[121] match rather well the properties of the TM cavity. The heterocyclic scaffold (quinolone, indole) could H-bond to the numerous polar side chains at TMs 1 and 2, the basic amine (piperidine, guanidine) may inter-

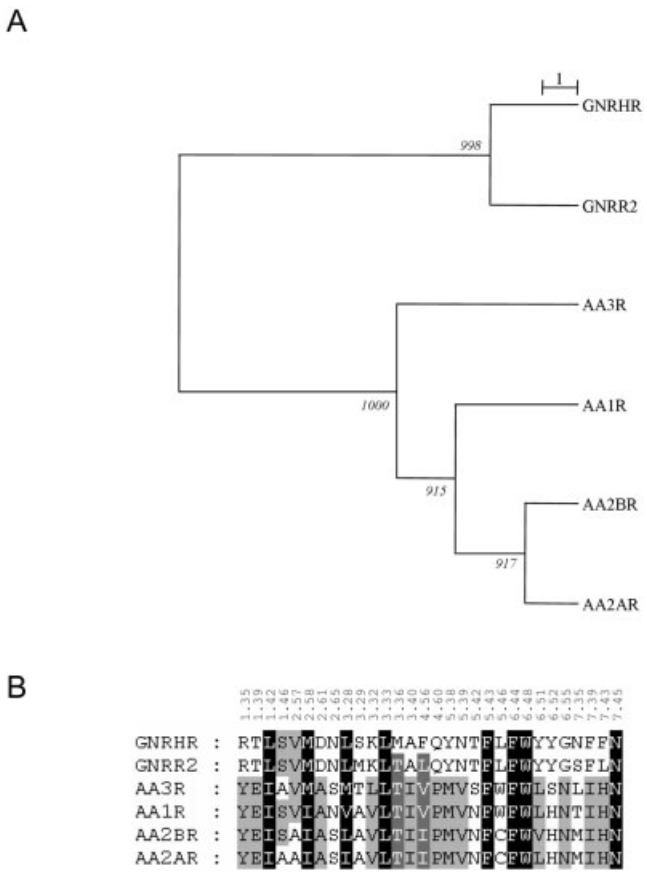


Fig. 16. The adenosine receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. AA1R: Adenosine A1 receptor; AA2AR: Adenosine A2a receptor; AA2BR: Adenosine A2b receptor; AA3R: Adenosine A3 receptor; GNRHR: Gonadotropin-releasing hormone receptor; GNRR2: Gonadotropin-releasing hormone II receptor.

act with Glu$^{2.53}$ or Asp$^{2.61}$, and one or two substituted aromatic rings could fill the aromatic subsite between TMs 5 and 6.

Adenosine receptors exhibit a quite conserved TM cavity [Fig. 16(B)]. Polar residues at positions 1.39 (Glu) and 6.55 (Asn), conserved in adenosine receptors, are believed to contribute both to interhelical hydrogen-bonding and ligand binding.[122] Interestingly, Glu$^{1.39}$ seems to be an anchoring point for agonists only,[123] whereas Asn$^{6.55}$ has been shown to be important for both agonist and antagonist binding.[122] The adenine moiety found in many nonselective adenosine receptor ligands is proposed to interact with conserved TM3 residues (e.g., Thr$^{3.36}$) whereas the ribose ring probably interacts with hydrophilic residues at TMs 3 and 7 (Thr$^{3.36}$, Ser$^{7.42}$, His$^{7.43}$).[124] Positions not taken into account herein (e.g., 3.31, 3.37) have been proposed to mediate ligand selectivity although it is unclear whether the observed mutagenesis effects are resulting form direct or indirect alteration of ligand binding.[124] The variable position 7.35 which has been shown to be accessible in the binding crevice of the AA1R receptor is likely to play an important role for directing the fine selectivity towards the adenosine receptor subtypes.

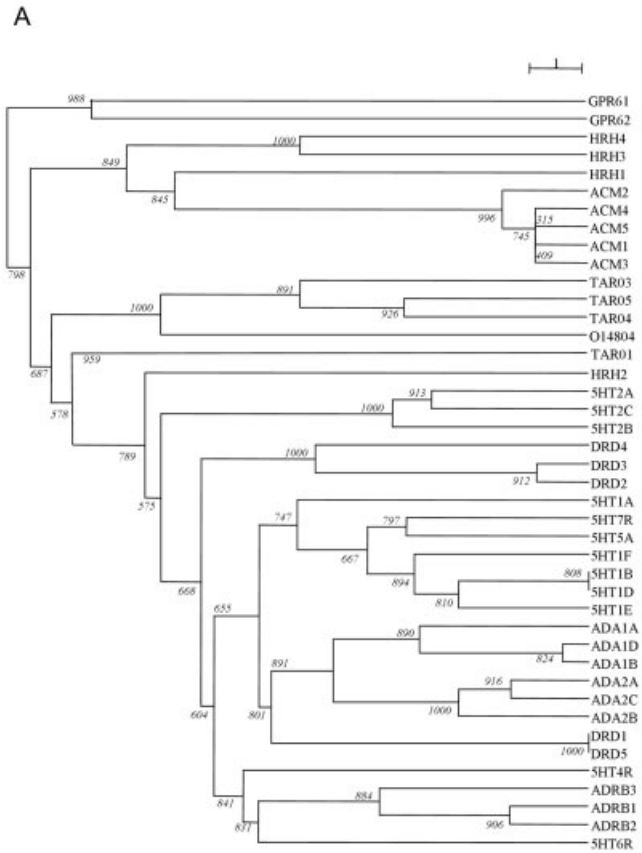
The Amines Receptor Cluster (42)

The amines receptor cluster is the largest one of our classification. Although it includes 42 representatives [Fig. 17(A)], it shows a remarkable homogeneity as indicated by the high bootstrap value (Fig. 2). Because of the extraordinarily vast literature on this receptor cluster, we will not discuss in details the binding site cavity of each receptor but rather concentrate on main general features explaining either selectivity or promiscuity.

Receptors from the amines receptor cluster present a well-defined TM cavity (e.g., 1070 Å^3 for 5HT2A) with two hydrophobic/aromatic subsites on both side of the conserved Asp$^{3.32}$. Subsite 2 between TMs 4, 5, and 6 (positions 4.56, 4.60, 5.38, 5.42, 5.43, 5.46, 6.44, 6.48, 6.51, 6.52, 7.35) has a strong aromatic character whereas subsite 1 contributes mainly by aliphatic residues [Val/Ile$^{1.35}$, Leu/Ile/Met/Thr$^{1.39}$, Val/Ile/Leu/Gly$^{1.42}$, Met/Leu/Val/Ile$^{2.58}$, Leu/Phe/Tyr/Trp$^{7.35}$, Val/Phe,Tyr/Asn$^{7.39}$, Tyr/Trp$^{7.43}$ and Ser/Asn$^{7.45}$; see Fig. 17(B)].

A typical hallmark of this cluster is a conserved aspartic acid at position 3.32 (excepted for the two orphans GPR61 and GPR62 which present a serine) and a cluster of aromatic residues at TM6 (Phe/Tyr$^{6.44}$, Trp$^{6.48}$, Phe/Tyr$^{6.51}$, Phe$^{6.52}$) and TM7 (Tyr/Trp$^{7.43}$). The carboxylate of Asp$^{3.32}$ is the counter ion of the basic amine,[17] and all above-cited aromatic amino acids are known to be crucial for recognition of the aromatic moiety common to biogenic amines.[125–127] For the 5-HT4 receptor, a network of intramolecular interactions between Asp$^{3.32}$, Trp$^{6.48}$, and Phe$^{6.51}$ has been proposed to stabilize the receptor in a silent state suitable for inverse agonist binding.[128] The aromatic cluster (notably Trp$^{6.48}$) seems to act as a molecular switch in triggering receptor activation by a cascade of changes. The modification of the rotameric states of these aromatic residues is followed by the alteration of the Pro$^{6.50}$ kink[129] and the disruption of an ionic lock between TMs3 and 6.[130]

The nature of three residues on top of TM5 (5.42, 5.43, 5.46) allows a clear differentiation of amine receptors. The 5-HT1 branch (including 5HT5A and 5HT7R) presents a "STA" fingerprint whereas 5-HT2 subtype presents a "GSA/S" sequence [Fig. 17(B)]. Receptors for catecholamines (dopamine, adrenaline) are enriched in Ser residues (e.g., "SSS" for dopamine and adrenergic receptors). Receptors for acetylcholine exhibits a "TAA" sequence. Receptors for trace amine present a more diverse pattern of residues. This variability reflects the variety of substituents on the aromatic rings of the endogenous ligands (phenyl, phenol, catechol, imidazole, indole).[131] Some additional specific amino acids also account for the selectivity of ligand recognition. Hence, Asn$^{6.52}$ is typical of acetylcholine receptors. It is important for both agonist and antagonist binding, with a more pronounced effect for antagonist recognition.[132] Likewise, Asn$^{6.55}$ found mainly in β-adrenergic receptors has been shown to account for the stereospecific recognition of β-adrenergic receptor agonists.[133] In the case of histamine receptors, although residues at positions 5.42, 5.43, and 5.46 of TM5 are not strictly conserved, they do participate to the selective



recognition of the histamine imidazole ring. Asp$^{5.42}$ and Glu$^{5.46}$ are responsible for the selective binding of the histamininergic agonists to HRH2, and HRH3/HRH4 subtypes respectively[134,135] by establishing an ion pair with the protonated imidazole ring. The HRH1 subtype lacks a negatively charged residue at TM5 yet utilizes an asparagine side chain (Asn$^{5.46}$) for H-bonding the imidazole ring of histamine. Thus, the example of histamine receptors demonstrates that the endogenous ligand must adapt its conformation and binding mode to the local TM5 environment.[134] The HRH1 subtype is unique among histamine receptors in possessing a Lys$^{5.39}$ which has been shown to be a selective anchoring point for second-generation HRH1 antagonists sharing an acidic moiety (e.g., acrivastine, etirizine).[136] It is now widely accepted that agonists and antagonists of biogenic amine receptors share overlapping but not identical binding sites; both agonists and antagonists occupy subsite 2 whereas subsite 1 is predominantly filled by antagonists.[137–139] Out of the four TM2 residues used in our clustering scheme, three positions (2.57, 2.61, and 2.65) have been demonstrated to be occluded upon antagonist binding.[138] The pseudo-symmetrical distribution of two hydrophobic cavities around the central Asp$^{3.32}$ explains why many biogenic amine receptor ligands also exhibit nonspecific chemotypes (a symmetrical distribution of aromatic rings around a central basic amine).[19,139] However, subsite 1 can be used to design selective antagonists since it has been shown to be responsible for the selective binding of dopamine D2 versus D4 antagonists.[139] Interestingly, in the two orphan receptors GPR61 and GPR62, the prototypical Asp$^{3.32}$ is replaced by a serine but a conserved Glu$^{7.35}$ might be an alternative anchoring residue for putative ligands of these receptors.

The Melanocortins Receptor Cluster (5)

The melanocortins receptor cluster comprises five related receptors for melanocortins [Fig. 18(A)]. A unique feature of this cluster is the availability of an endogenous

Fig. 17. The amines receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. 5HT1A: 5-hydroxytryptamine 1A receptor; 5HT1B: 5-hydroxytryptamine 1B receptor; 5HT1D: 5-hydroxytryptamine 1D receptor; 5HT1E: 5-hydroxytryptamine 1E receptor; 5HT1F: 5-hydroxytryptamine 1F receptor; 5HT2A: 5-hydroxytryptamine 2A receptor; 5HT2B: 5-hydroxytryptamine 2B receptor; 5HT2C: 5-hydroxytryptamine 2C receptor; 5HT4R: 5-hydroxytryptamine 4 receptor; 5HT5A: 5-hydroxytryptamine 5A receptor; 5HT6R: 5-hydroxytryptamine 6 receptor; 5HT7R: 5-hydroxytryptamine 7 receptor; ACM1: Muscarinic acetylcholine receptor M1; ACM2:Muscarinic acetylcholine receptor M2; ACM3: Muscarinic acetylcholine receptor M3; ACM4: Muscarinic acetylcholine receptor M4; ACM5: Muscarinic acetylcholine receptor M5; ADA1A: Alpha-1A adrenergic receptor; ADA1B: Alpha-1B adrenergic receptor; ADA1D:Alpha-1D adrenergic receptor; ADA2A: Alpha-2A adrenergic receptor; ADA2B: Alpha-2B adrenergic receptor; ADA2C: Alpha-2C-adrenergic receptor; ADRB1: Beta-1 adrenergic receptor; ADRB2: Beta-2 adrenergic receptor; ADRB3:Beta-3 adrenergic receptor; DRD1: D1A dopamine receptor; DRD2: D2 dopamine receptor; DRD3: D3 dopamine receptor; DRD4: D4 dopamine receptor; DRD5: D1B (D5) dopamine receptor; GPR61: G protein-coupled receptor 61; GPR62: G protein-coupled receptor 62; HRH1: Histamine H1 receptor; HRH2: Histamine H2 receptor; HRH3: Histamine H3 receptor; HRH4: Histamine H4 receptor; O14804: Putative neurotransmitter receptor; TAR01: Trace amine receptor 1; TAR03; Trace amine receptor 3; TAR04: Trace amine receptor 4; TAR05: Trace amine receptor 5.

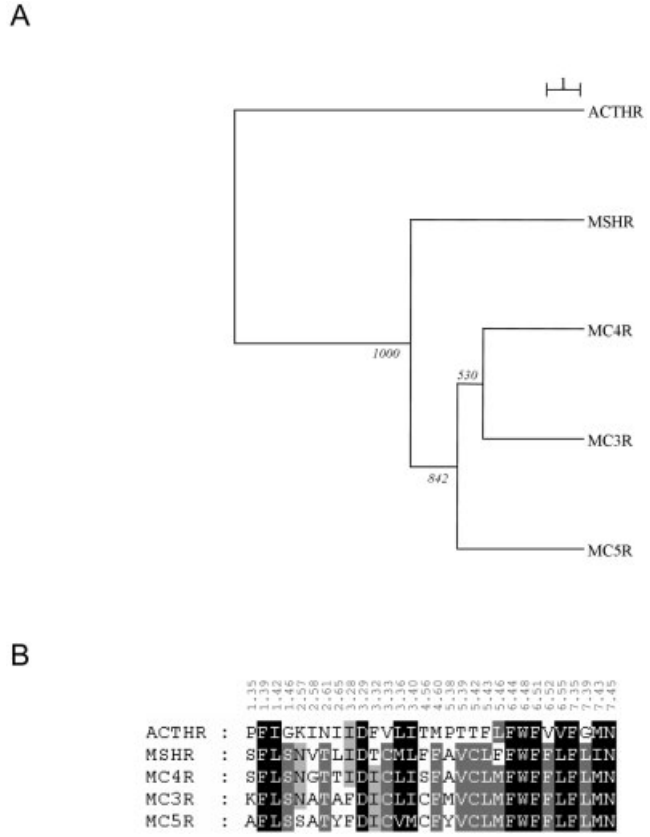


Fig. 18. The melanocortins receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. ACTHR: Adrenocorticotropic hormone receptor; MC3R: Melanocortin-3 receptor; MC4R: Melanocortin-4 receptor; MC5R: Melanocortin-5 receptor; MSHR: Melanocyte stimulating hormone receptor.

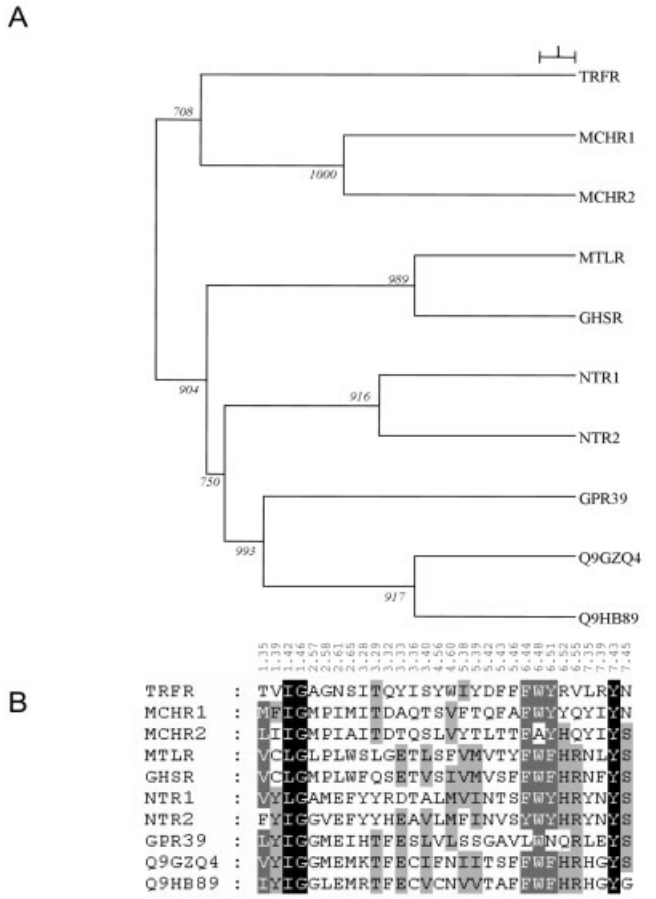


Fig. 19. The brain-gut peptides receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. GHSR: Growth hormone secretagogue receptor type 1; GPR39: G protein-coupled receptor 39; MCHR1: G protein-coupled receptor 24; MCHR2: Melanin-concentrating hormone 2 receptor; MTLR: Motilin receptor; NTR1: Neurotensin receptor type 1; NTR2: Neurotensin receptor type 2; Q9GZQ4: Neuromedin U receptor 2; Q9HB89: Neuromedin U receptor 1; TRFR: Thyrotropin-releasing hormone receptor.

inverse agonist (Agouti-related protein) for two receptor subtypes (MC3R and MC4R).[140] The sequence alignment shown in Figure 18(B) is characterized by a high proportion of hydrophobic or aromatic residues among conserved amino acids (Phe$^{1.39}$, lle/Leu$^{1.42}$, Ile/Phe$^{3.28}$, Val/Leu/Met$^{3.36}$, Ile/Leu/Met$^{3.40}$, Met/Phe$^{4.60}$, Phe/Leu$^{5.43}$, Leu/Met/Phe$^{5.46}$, Phe$^{6.44}$, Trp$^{6.48}$, Phe$^{6.51}$, Val/Phe$^{6.52}$, Val/Leu$^{6.55}$, Phe$^{7.35}$, Ile/Met$^{7.43}$). The predicted TM cavity (1,280 Å^3 for MSHR) is formed by two subsites around a conserved aspartic acid residue (Asp$^{3.29}$). Subsite 1 is delimited by hydrophobic side chains of TMs 1, 2 (at positions 1.39, 1.42 and 2.58) and two hydrophilic residues of TMs 2, 7 (Asn/Ser/Lys$^{2.57}$, Asn$^{7.45}$). Subsite 2 is more voluminous and defined by aromatic and aliphatic side chains (Leu/Val/Met$^{3.36}$, Ile/Leu/Met$^{3.40}$, Phe/Met$^{4.60}$, Leu/Phe$^{5.43}$, Trp$^{6.48}$, and Val/Phe$^{6.52}$). The proposed cavity agrees very well with site-directed mutagenesis data.[141–144] It demonstrates the importance of both the acidic residues (Asp$^{3.29}$, as well as Asp$^{3.25}$ not used for the classification yet conserved in the receptors of the cluster) and the numerous hydrophobic/aromatic side chains (3.28, 3.40, 4.56, 6.51, 7.35) for peptide agonist binding and fine subtype selectivity. It is also compatible with the chemotypes of known nonpeptide agonists or antagonists of melanocortin receptors (arylpiperazines, phenylguanidines, tetrahydroisoquinoleines)[145] that are likely to bind through an H-bond assisted salt bridge to one of the two above-cited important acidic residues

The Brain–Gut Peptides Receptor Cluster (10)

The brain-gut peptides receptor clusteris defined by ten receptors organized into two branches [Fig. 19(A)]. Our classification for this cluster significantly differs from that proposed by Fredriksson.[16] Hence, GPR38 is assigned to be the ghrelin receptor in the GRAFS classification and not the motilin receptor as herein or in the Swiss-Prot (the ghrelin receptor being the growth-hormone segretagogue receptor GHSR). The related GPR39 receptor is not classified in the GRAFS classification whereas we find it close to GPR38 (MTLR) in agreement with a previous report.[146] Last, the receptors for hypothalamic peptides (MCHR1, MCHR2, TRFR) are far form each other and from the neurotensin receptors in the GRAFS classification.

This cluster is characterized by conserved medium-sized and small aliphatic residues [Fig. 19(B)] in TM 1 (Ile/

Leu[1.42], Gly[1.46]) and a cluster of aromatic residues in TM6 (Phe/Tyr[6.44], Trp[6.48], and Phe/Tyr[6.51], His/Tyr[6.52]). Their cavity (e.g., 1250 Å^3 for GHSR) is typically formed by two hydrophobic subsites; subsite 1 between TMs 1, 2, and 7 involves positions 1.35, 1.42, 2.57, 2.65 and subsite 2 between TMs 5 and 6, centered around position 3.40 (generally apolar) involves positions 5.39, 5.42, 6.44, 6.48, and 6.52. At the center of the cavity, several polar residues (Asp/Arg[3.32], Asp/Glu[3.33], and Gln/Arg[6.55]) connect both hydrophobic subsites and provide anchoring atoms for putative ligands.

Position 3.33 is indeed an anchoring amino acid to the basic amine found in most GHSR ligands.[147] In agreement with our cavity model, a hydrophobic subsite involving Met[5.39] and His[6.52] has been confirmed by side-directed mutagenesis.[147] Our predictions are also consistent with the experimentally-determined binding site between TMs 6 (Tyr[6.51]) and 7 (Tyr[7.35]) for a neurotensin-1 receptor antagonist.[148] It should be noted that a charged residue (Arg[6.54]) not taken into account herein but vicinal to the above-described Arg[6.55], is proposed to neutralize the negative charge of the NTR1 agonist SR-48692.[148]

Several key residues (3.33, 5.46, 6.48, 6.51, 6.52, 7.39) have been found to map the thyrotropin-releasing hormone (TRH) binding site to its receptor (TRFR).[149–152] They are all among the 30 critical positions used for the current analysis. The binding mode of TRH (PyroGlu-His-Pro-NH_2) has been thoroughly investigated and is in perfect agreement with our TRFR predicted cavity. Tyr106[3.33] interacts with the pyroglutamyl carbonyl group, Tyr282[6.51] faces the His aromatic ring, and Arg306[7.39] is H-bonded to the C-terminal carboxamide moiety of the peptide. Amino acid at position 3.33 plays a noticeable role in ligand binding.[153,154] It is an acidic residue in seven receptors [GHSR, GPR39, MTLR, NTR1, NTR2, Q9GZQ4, and Q9HB89; Fig. 19(B)]. In two other receptors (MCHR1 and MCHR2), an aspartate is observed at the preceding position (Asp[3.32]). The concomitant presence of an acidic residue in TM3 and an aromatic cluster in TMs 5 and 6 (excepted for TRFR) explains why many receptor antagonists from this cluster (e.g., MCHR1) can be derived from biogenic amine receptor ligands.[155]

The Acids Receptor Cluster (5)

The acids receptor group comprises five receptors [Fig. 20(A)] for which few information is available, excepted that known ligands share a carboxylic acid moiety. G109B (HM74) and Q8TDS4 (HM74a) have recently been shown to bind nicotinic acid. Their function could be to decrease lipolysis in adipose tissues, although nicotinic acid is unlikely to be the endogenous ligand.[156] GPR31 and GPR81 are two orphan receptors of unknown function sharing 41% and 65% sequence identity to HM74, respectively. Last, Q8TDS5 has recently been demonstrated to be a chemotactic receptor for 5-oxo ETE, an inflammatory eicosanoid expressed in eosinophils and neutrophils.[157]

All these receptors exhibit some sequence identity to purine, chemokine, chemoattractant, and somatostatin receptors which are grouped nearby into four independent

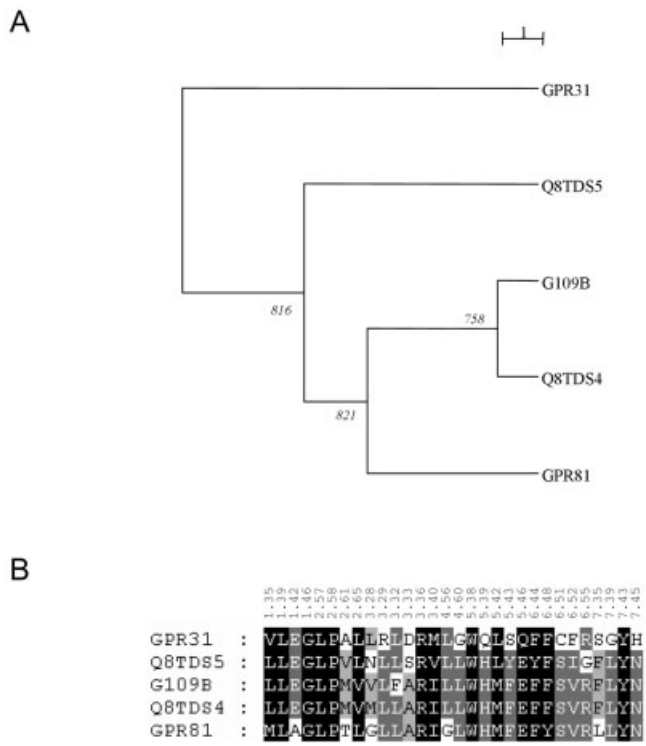


Fig. 20. The acids receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. G109B: G protein-coupled receptor HM74; GPR31: G protein-coupled receptor 31; GPR32: G protein-coupled receptor 32: GPR81: G protein-coupled receptor 31; Q8TDS4: G protein-coupled receptor HM74a; Q8TDS5: G protein-coupled receptor TG1019; UR2R: Urotensin II receptor (UR-II-R).

clusters (Fig. 2). Typical features from the latter four clusters, notably at TMs 1 and 2 [Gly[1.46], Leu[2.57], Pro[2.58], Ile/Val/Met[3.40], Fig. 20(B)] are also present in the current cluster. However, it significantly diverges from related clusters at TM1 (Glu[1.42] in four out of five members), TM3 (conserved Arg[3.36]) and TM5 (Glu/Gln[5.46])

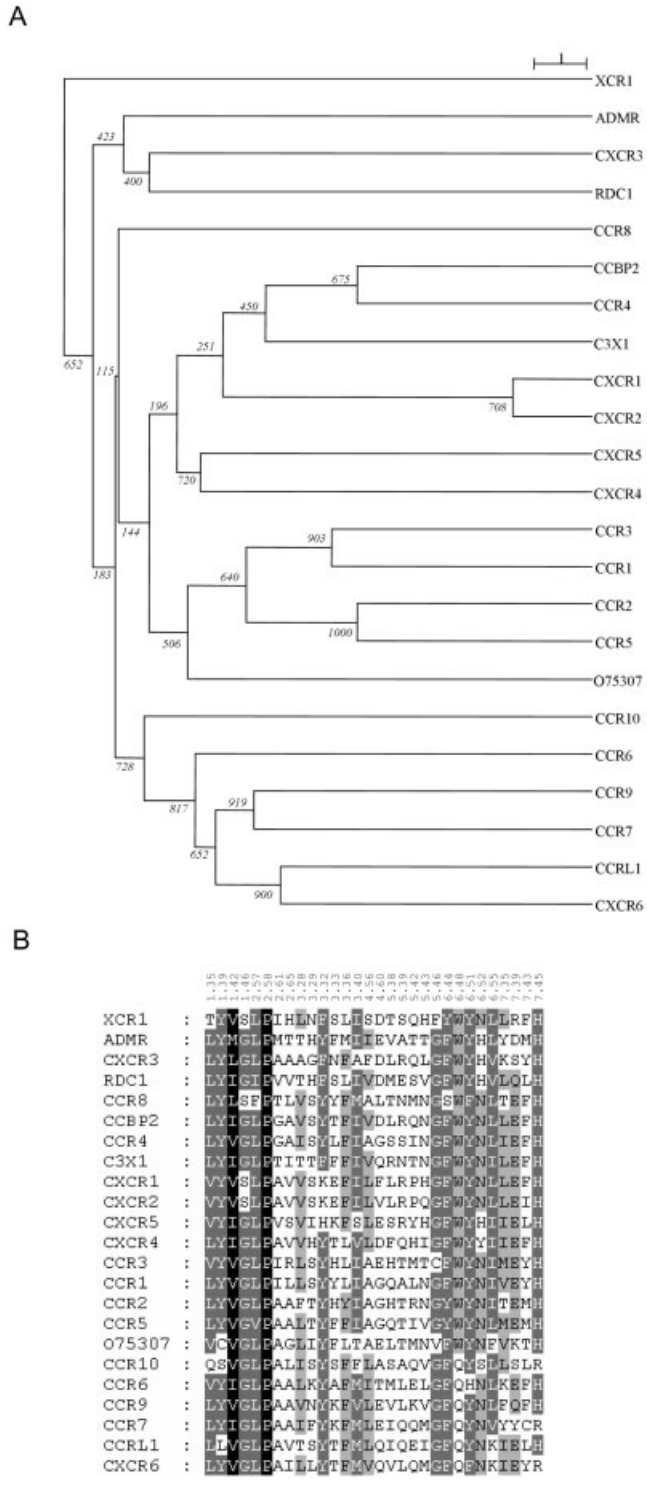
A representative cavity (e.g., G109B, volume of 1,310 Å^3) shows two clear subsites linked by a polar channel. Subsite 1 is a small hydrophobic pocket, with a polar contribution of the accessible Glu[1.42]. Its size is often restricted by the presence of a bulky Tyr[7.43] [Fig. 19(B)]. Subsite 2 is larger, hydrophobic and lined by medium-sized hydrophobic residues at 3.40 and 5.42 and a small side chain at 6.51.

G109B, GPR31, and GPR81 are likely receptors for anionic ligands due to the conserved presence of two arginine residues in the center of the cavity (positions 3.36 and 6.55). It is worth noting that receptors for carboxylic acids are also found in adjacent clusters (chemokines, chemoattractant, purinoreceptors, Fig. 2).

The Chemokines Receptor Cluster (23)

The chemokines receptor cluster contains 23 receptors [Fig. 21(A)] whose endogenous ligands are small-sized proteins (chemokines, adrenomedullin). The present cluster is nearly identical to that proposed by Fredriksson.[16] Members of the chemokines cluster present in their TM cavity (ca. 970 Å^3 for CCR5) a clear hydrophobic subsite 1

(Leu/Val$^{1.35}$, Tyr/Leu$^{1.39}$, Ala/Met/Val/Ile$^{2.61}$, Phe/Ala/Leu/Val/Ile$^{3.28}$, Tyr/Phe$^{3.32}$, Met/Leu/Ile/Val/Phe$^{7.35}$, and Met/Tyr/Phe/Leu/Ile$^{7.43}$). As in several other clusters (Opiates, Chemoattractants) a conserved Gly$^{1.46}$/Leu$^{2.57}$/Pro$^{2.58}$ motif is likely to induce a kink at the C-terminal part of TM2 [Fig. 21(B)]. In the center of the cavity, four aromatic side chains (Phe/Tyr/His$^{3.32}$, Phe/Tyr$^{3.36}$, Phe/Tyr$^{6.44}$, Trp$^{6.48}$, and Tyr/Phe$^{6.51}$) form an aromatic cage, typical of this cluster. Last, a much more polar subsite 2 is formed between TMs 3, 4 and 5 involving notably positions 3.33 (Glu, Lys), 4.60 (Asp, Glu, Gln), 5.42 (Glu, Arg, Lys, Gln), and 5.43 (Asn). The proposed TM cavities are in good agreement with experimental data. Likewise, many chemokine receptors (e.g., CCR2, CCR4, CCR5) have been shown to use a negatively-charged residue (Asp/Glu$^{4.60}$, Glu$^{7.39}$) to recognize the protonable amine of typical chemokine receptor antagonists.[158–161] Notably a negative charge at 7.39 is close in space to the prototypical position 3.32 used by biogenic amine receptors, thus explaining why chemokine receptor antagonists may also bind to biogenic amine receptors.[3] Besides the conserved Asp/Glu at 7.39, the above-described aromatic cage between TMs 3 and 6 is also an important anchoring site for aromatic moieties of chemokine receptor antagonists.[161]



The Opiates Receptor Cluster (13)

The opiates receptor cluster contains 13 receptors that can be divided into four main branches [Fig. 22(A)]. Endogenous ligands for receptors of this cluster are all short peptides (from 5 to 30 amino acids) usually rich in basic residues likely to interact with a conserved Asp$^{3.32}$ in TM3. The proposed classification is in agreement with that of Fredriksson.[16] This large cluster is characterized by conserved hydrophobic residues at TMs 1 and 2 [Ile/Leu/Val$^{1.35}$, Tyr$^{1.39}$, Ile/Val$^{1.42}$, Gly$^{1.46}$, Val/Leu/Met$^{2.57}$, Pro$^{2.58}$; Fig. 22(B)] and hydrophilic amino acids at TM 7 (His/Tyr$^{7.43}$, Asn$^{7.45}$). A conserved proline at 2.58, which might induce a kink at the TM2, is typical of this cluster and also present in chemokines and purinoreceptors clusters.

The somatostatin receptor cavity (ca. 1,070 Å^3 for SSR1) exhibits two hydrophobic subsites. Subsite 1 is delimited by Val$^{1.42}$, Val/Leu$^{2.57}$, Pro$^{2.58}$ and Tyr$^{7.43}$. Subsite 2 is lined by Leu/Met$^{3.29}$, Ala/Gly$^{3.33}$, Val/Met$^{4.56}$ Ile/Leu/Val$^{4.60}$, Phe$^{5.38}$, Trp$^{6.48}$ and Tyr/Phe$^{6.52}$. Phe$^{3.37}$, which is not included in the 30 selected amino acids, is conserved in all the somatostatin receptors and contributes to the second subsite. Both subsites are linked by a hydrophilic channel formed by two polar side chains conserved in almost all the cluster members (Asp$^{3.32}$ and Asn$^{7.45}$). Structure–function relationships studies indicate that the somatostatin (SST-14) core residues (spanning Phe^6 to Phe^{12}) interact with a binding pocket located between TMs

Fig. 21. The chemokines receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. ADMR: Adrenomedullin receptor; C3X1: CX3C chemokine receptor 1; CCBP2: Chemokine binding protein 2; CCR1: C-C chemokine receptor type 1; CCR10: C-C chemokine receptor type 10; CCR2: C-C chemokine receptor type 2; CCR3: C-C chemokine receptor type 3; CCR4: C-C chemokine receptor type 4; CCR5: C-C chemokine receptor type 5; CCR6: C-C chemokine receptor type 6; CCR7: C-C chemokine receptor type 7; CCR8: C- C chemokine receptor type 8; CCR9: C-C chemokine receptor type 9; CCRL1: C-C chemokine receptor type 11; CXCR1: High affinity interleukin-8 receptor A; CXCR2: High affinity interleukin-8 receptor B; CXCR3: C-X-C chemokine receptor type 3; CXCR4: C-X-C chemokine receptor type 4; CXCR5: C-X-C chemokine receptor type 5; CXCR6: C-X-C chemokine receptor type 6; O75307: Putative chemokine receptor; RDC1: G protein-coupled receptor RDC1 homolog; XCR1: Chemokine XC receptor 1.

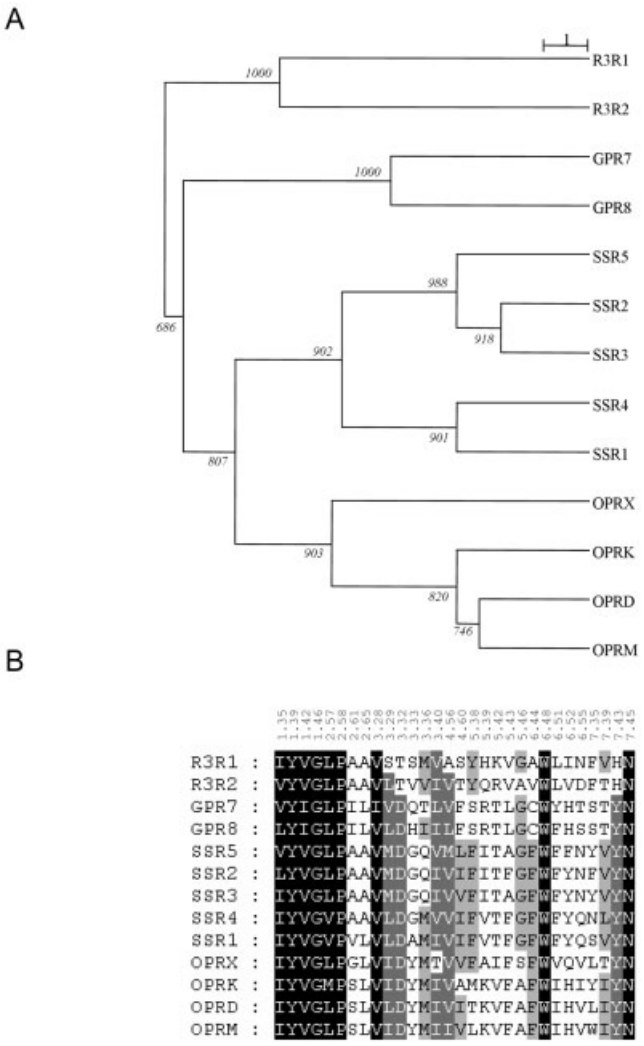


Fig. 22. The opiates receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. GPR7: G protein-coupled receptor 7; GPR8: G protein-coupled receptor 8; OPRD: Delta-type opioid receptor; OPRK: Kappa-type opioid receptor; OPRM: Mu-type opioid receptor; OPRX: Nociceptin receptor; R3R1: Somatostatin- and angiogenin-like peptide receptor; R3R2: Relaxin 3 receptor 2; SSR1: Somatostatin receptor type 1; SSR2: Somatostatin receptor type 2; SSR3: Somatostatin receptor type 3; SSR4: Somatostatin receptor type 4; SSR5: Somatostatin receptor type 5.

3-7.[162] Several studies agree to define positions 3.32, 6.55, and 7.35 as key residues for agonist binding.[163,164] Although some nonpeptide somatostatin receptor agonists[165] and antagonists[166] have been described recently, residues lining their binding pocket have still not been elucidated.

The opioid-receptor cavity resembles that of somatostatine receptors; it is centered around the conserved aspartic acid ($Asp^{3.32}$) although its role might not be as crucial as for somatostatin and biogenic amine receptors.[167] The subsite located between TMs 3, 5, and 6 is mostly composed of aromatic residues ($Tyr^{3.33}$, $Phe^{5.43}$, and Trp6.48). All these residues are conserved across δ, μ and κ-opioid receptors [Fig. 22(B)] suggesting a well conserved TM binding domain. In agreement with the modeled cavity, site-directed mutagenesis studies have clearly identified the important role of two tyrosine residues ($Tyr^{3.33}$, $Tyr^{7.43}$) for anchoring most δ-opioid receptor ligands.[168] Interestingly, a single amino acid variation between classical opioid receptors (OPRD, OPRK, OPRM) and the nociceptin receptor (OPRX) at position 5.39 [Lys vs. Ala, Fig. 22(B)] has been shown to be related to the selective binding of opioid alkaloids (e.g., naltrexone, bremazocine) to classical opioid receptors only.[169] The presence of a conserved aspartic acid at 3.32 and of aromatic residues at 6.48 and 6.51 for other poorly-studied members of that family (e.g., GPR7, GPR8,) suggests that biogenic amine ligands sharing an aromatic moiety and a protonated amine could bind to these two orphan receptors as well.

The Chemoattractants Receptor Cluster (17)

The chemoattractants receptor cluster groups receptors for chemoattractants, angiotensin II, and bradykinin into three branches [Fig. 23(A)]. This cluster is nearly identical to that proposed in the GRAFS classification.[16] A typical hallmark of this cluster is the very hydrophobic nature of the transmembrane cavity (positions 1.39, 2.57, 2.58, 3.32, 3.36, 3.40, 6.44, 6.48, 6.51, 7.43) topped by charged amino acids [e.g., Asp/$Lys^{3.33}$, $Arg^{5.38}$, Arg/Lys/$His^{5.42}$, Asp/$Glu^{7.35}$; see Fig. 23(B)]. Few site-directed mutagenesis studies have been undertaken to map the binding site of natural anaphylactic peptide ligands. All conclude that the N-terminal part of these ligands is recognized by charged residues of the receptor N-terminal domain[170] whereas the C-terminus interacts with charged residues at the top of the TM cavity ($Glu^{5.35}$, $Arg^{5.42}$, $Asp^{7.35}$).[171–173] A negatively-charged anchoring amino acid at 7.35 is specific of anaphylactic peptide receptors and may ensure the absolute specificity of the activation. Interestingly, the chemotypes of nonpeptidic C5a receptor ligands (agonists, antagonists) perfectly match observed physicochemical properties of the computed TM cavity. C5a agonists usually share a positively charged substituent that is likely to mimic one of the two important basic residues of C5a (Lys68, Arg74) that interacts with the conserved negatively charged side chains at the top of the 7-TM bundle ($Glu^{5.35}$, $Asp^{7.35}$).[174,175] C5a receptor antagonists[175] are much more hydrophobic and probably interact with the hydrophobic core of the cavity as suggested by one site-directed mutagenesis study.[172] $Ile^{3.32}$ and $Val^{7.39}$ are important anchoring residues, in addition to $Glu^{5.35}$ and $Asp^{7.35}$ that may also contact the positively charged group of the ligand.

The modeled TM cavity of formyl peptide receptors (FPR1R, FRRL1, FPRL2) is also in agreement with experimental data.[176] Charged residues, either conserved throughout the cluster ($Arg^{5.42}$) or among formyl peptide receptors ($Asp^{3.33}$, $Arg^{5.38}$) have been demonstrated to directly interact with both the N-formyl and the C-terminal carboxylate moieties of endogenous peptide ligands.[176] The peptide backbone is probably parallel to the main axis of TM5 with hydrophobic side chains contacting the numerous hydrophobic residues of the cavity, especially at TM2. Few antagonists of formyl peptide receptors (small peptides, bile acids) have been described up to

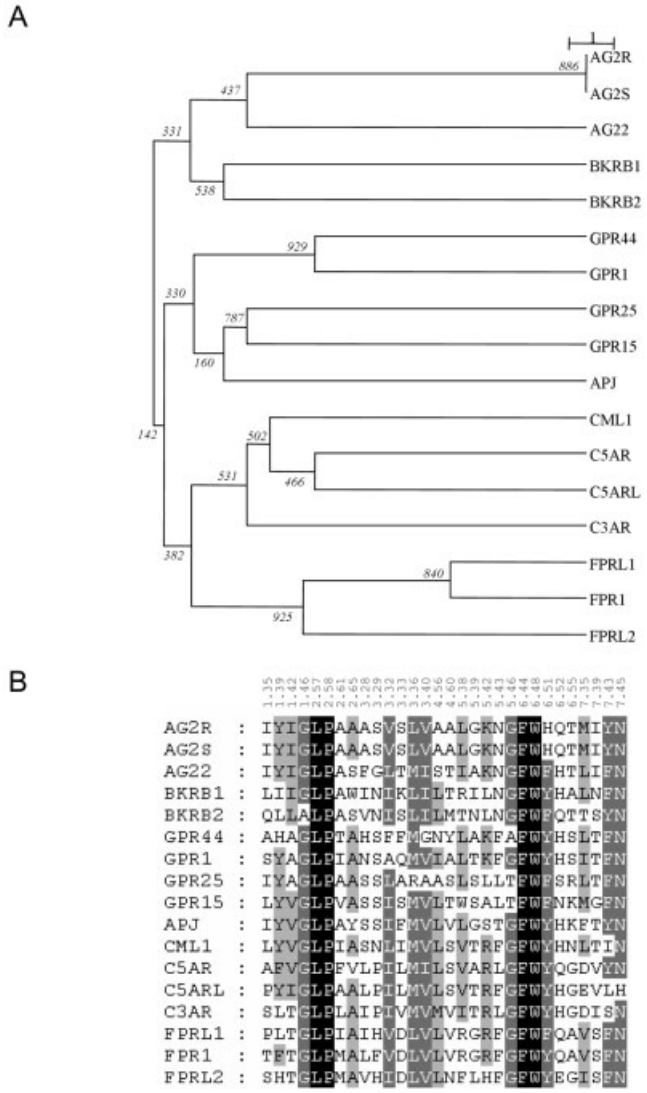


Fig. 23. The chemoattractants receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. AG22: Type-2 angiotensin II receptor; AG2R: Type-1 angiotensin II receptor; AG2S: Type-1B angiotensin II receptor; APJ: Apelin receptor; BKRB1: B1 bradykinin receptor; BKRB2: B2 bradykinin receptor; C3AR: C3a anaphylatoxin chemotactic receptor; C5AR: C5a anaphylatoxin chemotactic receptor; C5ARL: C5a anaphylatoxin chemotactic receptor C5L2; CML1: Chemokine receptor-like 1 (ChemR23); FPR1: fMet-Leu-Phe receptor; FPRL1: FMLP-related receptor I; FPRL2: FMLP-related receptor II; GPR1: G protein-coupled receptor 1; GPR15: G protein-coupled receptor 15 (BOB); GPR25: G protein-coupled receptor 25; GPR44: G protein-coupled receptor 44.

now.[177–179] Like previously described C5a receptor antagonists, they all share a very hydrophobic core and a negatively-charged group likely to interact with Arg$^{5.38}$ in the TM cavity.

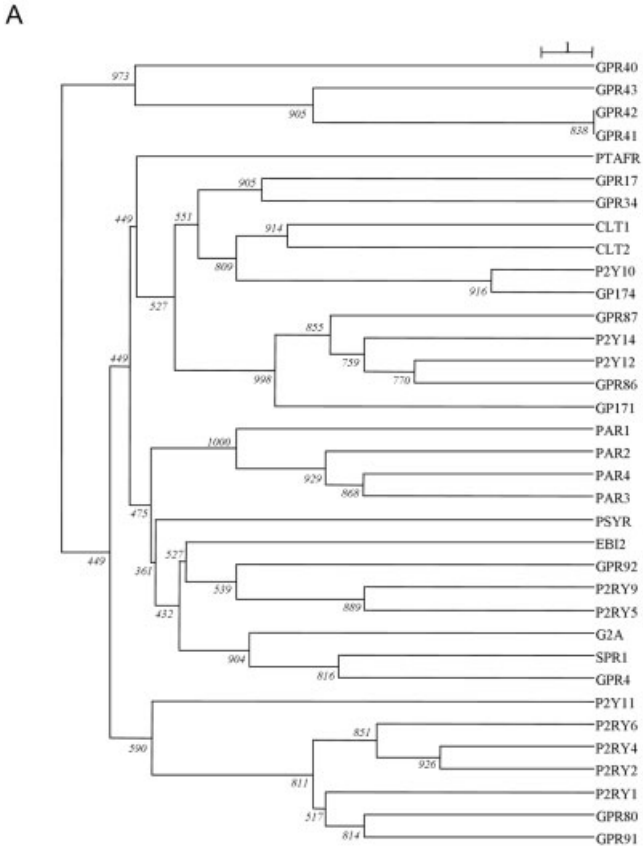
The predicted cavity for angiotensin II type 1 receptors (AG2R, AG2S) perfectly agrees with the site-directed mutagenesis used to map the binding sites for nonpeptide antagonists; for instance, insurmountable antagonism[180,181] is induced by the interaction of biphenyltetrazole substructure (e.g., Losartan) with both Lys$^{5.42}$ and the neighboring aromatic cage at TM6.

Bradikynin receptors are herein clustered with angiotensin II receptors. Significant similarities between these two receptor types have already been noticed.[182] Bradykinin type I and II receptors do not recognize the same ligand (des-Arg^{9}-bradykinine for BKRB1, bradykinin for BKRB2).[183] A simple explanation is given by looking at the corresponding TM cavities [Fig. 23(B)]. Lys$^{3.33}$ which is seen only for BKRB1 (replaced by Ser in BKRB2) prevents recognition of the C-terminal Arg9 of bradykinin.[184] No ligands have been described up to now for the four orphan receptors of this cluster. The identification of specific basic residues in the cavity of these receptors (Lys/Arg$^{6.55}$ for GPR15 and GPR25, Lys$^{5.42}$ for GPR1 and GPR44) close to the aromatic cage at TM6 suggests that aromatic carboxylic acids may represent a good starting point to design focused libraries towards these receptors.

The Purine Receptors Cluster (35)

The purine receptors cluster is one the largest in our classification [Fig. 24(A)] and contains 35 members with a significant proportion of orphan targets (nine in total). All known endogenous ligands are anionic molecules (carboxylates or phosphates). The major difference between the GRAFS classification and ours resides in the location of monocarboxylic acids receptors (GPR40, 41, 42, 43) which could not been classified by Fredriksson et al.[16] whereas we cluster them with known receptors for dicarboxylic acids (GPR80, GPR91).The large purinoreceptors cluster is characterized by basic residues [less than five per sequence; Fig. 24(B)] found at various spots in the TM cavity (2.65, 3.29, 3.32, 3.33, 4.56, 5.38, 5.39, 5.42, 6.52, 6.55, 7.39). Side chains of these arginine, lysine, and histidine residues are likely to neutralize the negative charges of the ligands. TMs 1, 2, and 3 contribute mainly to the hydrophobic part of the cavity (Phe/Tyr/Leu$^{1.39}$, Ile/Val$^{1.42}$, Gly$^{1.46}$, Leu/Phe$^{2.57}$, Pro$^{2.58}$, Phe/Tyr/Leu$^{3.33}$, Leu/Met/Ile$^{3.36}$, and Val/Ile/Met$^{3.40}$). A bunch of aromatic residues in TM6 (Phe/Tyr$^{6.44}$, Phe/Tyr$^{6.48}$, Phe/Tyr/His$^{6.51}$) is much conserved throughout the cluster. Interestingly, position 6.48, which is known to play a well-defined role in molecular activation of rhodopsin-like GPCRs is a phenylalanine for most members of this cluster.

In the case of carboxylic acid receptors, the distribution of charges at the protein surface is consistent with the structure of their ligands. Receptors for dicarboxylic acids (GPR80, GPR91) present five basic residues organized into two positively charges areas whereas receptor for monocarboxylic acids (GPR40, GPR41, GPR42, GPR43) present only four basic residues all involved in a single positive spot around Arg$^{5.39}$ and Arg$^{7.35}$. The first basic area of dicarboxylic acid receptors is located at TM3 (Arg$^{3.29}$, His$^{3.33}$) and the second one involves TMs 6 and 7 (His$^{6.52}$, Arg$^{6.55}$, Arg$^{7.39}$). Four of these residues have been shown to be anchoring points for succininic acid and α-ketoglutaric acid to GPR80/GPR91.[185] The modeled TM cavity of carboxylic acid receptors is also coherent with respect to ligands' size and shape. Hence, the cavity of the long chain fatty acid receptor (GPR40) extends much deeper into an hydrophobic subsite between TMs 3, 5, and 6 than the



corresponding cavity of short chain fatty acid receptors (GPR41, GPR43). The bulkiness of side chains at positions 6.44 and 6.48 is the suggested molecular explanation for this feature [Fig. 24(B)].

Receptors for cationic glyco- and phospholipids (G2A, GPR4, SPR1, and PSYR) share a conserved $Arg^{5.42}$ that is likely the anchoring residue of the phosphate group of endogenous ligands (LPC: lysophosphatidylcholine, SPC: sphingosylphosphorylcholine, psychosine). No negatively-charged residues are present in the TM cavity for counterbalancing the choline positive charge. The choline moiety therefore probably interacts with either the aromatic cluster at TM6 ($Phe^{6.44}$, $Phe/Tyr^{6.48}$, $Phe/Tyr^{6.51}$, $His^{6.55}$) or the very acidic second extracellular loop of SPC/LPC receptors. Last, the fatty carbon chain of SPC/LPC is proposed to fill a shallow hydrophobic needle between TMs 1 and 2 terminating at conserved positions between TM1 ($Val^{1.42}$, $Gly/Ser^{1.46}$) and TM2 ($Tyr^{2.53}$, $Gly^{7.46}$).

Recent studies[186] demonstrated that a patch of basic residues ($Arg^{3.29}$, $Arg/Lys^{6.55}$, $Arg^{7.35,}$ $Arg^{7.39}$) in nucleotide receptors participate to ligand recognition.[187,188] Interestingly, Gq-coupled receptor subtypes (P2RY1, P2RY2, P2RY4, P2RY6) are well separated from Gi-coupled receptor subtypes (GPR86, P2Y12, P2Y14) and present slightly different basic patches. The first subgroup share basic residues at positions 3.29, 6.55, and 7.39. The second subgroup share basic amino acids at positions 6.55 and 7.35. A thorough study of purinergic receptors recently agrees with our conclusion and additionally proposes that a Lys residue in the second extracellular loop to replace $Arg^{3.29}$ for Gi-coupled subtypes.[186] Other purinoreceptor subtypes (P2Y10, P2RY5, P2RY9), which do not recognize nucleotides present much less accessible basic residues in the TM cavity. For example, P2RY9 has recently been shown to be a receptor for lysophosphatidic acid (LPA).[189] Interestingly, it presents together with other atypical P2Y subtypes and related GPCRs (GP174, GPR92, P2Y10) a conserved glutamic acid at position 5.43. P2Y9 is not related to EDG receptors (EDG2, EDG4, and EDG7) which also recognize LPA but use $Arg^{3.28}$ and $Gln^{3.29}$ as main anchoring residues (see Lipids cluster). These findings suggest that a single ligand can be recognized by unrelated

Fig. 24. The purines receptor cluster: (**A**) Phylogenetic tree, (**B**) Cavity-based alignment. CLTR1: Cysteinyl leukotriene receptor 1; CLTR2: Cysteinyl leukotriene receptor 2; EBI2: EBV-induced G protein-coupled receptor 2; G2A: Lysophosphatidylcholine receptor G2A; GP171: G protein-coupled receptor 171; GP174: Putative P2Y purinoceptor FKSG79; GPR17: P2Y purinoceptor GPR17; GPR34: G protein-coupled receptor 34; GPR4: G protein-coupled receptor 4; GPR40: G protein-coupled receptor 40; GPR41: G protein-coupled receptor 41; GPR42: G protein-coupled receptor 42; GPR43: G protein-coupled receptor GPR43; GPR80: G protein-coupled receptor 80; GPR86: FKSG77 protein; GPR87: FKSG88 protein; GPR91: G protein-coupled receptor 91; GPR92: G protein-coupled receptor 92; P2RY1: P2Y purinoceptor 1; P2RY2: P2Y purinoceptor 2; P2RY4: P2Y purinoceptor 4; P2RY5: P2Y purinoceptor 5; P2RY6: P2Y purinoceptor 6; P2RY9: P2Y purinoceptor 9; P2Y10: P2Y purinoceptor 10; P2Y11: P2Y purinoceptor 11; P2Y12: P2Y purinoceptor 12; P2Y14: UDP-glucose receptor; PAR1: Proteinase activated receptor 1; PAR2: Proteinase activated receptor 2; PAR3: Proteinase activated receptor 3; PAR4: Proteinase activated receptor 4; PSYR: T cell-death associated protein; PTAFR: Platelet activating factor receptor; SPR1: G protein-coupled receptor 68.

GPCRs using quite different binding modes. It is likely that P2RY5 which shares 70% sequence identity with P2RY9 on our 30 consensus positions is also a phospholipid receptor.

PAR receptors exhibit a much more hydrophobic cavity, notably at the center between TMs 3 and 6, with a few conserved small-sized polar residues (Ser/Thr$^{6.51}$, Asn$^{6.52}$, and Ser/Thr$^{7.43}$). Accordingly, known PAR antagonists are quite hydrophobic with several aromatic rings along a pseudopeptide main chain.[190]

CONCLUSIONS

We have reported an exhaustive classification of human G protein-coupled receptors based on the analysis of 30 critical positions supposed to delimit the binding cavity of the typical transmembrane domain of ground-state receptors. The proposed phylogenetic tree is coherent and compatible with numerous experimental data, notably in identifying important residues for small molecular-weight ligand recognition. The present study does not claim that only the 30 selected residues are of interest for explaining ligand binding. It simply attempts to draw general structure-binding relationships over the entire family of human receptors on a minimal set of common positions.

We are confident about the alignment of members of a particular family (e.g., rhodopsin-like receptors) because the alignment method matches conserved fingerprints.[24] The alignment of members from different families (e.g., rhodopsin-like versus glutamate-like receptors) is more questionable. Recent mutagenesis studies on glutamate-like receptors suggest that the proposed alignment of class C to class A receptors is compatible with many experimental data.[46,47,53,54] However, we admit that the comparison of class A/class C with class B receptors is still speculative although many of the consensus positions used herein for the clustering have been experimentally shown to be of importance for antagonist binding. The current study pinpoints relationships between a few critical positions of the 7-TM domain and ligand binding. The confomational dynamics of the cavity that is likely to play a key role in ligand recognition has not been investigated here for many reasons (irrelevance of unrestrained MD simulations regarding the paucity of 3D information about the membrane-embbeded receptor, unknown status of oligomerization, unknown membrane regulatory proteins). Studying the plasticity of the ligand–receptor interactions should thus be realized on a case-by-case basis. Our high-throughput 3D models provide a reasonable start for such studies at the condition that extracellular loops are build and joined to the 7-TM helices.

Privileged Structures of GPCR Ligands Can Be Matched with TM Cavity Hotspots

Reducing the complexity of chemogenomic data is likely to facilitate the rational design of ligands or focused libraries and to better predict selectivity towards a family of therapeutically relevant macromolecular targets. Relating cluster members to precise molecular features is here greatly facilitated by the analysis of a small subset of amino acids. For each of the 22 clusters, there is often a clear relationship between known ligand chemotypes (e.g., amines, carboxylic acids, phosphates, peptides, eicosanoids, and lipids) and the cognate TM cavities. For example, receptors for bulky ligands (e.g., phospholipids, prostanoids) have a TM cavity significantly larger than that for smaller compounds (e.g., biogenic amines, nucleotides). Receptors for charged ligands (cationic amines, phosphates, mono and di-carboxylic acids) always present, among the 30 critical residues, one or more conserved amino acid exhibiting the opposite charge (e.g., Asp$^{3.32}$ for biogenic amines; Asp$^{4.60}$/Glu$^{7.39}$ for chemokines; Arg$^{3.29}$/Lys$^{6.55}$/Arg$^{7.35}$ for nucleotides). Such complementarity has already been pointed out by several previous studies on subsets of the GPCR proteome for which numerous data exist for both receptors and cognate ligands.[19–22,31,186] However, the present report exemplifies for the first time such considerations to the entire collection of nonolfactive human GPCRs.

GPCR ligands sharing a common substructure ("privileged structure") and exhibiting promiscuous binding to unrelated GPCRs are a current important source for GPCR library design.[3,31] Assuming that conserved moieties of the ligands are likely to bind to conserved subsites of the targets,[31] matching privileged structures with TM hotspots can be achieved very easily without biasing the match by a manual or automated 3D docking. What is required is access to a GPCR ligand database[191–193] in which both ligand structures and known GPCR targets are listed.

As an example, biphenyltetrazoles and biphenylcarboxylic acids are known to bind to at least six GPCRs (AG22, AG2R, AG2S, GHSR, LT4R1, LT4R2).[181,194,195] Fine details of 3D recognition of this privileged substructure by GPCR hotspots have been recently investigated by a thorough mutagenesis-guided manual docking of several GPCR ligands.[31] We propose here a much simpler approach leading to the same outcome; looking at the 30 residues lining the TM cavity of the later six GPCRs allows us to clearly identify putative TM residues able to interact with this substructure (Fig. 25). Conserved aromatic residues are likely to interact with the biaryl moiety cluster between TMs 6 and 7 (Phe$^{6.44}$, Trp$^{6.48}$, Phe/Tyr/His$^{6.51}$, Phe/Tyr$^{7.43}$). A positively-charged residue that probably interacts with the bioisosteric tetrazole and carboxylate groups should be located nearby the aromatic cluster. Hence, three basic residues (Lys$^{5.42}$, Arg$^{6.55}$, and Arg$^{7.35}$) fulfill this requirement. Last, a polar side chain at position 6.52 (His/Gln) is conserved for the six investigated GPCRs and might H-bond to the acidic moiety of the privileged structure. By simply looking at sequence alignments of TM cavity-lining amino acids, and without relying on any 3D docking data, we managed to find out the same important anchoring residues than Bondensgaard et al.[31] Searching our TM cavity database for additional GPCRs fulfilling the above-described requirements (Phe$^{6.44}$, Trp$^{6.48}$, Phe/Tyr/His$^{6.51}$, Phe/Tyr$^{7.43}$, and Lys$^{5.42}$ or Arg$^{6.55}$ or Arg$^{7.35}$ and His/Gln$^{6.52}$) extracts 17 new GPCRs that might accommodate biphenyl-tetrazoles and biphenyl-carboxylic acids (Fig.

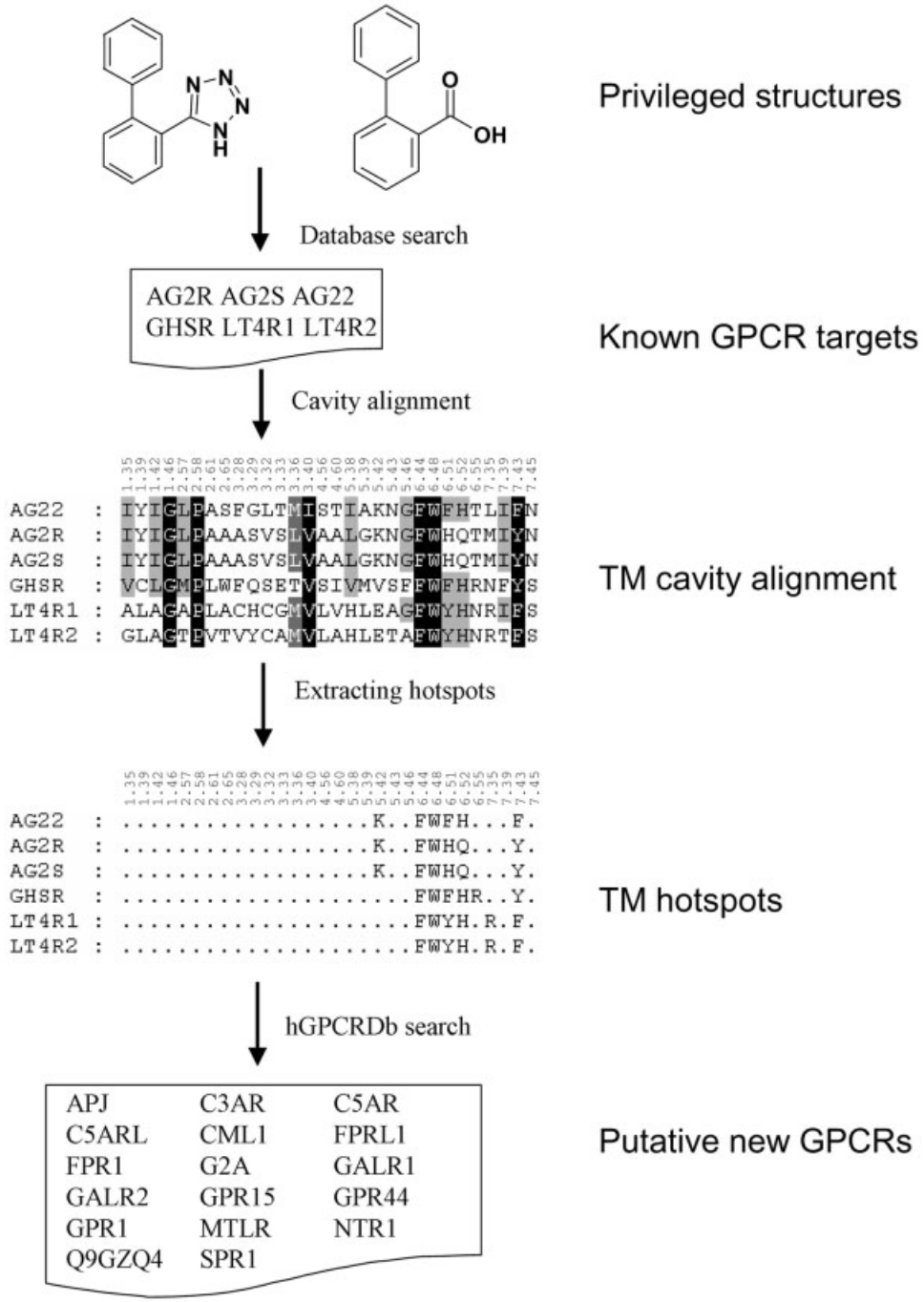


Fig. 25. Matching privileged structures of known GPCR ligands to TM hotspots. An in-house GPCR ligand database is searched to retrieve privileged structures common to multiple GPCRs and to find conserved residues within the 7-TM cavity of selected entries. Browsing the in-house GPCR cavity database (sequence of 30 critical positions lining the 7-TM cavity of 369 human GPCRs) allow to retrieve new GPCR entries satisfying the query and likely to accommodate the privileged structure.

25). Among putative targets are ten chemoattractant receptors (APJ, C3AR, C5AR, C5ARL, CML1, FPR1, FPRL1, GPR15, GPR44, and GPR1), three brain-gut peptide receptors (MTLR, NTR1, and Q9GZQ4), two cationic phospholipid receptors (G2A, SPR1) and two peptide receptors (GALR1, GALR2). This target list contains receptors recently identified by Bondensgaard et al. (e.g., APJ, NTR1).[31] It also suggests totally new putative targets for the investigated privileged structure that might serve as a common scaffold for small-sized combinatorial libraries targeting the new receptors list. Several ligand-based approaches based on known privileged structures have been reported for designing GPCR focused libraries.[196,197] However, the use of "generic" privileged structures (e.g.,

TABLE I. Possible Ligand Source for Some Orphan GPCRs

Orphan receptor(s)	Cluster	Source
GPR88, Q9NFN8	Glutamate	GABA-B allosteric ligands
Q8NHZ9, Q8TDU1	Glutamate	CaSR allosteric ligands
LRG4, LRG5, LRG6	Glycoproteins	LH/FSH nonpeptide ligands
GP119	Lipids	Cannabinoid receptors ligands
GPR19, GPR83	Peptides	Tachykinin receptors ligands
KISSR	Peptides	Galanine receptor ligands
Q6W5P4, PKR1, PKR2	Vasopeptides	Oxytocin/vasopressin receptor ligands
O14804	Amines	Biogenic amine receptors ligands
GPR39	Brain-gut peptides	Neuromedin U receptors ligands
O75307, RDC1	Chemokines	Chemokine receptor ligands
GPR7, GPR8	Opiates	Somatostatine receptor ligands
GPR15, GPR25, GPR44, GPR1	Chemoattractants	Angiotensin II receptor ligands
EBI2, GPR92, P2RY5	Purines	LPC/SPC receptor ligands
GP171, GPR87	Purines	Purinergic nucleotide receptor ligands
GPR17, GPR34, GP174	Purines	Cysteinyl Leukotriene receptor ligands

aryl-4-piperazine, biphenyl) is conceptually incompatible with the design of innovative compounds. It is therefore of outmost importance to generate novel templates by incorporating structure-based knowledge. The herein proposed approach may contribute to identify such substructures by an exhaustive prioritization of fragments according to their likelihood to bind predicted TM cavities of particular clusters.

Prioritizing Ligand/Library Selection and Design for Receptor Deorphanization

Matching TM hotspots to privileged structures may be useful for finding out ligands of orphan targets. Considering not only the target orphan receptor but also other receptors present in the same cluster enables to derive sufficient information to generate compound libraries. Receptors very close in the current phylogenetic tree present significant similarities in their TM binding cavity. A rationale source for putative ligands of an orphan receptor is thus to evaluate first known ligands of GPCRs which are the closest to the orphan target. There are still numerous orphan GPCRs spread over nearly all clusters presented in the current study (see a nonexhaustive list in Table I).

GPR88 represents an interesting example. Although rather close to dopamine D1 and D5 receptors when considering the 7-TM domain, this receptor clusters with class C GPCRs [Fig. 6(A)] when looking at the 30 residues lining the putative TM binding site. Finding starting hits for this receptor could then be addressed by evaluating first, known noncompetitive ligands for class C GPCRs, especially allosteric ligands of the GABA-B receptors. Likewise, known ligands for Angiotensin II receptors (Fig. 23) should represent good starting points for identifying putative ligands of closely related orphan targets (APJ, GPR15, GPR25, GPR44, GPR1). An experimental validation of this approach has been recently reported by the identification of GPR44 ligands by evaluating known Angiotensin II receptor ligands.[198] In the absence of any strongly conserved anchoring residues within the GPCR subfamily under investigation, a more systematic approach could be to dock a library of preselected scaffolds to pick out the best scored structures and prioritize scaffolds enriched among the top scorers when considering the target receptor or even better the whole cluster branch to which this receptor belongs to.

As a conclusion, the herein presented phylogenetic tree can be used to study the selectivity profile of either a ligand or a receptor under investigation. Knowing whether the ligand (receptor) is likely to be permissive for numerous receptors (ligands) is of crucial importance for drug discovery. The location of a precise receptor on the tree quickly indicates the number and identity of close GPCR neighbors. Looking then at the TM cavity alignment of this GPCR subset may help in the identification of selective or permissive residues that may favor or hinder design of selective ligands. Conversely, designing promiscuous ligands for either related or unrelated receptors may be guided by the analysis of their TM cavity and the identification of common anchoring amino acids that direct library design.

ACKNOWLEDGMENTS

This work is supported by grants from the French Ministry of Research and Technology, the European Union (Marie Curie fellowship HPMF-CT-2002-02141) and the Centre National de la Recherche Scientifique (CNRS, Programme Protéomique et génie des protéines). Prof. H. Schiöth (Uppsala University, Sweden) is sincerely acknowledged for critical reading of the manuscript. All data (full sequences, sequence of TM residues, sequence of cavity-lining residues, and phylogenetic tree in Newick format) on 369 human GPCRs can be obtained from the following URL: http://bioinfo-pharma.u-strasbg.fr/gpcrdb/gpcrdb_form/html

REFERENCES

1. Horn F, Bettler E, Oliveira L, Campagne F, Cohen FE, Vriend G. GPCRDB information system for G protein-coupled receptors. Nucleic Acids Res 2003;31:294–297.
2. Kristiansen K. Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled

receptors: molecular modeling and mutagenesis approaches to receptor structure and function. Pharmacol Ther 2004;103:21–80.
3. Klabunde T, Hessler G. Drug design strategies for targeting G-protein-coupled receptors. Chembiochem 2002;3:928–944.
4. Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev 2000;21:90–113.
5. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.
6. Wise A, Jupe SC, Rees S. The identification of ligands at orphan G-protein coupled receptors. Annu Rev Pharmacol Toxicol 2004;44: 43–66.
7. Kolakowski Jr. LF. GCRDb: a G-protein-coupled receptor database. Recept Channels 1994;2:1–7.
8. Attwood TK, Findlay JB. Fingerprinting G-protein-coupled receptors. Protein Eng 1994;7:195–203.
9. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001;305:567–580.
10. Joost P, Methner A. Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands. Genome Biol 2002;3:RESEARCH0063.
11. Karchin R, Karplus K, Haussler D. Classifying G-protein coupled receptors with support vector machines. Bioinformatics 2002;18: 147–159.
12. Qian B, Soyer OS, Neubig RR, Goldstein RA. Depicting a protein's two faces: GPCR classification by phylogenetic tree-based HMMs. FEBS Lett 2003;554:95–99.
13. Attwood TK. A compendium of specific motifs for diagnosing GPCR subtypes. Trends Pharmacol Sci 2001;22:162–165.
14. Elrod DW, Chou KC. A study on the correlation of G-protein-coupled receptor types with amino acid composition. Protein Eng 2002;15:713–715.
15. Lapinsh M, Gutcaits A, Prusis P, Post C, Lundstedt T, Wikberg JE. Classification of G-protein coupled receptors by alignment-independent extraction of principal chemical properties of primary amino acid sequences. Protein Sci 2002;11:795–805.
16. Fredriksson R, Lagerstrom MC, Lundin LG, Schioth HB. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. Mol Pharmacol 2003;63:1256–1272.
17. Strader CD, Sigal IS, Dixon RA. Genetic approaches to the determination of structure-function relationships of G protein-coupled receptors. Trends Pharmacol Sci 1989;Suppl:26–30.
18. Trumpp-Kallmeyer S, Hoflack J, Bruinvels A, Hibert M. Modeling of G-protein-coupled receptors: application to dopamine, adrenaline, serotonin, acetylcholine, and mammalian opsin receptors. J Med Chem 1992;35:3448–3462.
19. Jacoby E. A novel chemogenomics knowledge-based ligand design strategy—application to G Protein coupled receptors. Quant Struct-Act Relat 2001;20:115–123.
20. Bywater RP. Location and nature of the residues important for ligand recognition in G-protein coupled receptors. J Mol Recognit 2005;18:60–72.
21. Jacoby E, Schuffenhauer A, Acklin P. In: Kubinyi H, Müller G, editors. Chemogenomics in drug discovery—a medicinal chemistry perspective. Weinheim: Wiley-VCH; 2004. p 139–166.
22. Crossley R. The design of screening libraries targeted at G-protein coupled receptors. Curr Top Med Chem 2004;4:581–588.
23. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, et al. UniProt: the Universal Protein knowledgebase. Nucleic Acids Res 2004;32:D115–119.
24. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, Madden TL, Pontius JU, et al. Database resources of the National Center for Biotechnology Information: update. Nucleic Acids Res 2004;32:D35–40.
25. Bissantz C, Logean A, Rognan D. High-throughput modelling of human G-Protein coupled receptors: amino acid sequence alignment, three-dimensional model building, and receptor library screening. J Chem Info Comput Sci 2004;44:1162–1176.
26. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, et al. Crystal structure of rhodopsin: A G protein-coupled receptor. Science 2000;289:739–745.
27. Sneath PH, Sokal RR. Numerical taxonomy. Nature 1962;193:855–860.
28. Perriere G, Gouy M. WWW-query: an on-line retrieval system for biological sequence banks. Biochimie 1996;78:364–369.
29. Exner TE, Keil M, Moeckel G, Brickmann J. Identification of substrate channels and protein cavities. J Mol Model 1998;98:440–443.
30. SYBYL. 6.91., TRIPOS, Assoc., Inc., St-Louis, MO.
31. Bondensgaard K, Ankersen M, Thogersen H, Hansen BS, Wulff BS, Bywater RP. Recognition of privileged structures by G-protein coupled receptors. J Med Chem 2004;47:888–899.
32. Ballesteros JA, Shi L, Javitch JA. Structural mimicry in G protein-coupled receptors: implications of the high-resolution structure of rhodopsin for structure-function analysis of rhodopsin-like receptors. Mol Pharmacol 2001;60:1–19.
33. Miller JR. The Wnts. Genome Biol 2002;3:REVIEWS3001.
34. Williams JA, Guicherit OM, Zaharian BI, Xu Y, Chai L, Wichterle H, Kon C, et al. Identification of a small molecule inhibitor of the hedgehog signaling pathway: effects on basal cell carcinoma-like lesions. Proc Natl Acad Sci USA 2003;100:4616–4621.
35. Chen JK, Taipale J, Young KE, Maiti T, Beachy PA. Small molecule modulation of Smoothened activity. Proc Natl Acad Sci USA 2002;99:14071–14076.
36. Harmar AJ. Family-B G-protein-coupled receptors. Genome Biol 2001;2:REVIEWS3013.
37. Grace CR, Perrin MH, DiGruccio MR, Miller CL, Rivier JE, Vale WW, Riek R. NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor. Proc Natl Acad Sci USA 2004;101:12836–12841.
38. Collins JL, Dambek PJ, Goldstein SW, Faraci, W. Stephen. CP-99,711: a non-peptide glucagon receptor antagonist. Bioorg Med Chem Lett 1992;2:915–918.
39. Cascieri MA, Koch GE, Ber E, Sadowski SJ, Louizides D, de Laszlo SE, Hacker C, et al. Characterization of a novel, non-peptidyl antagonist of the human glucagon receptor. J Biol Chem 1999;274: 8694–8697.
40. Kurukulasuriya R, Sorensen BK, Link JT, Patel JR, Jae H-S, Winn MX, Rohde JR, Grihalde ND, Lin CW, Ogiela CA. Biaryl amide glucagon receptor antagonists. Bioorg Med Chem Lett 2004;14:2047–2050.
41. Hoare SR. Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. Drug Discov Today 2005; 10:417–427.
42. Liaw CW, Grigoriadis DE, Lorang MT, De Souza EB, Maki RA. Localization of agonist- and antagonist-binding domains of human corticotropin-releasing factor receptors. Mol Endocrinol 1997; 11:2048–2053.
43. Kwakkenbos MJ, Kop EN, Stacey M, Matmati M, Gordon S, Lin HH, Hamann J. The EGF-TM7 family: a postgenomic view. Immunogenetics 2004;55:655–666.
44. Bjarnadottir TK, Fredriksson R, Hoglund PJ, Gloriam DE, Lagerstrom MC, Schioth HB. The human and mouse repertoire of the adhesion family of G-protein-coupled receptors. Genomics 2004;84: 23–33.
45. Pin JP, Galvez T, Prezeau L. Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. Pharmacol Ther 2003;98:325–354.
46. Goudet C, Gaven F, Kniazeff J, Vol C, Liu J, Cohen-Gonsaud M, Acher F, Prezeau L, Pin JP. Heptahelical domain of metabotropic glutamate receptor 5 behaves like rhodopsin-like receptors. Proc Natl Acad Sci USA 2004;101:378–383.
47. Malherbe P, Kratochwil N, Knoflach F, Zenner MT, Kew JN, Kratzeisen C, Maerki HP, Adam G, Mutel V. Mutational analysis and molecular modeling of the allosteric binding site of a novel, selective, noncompetitive antagonist of the metabotropic glutamate 1 receptor. J Biol Chem 2003;278:8340–8347.
48. Petrel C, Kessler A, Maslah F, Dauban P, Dodd RH, Rognan D, Ruat M. Modeling and mutagenesis of the binding site of Calhex 231, a novel negative allosteric modulator of the extracellular Ca(2+)-sensing receptor. J Biol Chem 2003;278:49487–49494.
49. Petrel C, Kessler A, Dauban P, Dodd RH, Rognan D, Ruat M. Positive and negative allosteric modulators of the Ca2+ -sensing receptor interact within overlapping but not identical binding sites in the transmembrane domain. J Biol Chem 2004;279:18990–18997.
50. Knoflach F, Mutel V, Jolidon S, Kew JN, Malherbe P, Vieira E, Wichmann J, Kemp JA. Positive allosteric modulators of metabotropic glutamate 1 receptor: characterization, mechanism of action, and binding site. Proc Natl Acad Sci USA 2001;98:13402–13407.
51. Kniazeff J, Galvez T, Labesse G, Pin J-P. No ligand binding in the

GB2 subunit of the GABAB receptor is required for activation and allosteric interaction between the subunits. J Neurosci 2002;22: 7352–7361.
52. Binet V, Brajon C, Le Corre L, Acher F, Pin JP, Prezeau L. The heptahelical domain of GABA(B2) is activated directly by CGP7930, a positive allosteric modulator of the GABA(B) receptor. J Biol Chem 2004;279:29085–29091.
53. Urwyler S, Mosbacher J, Lingenhoehl K, Heid J, Hofstetter K, Froestl W, Bettler B, et al. Positive allosteric modulation of native and recombinant gamma-aminobutyric acid(B) receptors by 2,6-Di-tert-butyl-4-(3-hydroxy-2,2-dimethyl-propyl)-phenol (CGP7930) and its aldehyde analog CGP13501. Mol Pharmacol 2001;60:963–971.
54. Urwyler S, Pozza MF, Lingenhoehl K, Mosbacher J, Lampert C, Froestl W, Koller M, et al. N,N'-Dicyclopentyl-2-methylsulfanyl-5-nitro-pyrimidine-4,6-diamine (GS39783) and structurally related compounds: novel allosteric enhancers of gamma-aminobutyric acidB receptor function. J Pharmacol Exp Ther 2003;307:322–330.
55. Toh H, Ichikawa A, Narumiya S. Molecular evolution of receptors for eicosanoids. FEBS Lett 1995;361:17–21.
56. Regan JW, Bailey TJ, Pepperl DJ, Pierce KL, Bogardus AM, Donello JE, Fairbairn CE, et al. Cloning of a novel human prostaglandin receptor with characteristics of the pharmacologically defined EP2 subtype. Mol Pharmacol 1994;46:213–220.
57. Narumiya S, Sugimoto Y, Ushikubi F. Prostanoid receptors: structures, properties, and functions. Physiol Rev 1999;79:1193–1226.
58. Stitham J, Stojanovic A, Merenick BL, O'Hara KA, Hwa J. The unique ligand-binding pocket for the human prostacyclin receptor. Site-directed mutagenesis and molecular modeling. J Biol Chem 2003;278:4250–4257.
59. Kobayashi T, Ushikubi F, Narumiya S. Amino acid residues conferring ligand binding properties of prostaglandin I and prostaglandin D receptors. Identification by site-directed mutagenesis. J Biol Chem 2000;275:24294–24303.
60. Kobayashi T, Kiriyama M, Hirata T, Hirata M, Ushikubi F, Narumiya S. Identification of domains conferring ligand binding specificity to the prostanoid receptor. Studies on chimeric prostacyclin/prostaglandin D receptors. J Biol Chem 1997;272:15154–15160.
61. Elworthy TR, Kertesz DJ, Kim W, Roepel MG, Quattrocchio-Setti L, Smith DB, Tracy JL, et al. Lactams as EP(4) prostanoid receptor subtype selective agonists. Part 1: 2-pyrrolidinones-stereochemical and lower side-chain optimization. Bioorg Med Chem Lett 2004;14:1655–1659.
62. Mitsumori S, Tsuri T, Honma T, Hiramatsu Y, Okada T, Hashizume H, Inagaki M, et al. Synthesis and biological activity of various derivatives of a novel class of potent, selective, and orally active prostaglandin D2 receptor antagonists. 1. Bicyclo[2.2.1]heptane derivatives. J Med Chem 2003;46:2436–2445.
63. Vassart G, Pardo L, Costagliola S. A molecular dissection of the glycoprotein hormone receptors. Trends Biochem Sci 2004;29:119–126.
64. Lembo PM, Grazzini E, Groblewski T, O'Donnell D, Roy MO, Zhang J, Hoffert C, et al. Proenkephalin A gene products activate a new family of sensory neuron—specific GPCRs. Nat Neurosci 2002;5:201–209.
65. Dong X, Han S, Zylka MJ, Simon MI, Anderson DJ. A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons. Cell 2001;106:619–632.
66. Santos RA, Simoes e Silva AC, Maric C, Silva DM, Machado RP, de Buhr I, Heringer-Walther S, et al. Angiotensin-(1-7) is an endogenous ligand for the G protein-coupled receptor Mas. Proc Natl Acad Sci USA 2003;100:8258–8263.
67. Shinohara T, Harada M, Ogi K, Maruyama M, Fujii R, Tanaka H, Fukusumi S, et al. Identification of a G Protein-coupled Receptor Specifically Responsive to {beta}-Alanine. J Biol Chem 2004;279: 23559–23564..
68. Bender E, Buist A, Jurzak M, Langlois X, Baggerman G, Verhasselt P, Ercken M, et al. Characterization of an orphan G protein-coupled receptor localized in the dorsal root ganglia reveals adenine as a signaling molecule. Proc Natl Acad Sci USA 2002;99: 8573–8578
69. Han SK, Dong X, Hwang JI, Zylka MJ, Anderson DJ, Simon MI. Orphan G protein-coupled receptors MrgA1 and MrgC11 are distinctively activated by RF-amide-related peptides through the Galpha q/11 pathway. Proc Natl Acad Sci USA 2002;99:14740–14745.
70. Robas N, Mead E, Fidock M. MrgX2 is a high potency cortistatin receptor expressed in dorsal root ganglion. J Biol Chem 2003;278: 44400–44404.
71. Matsumoto M, Saito T, Takasaki J, Kamohara M, Sugimoto T, Kobayashi M, Tadokoro M, Matsumoto S, Ohishi T, Furuichi K. An evolutionarily conserved G-protein coupled receptor family, SREB, expressed in the central nervous system. Biochem Biophys Res Commun 2000;272:576–582.
72. Shen D, Jiang M, Hao W, Tao L, Salazar M, Fong HK. A human opsin-related gene that encodes a retinaldehyde-binding protein. Biochemistry 1994;33:13117–13125.
73. Fredriksson R, Hoglund PJ, Gloriam DE, Lagerstrom MC, Schioth HB. Seven evolutionarily conserved human rhodopsin G protein-coupled receptors lacking close relatives. FEBS Lett 2003;554:381–388.
74. Uhlenbrock K, Gassenhuber H, Kostenis E. Sphingosine 1-phosphate is a ligand of the human gpr3, gpr6 and gpr12 family of constitutively active G protein-coupled receptors. Cell Signal 2002;14:941–953.
75. Goto S, Okuno Y, Hattori M, Nishioka T, Kanehisa M. LIGAND: database of chemical compounds and reactions in biological pathways. Nucleic Acids Res 2002;30:402–404.
76. Wang DA, Lorincz Z, Bautista DL, Liliom K, Tigyi G, Parrill AL. A single amino acid determines lysophospholipid specificity of the S1P1 (EDG1) and LPA1 (EDG2) phospholipid growth factor receptors. J Biol Chem 2001;276:49213–49220.
77. Tigyi G. Selective ligands for lysophosphatidic acid receptor subtypes: gaining control over the endothelial differentiation gene family. Mol Pharmacol 2001;60:1161–1164.
78. Chin CN, Murphy JW, Huffman JW, Kendall DA. The third transmembrane helix of the cannabinoid receptor plays a role in the selectivity of aminoalkylindoles for CB2, peripheral cannabinoid receptor. J Pharmacol Exp Ther 1999;291:837–844.
79. Gouldson P, Calandra B, Legoux P, Kerneis A, Rinaldi-Carmona M, Barth F, Le Fur G, Ferrara P, Shire D. Mutational analysis and molecular modelling of the antagonist SR 144528 binding site on the human cannabinoid CB(2) receptor. Eur J Pharmacol 2000;401: 17–25.
80. Blaker M, Ren Y, Seshadri L, McBride EW, Beinborn M, Kopin AS. CCK-B/Gastrin receptor transmembrane domain mutations selectively alter synthetic agonist efficacy without affecting the activity of endogenous peptides. Mol Pharmacol 2000;58:399–406.
81. Escrieut C, Gigoux V, Archer E, Verrier S, Maigret B, Behrendt R, Moroder L, Bignon E, Silvente-Poirot S, Pradayrol L, Fourmy D. The biologically crucial C terminus of cholecystokinin and the non-peptide agonist SR-146,131 share a common binding Site in the Human CCK1 Receptor. Evidence for a crucial role of Met121 in the activation process. J Biol Chem 2002;277:7546–7555.
82. Berthold M, Bartfai T. Modes of peptide binding in G protein-coupled receptors. Neurochem Res 1997;22:1023–1031.
83. Berthold M, Kahl U, Jureus A, Kask K, Nordvall G, Langel U, Bartfai T. Mutagenesis and ligand modification studies on galanin binding to its GTP-binding-protein-coupled receptor GalR1. Eur J Biochem 1997;249:601–606.
84. Sautel M, Martinez R, Munoz M, Peitsch MC, Beck-Sickinger AG, Walker P. Role of a hydrophobic pocket of the human Y1 neuropeptide Y receptor in ligand binding. Mol Cell Endocrinol 1995;112: 215–222.
85. Breu V, Hashido K, Broger C, Miyamoto C, Furuichi Y, Hayes A, Kalina B, et al. Separable binding sites for the natural agonist endothelin-1 and the non-peptide antagonist bosentan on human endothelin-A receptors. Eur J Biochem 1995;231:266–270.
86. Giolitti A, Cucchi P, Renzetti AR, Rotondaro L, Zappitelli S, Maggi CA. Molecular determinants of peptide and nonpeptide NK-2 receptor antagonists binding sites of the human tachykinin NK-2 receptor by site-directed mutagenesis. Neuropharmacology 2000; 39:1422–1429.
87. Gouldson P, Legoux P, Carillon C, Delpech B, Le Fur G, Ferrara P, Shire D. Contrasting roles of leu(356) in the human CCK(1) receptor for antagonist SR 27897 and agonist SR 146131 binding. Eur J Pharmacol 1999;383:339–346.
88. Greenfeder S, Cheewatrakoolpong B, Billah M, Egan RW, Keene E, Murgolo NJ, Anthes JC. The neurokinin-1 and neurokinin-2

receptor binding sites of MDL103,392 differ. Bioorg Med Chem 1999;7:2867–2876.
89. Donohue PJ, Sainz E, Akeson M, Kroog GS, Mantey SA, Battey JF, Jensen RT, Northup JK. An aspartate residue at the extracellular boundary of TMII and an arginine residue in TMVII of the gastrin-releasing peptide receptor interact to facilitate heterotrimeric G protein coupling. Biochemistry 1999;38:9366–9372.
90. Akeson M, Sainz E, Mantey SA, Jensen RT, Battey JF. Identification of four amino acids in the gastrin-releasing peptide receptor that are required for high affinity agonist binding. J Biol Chem 1997;272:17405–17409.
91. Tokita K, Katsuno T, Hocart SJ, Coy DH, Llinares M, Martinez J, Jensen RT. Molecular basis for selectivity of high affinity peptide antagonists for the gastrin-releasing peptide receptor. J Biol Chem 2001;276:36652–36663.
92. Lin Y, Jian X, Lin Z, Kroog GS, Mantey S, Jensen RT, Battey J, Northup J. Two amino acids in the sixth transmembrane segment of the mouse gastrin-releasing peptide receptor are important for receptor activation. J Pharmacol Exp Ther 2000;294:1053–1062.
93. Tokita K, Hocart SJ, Katsuno T, Mantey SA, Coy DH, Jensen RT. Tyrosine 220 in the 5th transmembrane domain of the neuromedin B receptor is critical for the high selectivity of the peptoid antagonist PD168368. J Biol Chem 2001;276:495–504.
94. Wu C, Chan MF, Stavros F, Raju B, Okun I, Castillo RS. Structure-activity relationships of N2-aryl-3-(isoxazolylsulfamoyl)-2-thiophenecarboxamides as selective endothelin receptor-A antagonists. J Med Chem 1997;40:1682–1689.
95. Funk OF, Kettmann V, Drimal J, Langer T. Chemical function based pharmacophore generation of endothelin-A selective receptor antagonists. J Med Chem 2004;47:2750–2760.
96. Lee JA, Brinkmann JA, Longton ED, Peishoff CE, et al. Lysine 182 of endothelin B receptor modulates agonist selectivity and antagonist affinity: evidence for the overlap of peptide and nonpeptide ligand binding sites. Biochemistry 1994;33:14543–14549.
97. Webb ML, Patel PS, Rose PM, Liu EC, Stein PD, Barrish J, Lach DA, et al. Mutational analysis of the endothelin type A receptor (ETA): interactions and model of selective ETA antagonist BMS-182874 with putative ETA receptor binding cavity. Biochemistry 1996;35:2548–2556.
98. Huang RR, Yu H, Strader CD, Fong TM. Interaction of substance P with the second and seventh transmembrane domains of the neurokinin-1 receptor. Biochemistry 1994;33:3007–3013.
99. Bhogal N, Donnelly D, Findlay JB. The ligand binding site of the neurokinin 2 receptor. Site-directed mutagenesis and identification of neurokinin A binding residues in the human neurokinin 2 receptor. J Biol Chem 1994;269:27269–27274.
100. Labrou NE, Bhogal N, Hurrell CR, Findlay JB. Interaction of Met297 in the seventh transmembrane segment of the tachykinin NK2 receptor with neurokinin A. J Biol Chem 2001;276:37944–37949.
101. Renzetti AR, Catalioto RM, Criscuoli M, Cucchi P, Ferrer C, Giolitti A, Guelfi M, et al. Relevance of aromatic residues in transmembrane segments V to VII for binding of peptide and nonpeptide antagonists to the human tachykinin NK(2) receptor. J Pharmacol Exp Ther 1999;290:487–495.
102. Sautel M, Rudolf K, Wittneben H, Herzog H, Martinez R, Munoz M, Eberlein W, et al. Neuropeptide Y and the nonpeptide antagonist BIBP 3226 share an overlapping binding site at the human Y1 receptor. Mol Pharmacol 1996;50:285–292.
103. Kannoa T, Kanatani A, Keen SL, Arai-Otsuki S, Haga Y, Iwama T, Ishihara A, et al. Different binding sites for the neuropeptide Y Y1 antagonists 1229U91 and J-104870 on human Y1 receptors. Peptides 2001;22:405–413.
104. Conway S, Canning SJ, Barrett P, Guardiola-Lemaitre B, Delagrange P, Morgan PJ. The roles of valine 208 and histidine 211 in ligand binding and receptor function of the ovine Mel1a beta melatonin receptor. Biochem Biophys Res Commun 1997;239:418–423.
105. Conway S, Drew JE, Mowat ES, Barrett P, Delagrange P, Morgan PJ. Chimeric melatonin mt1 and melatonin-related receptors. Identification of domains and residues participating in ligand binding and receptor activation of the melatonin mt1 receptor. J Biol Chem 2000;275:20602–20609.
106. Kokkola T, Foord SM, Watson MA, Vakkuri O, Laitinen JT. Important amino acids for the function of the human MT1 melatonin receptor. Biochem Pharmacol 2003;65:1463–1471.
107. Kokkola T, Watson MA, White J, Dowell S, Foord SM, Laitinen JT. Mutagenesis of human Mel1a melatonin receptor expressed in yeast reveals domains important for receptor function. Biochem Biophys Res Commun 1998;249:531–536.
108. Gubitz AK, Reppert SM. Chimeric and point-mutated receptors reveal that a single glycine residue in transmembrane domain 6 is critical for high affinity melatonin binding. Endocrinology 2000;141:1236–1244.
109. Gerdin MJ, Mseeh F, Dubocovich ML. Mutagenesis studies of the human MT2 melatonin receptor. Biochem Pharmacol 2003;66:315–320.
110. Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, Makela S, et al. Characterization of a Common Susceptibility Locus for Asthma-Related Traits. Science 2004;304:300–304.
111. Hibert M, Hoflack J, Trumpp-Kallmeyer S, Mouillac B, Chini B, Mahe E, Cotte N, et al. Functional architecture of vasopressin/oxytocin receptors. J Recept Signal Transduct Res 1999;19:589–596.
112. Phalipou S, Seyer R, Cotte N, Breton C, Barberis C, Hibert M, Mouillac B. Docking of linear peptide antagonists into the human V(1a) vasopressin receptor. Identification of binding domains by photoaffinity labeling. J Biol Chem 1999;274:23316–23327.
113. Tahtaoui C, Balestre MN, Klotz P, Rognan D, Barberis C, Mouillac B, Hibert M. Identification of the binding sites of the SR49059 nonpeptide antagonist into the V1a vasopressin receptor using sulfydryl-reactive ligands and cysteine mutants as chemical sensors. J Biol Chem 2003;278:40010–40019.
114. Derick S, Pena A, Durroux T, Wagnon J, Serradeil-Le Gal C, Hibert M, Rognan D, et al. Key amino acids located within the transmembrane domains 5 and 7 account for the pharmacological specificity of the human V1b vasopressin receptor. Mol Endocrinol 2004;18:2777–2789.
115. Murugesan N, Tellew JE, Gu Z, Kunst BL, Fadnis L, Cornelius LA, Baska RA, et al. Discovery of N-isoxazolyl biphenylsulfonamides as potent dual angiotensin II and endothelin A receptor antagonists. J Med Chem 2002;45:3829–3835.
116. Kakefuda A, Suzuki T, Tobe T, Tsukada J, Tahara A, Sakamoto S, Tsukamoto S. Synthesis and pharmacological evaluation of 5-(4-biphenyl)-3-methyl-4-phenyl-1,2,4-triazole derivatives as a novel class of selective antagonists for the human vasopressin V(1A) receptor. J Med Chem 2002;45:2589–2598.
117. Pitt G, Batt A, Haigh R, Penson A, Robson P, Rooker D, Tartar A, et. al. Non-peptide oxytocin agonists. Bioorg Med Chem Lett 2004;14:4585–4589.
118. Flanagan CA, Millar RP, Illing N. Advances in understanding gonadotrophin-releasing hormone receptor structure and ligand interactions. Rev Reprod 1997;2:113–120.
119. Hoffmann SH, ter Laak T, Kuhne R, Reilander H, Beckers T. Residues within transmembrane helices 2 and 5 of the human gonadotropin-releasing hormone receptor contribute to agonist and antagonist binding. Mol Endocrinol 2000;14:1099–1115.
120. Hovelmann S, Hoffmann SH, Kuhne R, ter Laak T, Reilander H, Beckers T. Impact of aromatic residues within transmembrane helix 6 of the human gonadotropin-releasing hormone receptor upon agonist and antagonist binding. Biochemistry 2002;41:1129–1136.
121. Jiang J, DeVita RJ, Goulet MT, Wyvratt MJ, Lo JL, Ren N, Yudkovitz JB, et al. Syntheses and structure-activity relationship studies of piperidine-substituted quinolones as nonpeptide gonadotropin releasing hormone antagonists. Bioorg Med Chem Lett 2004;14:1795–1798.
122. Gao ZG, Chen A, Barak D, Kim SK, Muller CE, Jacobson KA. Identification by site-directed mutagenesis of residues involved in ligand recognition and activation of the human A3 adenosine receptor. J Biol Chem 2002;277:19056–19063.
123. Barbhaiya H, McClain R, Ijzerman A, Rivkees SA. Site-directed mutagenesis of the human A1 adenosine receptor: influences of acidic and hydroxy residues in the first four transmembrane domains on ligand binding. Mol Pharmacol 1996;50:1635–1642.
124. Rivkees SA, Barbhaiya H, AP IJ. Identification of the adenine binding site of the human A1 adenosine receptor. J Biol Chem 1999;274:3617–3621.
125. Cho W, Taylor LP, Mansour A, Akil H. Hydrophobic residues of the D2 dopamine receptor are important for binding and signal transduction. J Neurochem 1995;65:2105–2115.
126. Javitch JA, Ballesteros JA, Weinstein H, Chen J. A cluster of aromatic residues in the sixth membrane-spanning segment of

the dopamine D2 receptor is accessible in the binding-site crevice. Biochemistry 1998;37:998–1006.
127. Roth BL, Willins DL, Kristiansen K, Kroeze WK. 5-Hydroxytryptamine2-family receptors (5-hydroxytryptamine2A, 5-hydroxytryptamine2B, 5-hydroxytryptamine2C): where structure meets function. Pharmacol Ther 1998;79:231–257.
128. Joubert L, Claeysen S, Sebben M, Bessis AS, Clark RD, Martin RS, Bockaert J, et al. A 5-HT4 receptor transmembrane network implicated in the activity of inverse agonists but not agonists. J Biol Chem 2002;277:25502–25511.
129. Shi L, Liapakis G, Xu R, Guarnieri F, Ballesteros JA, Javitch JA. Beta2 adrenergic receptor activation. Modulation of the proline kink in transmembrane 6 by a rotamer toggle switch. J Biol Chem 2002;277:40989–40996.
130. Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, Gether U, Javitch JA. Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. J Biol Chem 2001;276: 29171–29177.
131. Kobilka B. Agonist binding: a multistep process. Mol Pharmacol 2004;65:1060–1062.
132. Wess J, Blin N, Mutschler E, Bluml K. Muscarinic acetylcholine receptors: structural basis of ligand binding and G protein coupling. Life Sci 1995;56:915–922.
133. Wieland K, Zuurmond HM, Krasel C, Ijzerman AP, Lohse MJ. Involvement of Asn-293 in stereospecific agonist recognition and in activation of the beta 2-adrenergic receptor. Proc Natl Acad Sci USA 1996;93:9276–9281.
134. Shin N, Coates E, Murgolo NJ, Morse KL, Bayne M, Strader CD, Monsma Jr FJ. Molecular modeling and site-specific mutagenesis of the histamine-binding site of the histamine H4 receptor. Mol Pharmacol 2002;62:38–47.
135. Gantz I, DelValle J, Wang LD, Tashiro T, Munzert G, Guo YJ, Konda Y, et al. Molecular basis for the interaction of histamine with the histamine H2 receptor. J Biol Chem 1992;267:20840–20843.
136. Wieland K, Laak AM, Smit MJ, Kuhne R, Timmerman H, Leurs R. Mutational analysis of the antagonist-binding site of the histamine H(1) receptor. J Biol Chem 1999;274:29994–30000.
137. Bissantz C, Bernard P, Hibert M, Rognan D. Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets? Proteins 2003; 50:5–25.
138. Javitch JA, Ballesteros JA, Chen J, Chiappa V, Simpson MM. Electrostatic and aromatic microdomains within the binding-site crevice of the D2 receptor: contributions of the second membrane-spanning segment. Biochemistry 1999;38:7961–7968.
139. Simpson MM, Ballesteros JA, Chiappa V, Chen J, Suehiro M, Hartman DS, Godel T, et al. Dopamine D4/D2 receptor selectivity is determined by A divergent aromatic microdomain contained within the second, third, and seventh membrane-spanning segments. Mol Pharmacol 1999;56:1116–1126.
140. Nijenhuis WAJ, Oosterom J, Adan RAH. AgRP(83-132) acts as an inverse agonist on the human-melanocortin-4 receptor. Mol Endocrinol 2001;15:164–171.
141. Yang YK, Fong TM, Dickinson CJ, Mao C, Li JY, Tota MR, Mosley R, et al. Molecular determinants of ligand binding to the human melanocortin-4 receptor. Biochemistry 2000;39:14900–14911.
142. Nickolls SA, Cismowski MI, Wang X, Wolff M, Conlon PJ, Maki RA. Molecular determinants of melanocortin 4 receptor ligand binding and MC4/MC3 receptor selectivity. J Pharmacol Exp Ther 2003;304:1217–1227.
143. Schioth HB, Muceniece R, Szardenings M, Prusis P, Lindeberg G, Sharma SD, Hruby VJ, et al. Characterisation of D117A and H260A mutations in the melanocortin 1 receptor. Mol Cell Endocrinol 1997;126:213–219.
144. Yang Y, Dickinson C, Haskell-Luevano C, Gantz I. Molecular basis for the interaction of [Nle4,D-Phe7]melanocyte stimulating hormone with the human melanocortin-1 receptor. J Biol Chem 1997;272:23000–23010.
145. Chen C, Yu J, Fleck BA, Hoare SRJ, Saunders J, Foster AC. Phenylguanidines as selective onpeptide melanocortin-5 receptor antagonists. J Med Chem 2004;47:4083–4088.
146. McKee KK, Tan CP, Palyha OC, Liu J, Feighner SD, Hreniuk DL, Smith RG, et al. Cloning and characterization of two human G protein-coupled receptor genes (GPR38 and GPR39) related to the growth hormone secretagogue and neurotensin receptors. Genomics 1997;46:426–434.
147. Feighner SD, Howard AD, Prendergast K, Palyha OC, Hreniuk DL, Nargund R, Underwood D, et al. Structural requirements for the activation of the human growth hormone secretagogue receptor by peptide and nonpeptide secretagogues. Mol Endocrinol 1998;12:137–145.
148. Labbe-Jullie C, Barroso S, Nicolas-Eteve D, Reversat JL, Botto JM, Mazella J, Bernassau JM, et al. Mutagenesis and modeling of the neurotensin receptor NTR1. Identification of residues that are critical for binding SR 48692, a nonpeptide neurotensin antagonist. J Biol Chem 1998;273:16351–16357.
149. Colson AO, Perlman JH, Jinsi-Parimoo A, Nussenzveig DR, Osman R, Gershengorn MC. A hydrophobic cluster between transmembrane helices 5 and 6 constrains the thyrotropin-releasing hormone receptor in an inactive conformation. Mol Pharmacol 1998;54:968–978.
150. Perlman JH, Laakkonen L, Osman R, Gershengorn MC. Distinct roles for arginines in transmembrane helices 6 and 7 of the thyrotropin-releasing hormone receptor. Mol Pharmacol 1995;47: 480–484.
151. Perlman JH, Laakkonen LJ, Guarnieri F, Osman R, Gershengorn MC. A refined model of the thyrotropin-releasing hormone (TRH) receptor binding pocket. Experimental analysis and energy minimization of the complex between TRH and TRH receptor. Biochemistry 1996;35:7643–7650.
152. Perlman JH, Thaw CN, Laakkonen L, Bowers CY, Osman R, Gershengorn MC. Hydrogen bonding interaction of thyrotropin-releasing hormone (TRH) with transmembrane tyrosine 106 of the TRH receptor. J Biol Chem 1994;269:1610–1613.
153. Clark DE, Higgs C, Wren SP, Dyke HJ, Wong M, Norman D, Lockey PM, Roach AG. A virtual screening approach to finding novel and potent anatgonists at the melanin-concentrating hormone 1 receptor. J Med Chem 2004;47:3962–3971.
154. Macdonald D, Murgolo N, Zhang R, Durkin JP, Yao X, Strader CD, Graziano MP. Molecular characterization of the melanin-concentrating hormone/receptor complex: identification of critical residues involved in binding and activation. Mol Pharmacol 2000;58:217–225.
155. Receveur JM, Bjurling E, Ulven T, Little PB, Norregaard PK, Hogberg T. 4-Acylamino-and 4-ureidobenzamides as melanin-concentrating hormone (MCH) receptor 1 antagonists. Bioorg Med Chem Lett 2004;14:5075–5080.
156. Tunaru S, Kero J, Schaub A, Wufka C, Blaukat A, Pfeffer K, Offermanns S. PUMA-G and HM74 are receptors for nicotinic acid and mediate its anti-lipolytic effect. Nat Med 2003;9:352–355.
157. Jones CE, Holden S, Tenaillon L, Bhatia U, Seuwen K, Tranter P, Turner J, et al. Expression and characterization of a 5-oxo-6E,8Z,11Z,14Z-eicosatetraenoic acid receptor highly expressed on human eosinophils and neutrophils. Mol Pharmacol 2003;63: 471–477.
158. Dragic T, Trkola A, Thompson DA, Cormier EG, Kajumo FA, Maxwell E, Lin SW, et al.. A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5. Proc Natl Acad Sci USA 2000;97:5639–5644.
159. Mirzadegan T, Diehl F, Ebi B, Bhakta S, Polsky I, McCarley D, Mulkins M, et al. Identification of the binding site for a novel class of CCR2b chemokine receptor antagonists: binding to a common chemokine receptor motif within the helical bundle. J Biol Chem 2000;275:25562–25571.
160. Hatse S, Princen K, Gerlach LO, Bridger G, Henson G, De Clercq E, Schwartz TW, et al. Mutation of Asp(171) and Asp(262) of the chemokine receptor CXCR4 impairs its coreceptor function for human immunodeficiency virus-1 entry and abrogates the antagonistic activity of AMD3100. Mol Pharmacol 2001;60:164–173.
161. Castonguay LA, Weng Y, Adolfsen W, Di Salvo J, Kilburn R, Caldwell CG, Daugherty BL, et al. Binding of 2-aryl-4-(piperidin-1-yl)butanamines and 1,3,4-trisubstituted pyrrolidines to human CCR5: a molecular modeling-guided mutagenesis study of the binding pocket. Biochemistry 2003;42:1544–1550.
162. Patel YC, Greenwood MT, Panetta R, Demchyshyn L, Niznik H, Srikant CB. The somatostatin receptor family. Life Sci 1995;57: 1249–1265.
163. Simon A, Czajlik A, Perczel A, Keri G, Nyikos L, Emri Z, Kardos J. Binding crevice for TT-232 in a homology model of type 1

somatostatin receptor. Biochem Biophys Res Commun 2004;316: 1059–1064.
164. Kaupmann K, Bruns C, Raulf F, Weber HP, Mattes H, Lubbert H. Two amino acids, located in transmembrane domains VI and VII, determine the selectivity of the peptide agonist SMS 201-995 for the SSTR2 somatostatin receptor. EMBO J 1995;14:727–735.
165. Rohrer SP, Birzin ET, Mosley RT, Berk SC, Hutchins SM, Shen DM, Xiong Y, et al.Rapid identification of subtype-selective agonists of the somatostatin receptor through combinatorial chemistry. Science 1998;282:737–740.
166. Poitout L, Roubert P, Contour-Galcera MO, Moinet C, Lannoy J, Pommier J, Plas P, et al. Identification of potent non-peptide somatostatin antagonists with sst(3) selectivity. J Med Chem 2001;44:2990–3000.
167. Befort K, Tabbara L, Bausch S, Chavkin C, Evans C, Kieffer B. The conserved aspartate residue in the third putative transmembrane domain of the delta-opioid receptor is not the anionic counterpart for cationic opiate binding but is a constituent of the receptor binding site. Mol Pharmacol 1996;49:216–223.
168. Befort K, Tabbara L, Kling D, Maigret B, Kieffer BL. Role of aromatic transmembrane residues of the delta-opioid receptor in ligand recognition. J Biol Chem 1996;271:10161–10168.
169. Meng F, Ueda Y, Hoversten MT, Taylor LP, Reinscheid RK, Monsma FJ, Watson SJ, et al. Creating a functional opioid alkaloid binding site in the orphanin FQ receptor through site-directed mutagenesis. Mol Pharmacol 1998;53:772–777.
170. Chen Z, Zhang X, Gonnella NC, Pellas TC, Boyar WC, Ni F. Residues 21-30 within the extracellular N-terminal region of the C5a receptor represent a binding domain for the C5a anaphylatoxin. J Biol Chem 1998;273:10411–10419.
171. DeMartino JA, Konteatis ZD, Siciliano SJ, Van Riper G, Underwood DJ, Fischer PA, Springer MS. Arginine 206 of the C5a receptor is critical for ligand recognition and receptor activation by C-terminal hexapeptide analogs. J Biol Chem 1995;270:15966–15969.
172. Gerber BO, Meng EC, Dotsch V, Baranski TJ, Bourne HR. An activation switch in the ligand binding pocket of the C5a receptor. J Biol Chem 2001;276:3394–3400.
173. Cain SA, Coughlan T, Monk PN. Mapping the ligand-binding site on the C5a receptor: arginine74 of C5a contacts aspartate282 of the C5a receptor. Biochemistry 2001;40:14047–14052.
174. de Laszlo SE, Allen EE, Li B, Ondeyka D, Rivero R, Malkowitz L, Molineaux C, et al. A nonpeptidic agonist ligand of the human C5a receptor: synthesis, binding affinity optimization and functional characterization. Bioorg Med Chem Lett 1997;7:213–218.
175. Astles PC, Brown TJ, Cox P, Halley F, Lockey PM, McCarthy C, McLay IM, et al. New non peptidic C5a receptor antagonists. BioorgMed Chem Lett 1997;7:907–912.
176. Mills JS, Miettinen HM, Cummings D, Jesaitis AJ. Characterization of the binding site on the formyl peptide receptor using three receptor mutants and analogs of Met-Leu-Phe and Met-Met-Trp-Leu-Leu. J Biol Chem 2000;275:39012–39017.
177. Chen X, Yang D, Shen W, Dong HF, Wang JM, Opppenheim JJ, Howard OMZ. Characterization of chenodeoxycholic acid as an endogenous antagonist of the G-coupled formyl peptide receptors. Inflamm Res 2000;49:744–755.
178. Liang TS, Gao J-L, Fatemi O, Lavigne M, Leto TL, Murphy PM. The endogenous opioid spinorphin blocks fMet-Leu-Phe-induced neutrophil chemotaxis by acting as a specific antagonist at the N-formylpeptide receptor subtype FPR. J Immunol 2001;167: 6609–6614.
179. Dalpiaz A, Ferretti ME, Vertuani G, Traniello S, Scatturin A, Spisani S. C- and N-terminal residue effect on peptide derivatives' antagonism toward the formyl-peptide receptor. Eur J Pharmacol 2002;436:187–196.
180. Fierens FL, Vanderheyden PM, Gaborik Z, Minh TL, Backer JP, Hunyady L, Ijzerman A, et al. Lys(199) mutation of the human angiotensin type 1 receptor differentially affects the binding of surmountable and insurmountable non-peptide antagonists. J Renin Angiotensin Aldosterone Syst 2000;1:283–288.
181. Ji H, Leung M, Zhang Y, Catt KJ, Sandberg K. Differential structural requirements for specific binding of nonpeptide and peptide antagonists to the AT1 angiotensin receptor. Identification of amino acid residues that determine binding of the antihypertensive drug losartan. J Biol Chem 1994;269:16533–16536.
182. Marie J, Richard E, Pruneau D, Paquet JL, Siatka C, Larguier R, Ponce C, et al. Control of conformational equilibria in the human B2 bradykinin receptor. Modeling of nonpeptidic ligand action and comparison to the rhodopsin structure. J Biol Chem 2001;276: 41100–41111.
183. Bascands JL, Schanstra JP, Couture R, Girolami JP. Bradykinin receptors: towards new pathophysiological roles. Med Sci (Paris) 2003;19:1093–1100.
184. Fathy DB, Mathis SA, Leeb T, Leeb-Lundberg LM. A single position in the third transmembrane domains of the human B1 and B2 bradykinin receptors is adjacent to and discriminates between the C-terminal residues of subtype-selective ligands. J Biol Chem 1998;273:12210–12218.
185. He W, Miao FJ, Lin DC, Schwandner RT, Wang Z, Gao J, Chen JL, et al. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. Nature 2004;429:188–193.
186. Costanzi S, Mamedova L, Gao ZG, Jacobson KA. Architecture of P2Y nucleotide receptors: structural comparison based on sequence analysis, mutagenesis, and homology modeling. J Med Chem 2004;47:5393–5404.
187. Jiang Q, Guo D, Lee BX, Van Rhee AM, Kim YC, Nicholas RA, Schachter JB, et al. A mutational analysis of residues essential for ligand recognition at the human P2Y1 receptor. Mol Pharmacol 1997;52:499–507.
188. Moro S, Guo D, Camaioni E, Boyer JL, Harden TK, Jacobson KA. Human P2Y1 receptor: molecular modeling and site-directed mutagenesis as tools to identify agonist and antagonist recognition sites. J Med Chem 1998;41:1456–1466.
189. Noguchi K, Ishii S, Shimizu T. Identification of p2y9/GPR23 as a novel G protein-coupled receptor for lysophosphatidic acid, structurally distant from the Edg family. J Biol Chem 2003;278:25600–25606.
190. Zhang HC, McComsey DF, White KB, Addo MF, Andrade-Gordon P, Derian CK, Oksenberg D, et al. Thrombin receptor (PAR-1) antagonists. Solid-phase synthesis of indole-based peptide mimetics by anchoring to a secondary amide. Bioorg Med Chem Lett 2001;11:2105–2109.
191. Schuffenhauer A, Zimmermann J, Stoop R, van der Vyver JJ, Lecchini S, Jacoby E. An ontology for pharmaceutical ligands and its application for in silico screening and library design. J Chem Inf Comput Sci 2002;42:947–955.
192. Balakin KV, Tkachenko SE, Lang SA, Okun I, Ivashchenko AA, Savchuk NP. Property-based design of GPCR-targeted library. J Chem Inf Comput Sci 2002;42:1332–1342.
193. von Korff M, Steger M. GPCR-tailored pharmacophore pattern recognition of small molecular ligands. J Chem Inf Comput Sci 2004;44:1137–1147.
194. Smith RG, Cheng K, Schoen WR, Pong SS, Hickey G, Jacks T, Butler B, et al. A nonpeptidyl growth hormone secretagogue. Science 1993;260:1640–1643.
195. Reiter LA, Koch K, Piscopio AD, Showell HJ, Alpert R, Biggers MS, Chambers RJ, et al. trans-3-Benzyl-4-hydroxy-7-chromanylbenzoic acid derivatives as antagonists of the leukotriene B4 (LTB4) receptor. Bioorg Med Chem Lett 1998;8:1781–1786.
196. Guo T, Hobbs DW. Privileged structure-based combinatorial libraries targeting G protein-coupled receptors. Assay Drug Dev Technol 2003;1:579–592.
197. Bleicher KH, Bohm HJ, Muller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. Nat Rev Drug Discov 2003;2:369–378.
198. Frimurer T, Ulven T, Elling CE, Gerlach LO, Kostenis E, Högberg T. A physicogenetic method to assign ligand-binding relationsships between 7TM receptors. Bioorg Med Chem Lett 2005;15:3707–3712.