EMBO MEMBERS’ REVIEW

Molecular tinkering of G protein-coupled receptors: an evolutionary success

Joel Bockaert1 and Jean Philippe Pin

UPR CNRS 9023, CCIPE, 141 rue de la Cardonille, 34094 Montpellier Cedex 5, France
1Corresponding author
e-mail: bockaert@ccipe.montp.inserm.fr

Among membrane-bound receptors, the G protein-coupled receptors (GPCRs) are certainly the most diverse. They have been very successful during evolution, being capable of transducing messages as different as photons, organic odorants, nucleotides, nucleosides, peptides, lipids and proteins. Indirect studies, as well as two-dimensional crystallization of rhodopsin, have led to a useful model of a common ‘central core’, composed of seven transmembrane helical domains, and its structural modifications during activation. There are at least six families of GPCRs showing no sequence similarity. They use an amazing number of different domains both to bind their ligands and to activate G proteins. The fine-tuning of their coupling to G proteins is regulated by splicing, RNA editing and phosphorylation. Some GPCRs have been found to form either homo- or heterodimers with a structurally different GPCR, but also with membrane-bound proteins having one transmembrane domain such as nina-A, odr-4 or RAMP, the latter being involved in their targeting, function and pharmacology. Finally, some GPCRs are unfaithful to G proteins and interact directly, via their C-terminal domain, with proteins containing PDZ and Enabled/VASP homology (EVH)-like domains.

Keywords: GPCRs/G proteins/signal transduction/transmembrane domains

Introduction

The evolution of multicellular organisms has been highly dependent on the capacity developed by their cells to communicate with each other and with their environment. One of the surprises of the last 15 years was the discovery that the membrane-bound receptors, dedicated to recognizing intercellular messenger molecules (such as hormones, neurotransmitters, growth and developmental factors), and several sensory messages (such as light, odors and gustative molecules), belong to four or five protein families; the most common one is the G protein-coupled receptor (GPCR) family. The diversity of structures and functions of GPCRs is a perfect illustration of F.Jacob’s famous idea: ‘evolution is molecular tinkering’. In vertebrates, this family contains 1000–2000 members (>1% of the genome) including >1000 coding for odorant and pheromone receptors. Similarly, the Caenorhabditis elegans genome encodes ~1100 GPCRs (5% of the genome, which is now entirely known) (Bargmann, 1998). In this organism, this is the family of genes which comprises the most members. GPCRs are certainly among the oldest devices devoted to signal transduction being present in plants (Plakidou-Dymock et al., 1998), yeast (Dohlman et al., 1991) and slime mold (Dictyostelium discoideum) (Devreotes, 1994), as well as in protozoa and the earliest diploblastic metazoa (Vernier et al., 1995; New and Wong, 1998). However, the range of sequence data is currently too small to be able to make reasonable inference as to the evolutionary emergence of most of the GPCRs present in vertebrates.

GPCRs are involved in the recognition and transduction of messages as diverse as light, Ca2+, odors, small molecules including amino-acid residues, nucleotides and peptides, as well as proteins (Figure 1A). They control the activity of enzymes, ion channels and transport of vesicles via the catalysis of the GDP–GTP exchange on heterotrimeric G proteins (Gα–βγ) (Figure 1A).

Many features of GPCR structures and functions have been reviewed recently (Bourne, 1997; Wess, 1997; Bockaert and Pin, 1998; Hamm, 1998). Here, we will highlight some recent data that may change our classical view of GPCR structures and functions.

A common core domain involved in G-protein recognition and activation

Sequence comparison between the different GPCRs revealed the existence of different receptor families sharing no sequence similarity (Figures 1 and 2). However, all these receptors have in common a central core domain constituted of seven transmembrane helices (TM-I through -VII) connected by three intracellular (i1, i2 and i3) and three extracellular (e1, e2 and e3) loops (Baldwin, 1993). Two cysteine residues (one in e1 and one in e2) which are conserved in most GPCRs, form a disulfide link which is probably important for the packing and for the stabilization of a restricted number of conformations of these seven TMs. Aside from sequence variations, GPCRs differ in the length and function of their N-terminal extracellular domain, their C-terminal intracellular domain and their intracellular loops. Each of these domains provide specific properties to these various receptor proteins, as discussed below.

The seven TM (7TM) region constitutes the core domain of these receptors, and a change in conformation of this domain is probably responsible for receptor activation (Figure 1B). Although no high resolution structure of such a domain has been determined yet, a low resolution (9 Å) electron diffraction structure of rhodopsin revealed the orientation of the transmembrane α-helices (Unger et al., 1997) (Figure 1B). Mutagenesis and biochemical analysis with model GPCRs like rhodopsin revealed that the switch
from the inactive to the active conformation is associated to a change in the relative orientation of TM-III and TM-VI (with a rotation of TM-VI and a separation from TM-III), which unmasks G protein-binding sites (Farrens et al., 1997; Bourne, 1997; Javitch et al., 1997). In family 1 GPCRs, one residue (Asp) in TM-II and a tripeptide (DRY or ERW) at the interface of TM-III and i2 (Figure 2A) are important for receptor activation (Oliviera et al., 1994; Scheer et al., 1996). Since these residues are not conserved in the other GPCR families, one may conclude that either the change in conformation of the core domain, or the molecular events leading to these changes are not conserved between members of these different families.

The change in conformation of the core domain generally affects the conformation of the i2 and i3 intracellular loop (which are directly linked to TM-III and TM-VI, respectively; see Figure 1B) that constitutes one of the key sites for G-protein recognition and activation (Spengler et al., 1993; Pin and Bockaert, 1995; Wess, 1997). In GnRH receptors, the i1 loop is essential for activation of Gs but not Gq (Arora et al., 1998). This perfectly illustrates the diversity of the solutions which have been selected during evolution to assume a good coupling to G proteins. It has been proposed that the C-terminal end of the G protein α-subunit binds in a pocket constituted by these intracellular loops in the various GPCR families.

**A large variety of molecular mechanisms allows the diverse ligands to activate the core domain**

Although we propose that a similar change in conformation of the core domain is associated with GPCR activation, a large diversity of molecular mechanisms have been selected during evolution to allow the natural ligand to induce this change in conformation (Figure 2). In the family 1 GPCRs activated by small ligands like catecholamines (subfamily 1a), the ligands bind in a cavity formed by TM-III to TM-VI. In the case of the light-activated receptor, rhodopsin, the target of photons, retinal, is covalently linked in this cavity, and its change in conformation induced by light activates the receptor. Other family 1 GPCRs (subfamily 1b) are activated by short peptides which interact with the extracellular loops and the N-terminal domain. However, the C-terminal end of these peptides has been proposed to interact within a cavity similar to that of the subfamily 1a GPCRs (Trumpp-Kallmeyer et al., 1995). For the family 2 receptors, which are activated by large peptides like glucagon or secretin, VIP or PACAP, the relatively long N-terminal domain also plays a role in the binding of the ligand (Pantaloni et al., 1996). Some GPCRs can also be activated by large proteins such as the subfamily 1c GPCRs. In this case, a large N-terminal extracellular domain recognizes and binds the glycoproteins, and allows them to activate the core domain via its interaction with e1 and e3 loops (Ji and Ji, 1995; Fernandez and Puett, 1996). The example of the family 3 receptors is very original. This receptor family comprises the metabotropic glutamate receptors (mGluRs), a receptor activated by extracellular Ca$^{2+}$ (Ca-sensing receptor), a group of putative pheromone receptors (termed VRs and G$_s$-VNrs reviewed in Bargmann, 1997) and the GABAB$_A$ receptor (Figure 2A). All these receptors possess a very large extracellular domain which shares a low but significant sequence similarity with periplasmic binding proteins of bacteria. Inside the bacteria, these proteins are involved in the transport of various types of molecules such as amino acids, ions, sugars or peptides. They are constituted of two lobes separated by a hinge region, and several studies including X-ray crystallography indicated that these two lobes closed like a Venus’ flytrap upon binding of the ligand. Based on detailed amino-acid sequence comparison, homology modeling and mutagenesis, the extracellular domain of mGluRs has been proposed to have a similar structure to the PBPs (O’Hara et al., 1993). In agreement with this proposal, the
Fig. 2. Classification and diversity of GPCRs. (A) Three main families (1, 2 and 3) can be easily recognized when comparing their amino-acid sequences. Receptors from different families share no sequence similarity, suggesting that we are in the presence of a remarkable example of molecular convergence. Family 1 contains most GPCRs including receptors for odorants. Group 1a contains GPCRs for small ligands including rhodopsin and β-adrenergic receptors. The binding site is localized within the seven TMs. Group 1b contains receptors for peptides whose binding site includes the N-terminal, the extracellular loops and the superior parts of TMs. Group 1c contains GPCRs for glycoprotein hormones. It is characterized by a large extracellular domain and a binding site which is mostly extracellular but at least with contact with extracellular loops e1 and e3. Family 2 GPCRs have a similar morphology to group 1c GPCRs, but they do not share any sequence homology. Their ligands include high molecular weight hormones such as glucagon, secretine, VIP-PACAP and the Black widow spider toxin, α-latrotoxin (Krasnoperov et al., 1997; Davletov et al., 1998). Family 3 contains mGluRs and the Ca\(^{2+}\) sensing receptors (Pin and Bockaert, 1995). Last year, however, GABA-B receptors (Kaupmann et al., 1997) and a group of putative pheromone receptors coupled to the G protein Go (termed VRs and Go-VN) became new members of this family (reviewed in Bargmann, 1997). (B) Family 4 comprises pheromone receptors (VNs) associated with Gi (Dulac and Axel, 1995). Family 5 includes the ‘frizzled’ and the ‘smoothened’ (Smo) receptors involved in embryonic development and in particular in cell polarity and segmentation. Finally, the cAMP receptors (cAR) have only been found in D.discoideum but its possible expression in vertebrate has not yet been reported. The dendrogram has been established with Clustal W.

The three main GPCR families are represented in the C.elegans genome. Family 1 contains ~150 members: 18 amine receptors, 50 peptide receptors and 80 orphan receptors. Family 2 has at least four members, whereas family 3 is represented by four metabotropic glutamate receptors and three GABA\(_B\) receptors. Among the 1000

N-terminal domain of mGluR1 produced by insect cells is a soluble protein and is still able to bind glutamate (Okamoto et al., 1998). Accordingly, it has been proposed that the large extracellular domain in a closed form may act as the activating ligand of their core domain. More work is necessary to demonstrate this hypothesis.
remaining orphan GPCRs, they probably encode 500 chemoreceptors, 200 genes with other functions and pseudogenes (Bargmann, 1998).

Homo- or heterodimerization: a revolution in current concepts of GPCR structures and functions

The classical view of GPCR/G-protein coupling stoichiometry is one receptor for one G-protein. However, the functional analysis of chimeric and mutated receptors revealed that they can dimerize (Maggio et al., 1993; Monnot et al., 1996) possibly via a coiled-coil interaction of their sixth TM. Indeed, a peptide corresponding to the sixth TM of the β2-adrenergic receptor inhibits both receptor dimerization and activation (Herbert et al., 1996), suggesting that GPCR dimerization may be important for G-protein activation. However, the functional significance of this phenomenon is still a matter of debate. The importance of 7TM receptor dimerization has recently been documented in the family 3 GPCRs. These receptors, including mGluRs and the Ca-sensing receptor, are homodimers disulfide linked at the level of their large extracellular domain (Romano et al., 1996; Bai et al., 1998; Okamoto et al., 1998). Very recently, the GABA-B receptor, which is related to family 3 GPCRs, has been shown to be a heterodimer. This receptor is constituted of two ‘subunits’ sharing sequence similarity (GABA-BR1 and GABA-BR2). None of these subunits gave rise to a functional receptor when expressed alone, but co-expression of both subunits gave rise to a GABA-B receptor efficiently coupled to G proteins (Jones et al., 1998; Kauppman et al., 1998; White et al., 1998). This observation further indicates that 7TM receptors function as dimers (or multimers), an idea first proposed by Rodbell (1992).

A new concept in cell biology and pharmacology arose only this spring. In order to be correctly folded, exported to the membrane, and in the case of CRLR (calcitonin receptor-like receptor) to obtain its final identity, some GPCRs need to form heterodimers with one TM domain proteins. It has been known for some years that nina A from Drosophila melanogaster and its vertebrate homologue, RanBP2, two cyclophilin-related proteins, bind opsins for folding and transport (Baker et al., 1994; Ferreira et al., 1996). In C.elegans, odr-4 and odr-8 are required to localize a subset of odorant GPCRs to cilia of olfactory neurons (Dwyer et al., 1998). More disturbing for pharmacologists is the report that CRLR is a virtual receptor which will generate the CGRP (calcitonin gene-related peptide) receptor when associated with RAMP1 (receptor-activity-modifying protein) and the adrenomedullin receptor when associated with RAMP2 (McLatchie et al., 1998). RAMP1 and -2 are also required for a correct glycosylation and transport of CRLR to the membrane. RAMPs, ord-4, and nina A have no sequence similarities.

All these data reveal a new level of complexity in the functioning of 7TM proteins, and may be of great help to elucidate the function of many GPCRs for which either the transduction pathway (5-HT5, angiotensine-II type 2 and dopamine D5 receptors for example) or the ligand (the so-called orphan receptors) is unknown.

Post-transcriptional and post-translational modifications of GPCRs for a fine tuning to G proteins

Splicing with insertions localized mainly at the third intracellular loop or C-terminal domain have been selected during evolution to modify or regulate the specificity and intensity of GPCRs coupling to G proteins (Journot et al., 1994). Recently, more subtle regulations of GPCR coupling to G proteins have been discovered. Transcripts
encoding the 5-HT$_{2C}$ receptor, a PLC-coupled receptor, undergo RNA editing events in which the genomically encoded adenosine residues are converted to inosines by a double-stranded RNA adenosine deaminase(s). Seven major 5-HT$_{2C}$ receptor isoforms are predicted, encoded by 11 distinct RNA species and differing in their second intracellular loops (Burns et al., 1997). This post-transcriptional modification leads to a 10- to 15-fold reduction in efficacy of the coupling of 5-HT$_{2C}$ to the G protein.

The post-transcriptional palmitoylations of some GPCRs at highly conserved Cys residues within the C-terminal tail and phosphorylations of the third intracellular loop or the C-terminal tail by GRKs (GPCR regulatory kinases) have been known for some time (O’Dowd et al., 1989; Carman et al., 1998). The phosphorylation of GPCRs by GRKs induce their functional desensitization. More recent reports indicate that phosphorylation may modify the coupling specificity of GPCRs. The $\beta_2$-adrenergic receptor is efficiently coupled to $\alpha_2$S but poorly to $\alpha_2$I. Once phosphorylated by PKA, it efficiently couples to $\alpha_2$I but weakly to $\alpha_2$S (Daaka et al., 1997). A putative critical phosphorylated site is the serine residue 262. The idea that a given receptor can be coupled to two different pathways depending on the duration of activation (and therefore potential desensitization) has also been highlighted in the case of mGluRs (Pin, 1998).

**GPCRs are unfaithful to G proteins**

It has been known for a long time that GPCRs interact directly not only with G proteins, but also with proteins called arrestins. These arrestins bind specifically to GPCRs phosphorylated by GRKs, an interaction which participates in the homologous desensitization of the receptor by disturbing their coupling to G proteins. Arrestins also target the receptors for internalization by virtue of their ability to interact with clathrin (for a review see Carman et al., 1998).

There are some particularly new and provocative observations indicating that GPCRs can also interact with the growing family of PDZ domain-containing proteins. In vertebrates, the prototype of PDZ domain proteins has been the PSD-95 family which organizes the NMDA receptors and some K$^+$ channels in post-synaptic densities (for a review see Kornau et al., 1997). These PDZ domains generally bind 3–4 amino-acid stretches of C-terminal sequences of target proteins having several motifs. The –S/TXV motif has been found in some PDZ target proteins but large variations around this theme have now been described. So far, three PDZ proteins have been found to interact with GPCRs, but they are obviously the first in the long list that is emerging. The most popular GPCR, the $\beta_2$-adrenergic receptor is unfaithful to G proteins in establishing, via its C-terminal domain (DSL), a contact with the first PDZ domain of the Na$^+$/H$^+$-exchanger regulatory factor (Figure 3A). This interaction has been shown to play a role in the regulation of the Na$^+$/H$^+$ exchange (Hall et al., 1998). MUPP1 is a multi-PDZ-domain protein of unknown function which interacts with the C-terminal domain of 5-HT$_{2C}$ (and probably 5-HT$_{2A}$, 2B) receptors (Ullmer et al., 1998).

Homer/Vesl proteins constitute a family of proteins with an Enabled/VASP homology (EVH)-like domain which interacts with group 1 mGluRs (mGluR1a, mGluR5) (Figure 3B) (Brakeman et al., 1997). One member, Homer 1a, is upregulated in electroconvulsive seizures and in long-term potentiation, a form of synaptic plasticity which may be involved in memory and seizures. Homer 1a is thus an immediate early gene (IEG). The other members, Homer 1b, 1c, Homer 2 and 3 and encode a C-terminal coiled-coil (CC) domain that confers the property of self- and hetero-multimerization (Xiao et al., 1998). A novel ‘Homer ligand’ (PPXXFR) present in the C-termini of mGluR1a and mGluR5 tightly interacts with the EVH domain of all Homer isoforms localized at their N-terminal domain (Tu et al., 1998). The Homer ligand is also present in the IP3 receptor and ryanodine receptor. Thus, it is likely that complexes containing mGluR1a, mGluR5, Homer 1b,c, Homer 2, Homer 3 and IP3 receptors (and possibly ryanodine receptors) can be formed (Figure 3B). Homer 1a, which does not multimerize because of the absence of the CC domain within its sequence, has been shown to disrupt the formation of multivalent complexes between other Homer proteins and mGluRs (Xiao et al., 1998).

This effect results in the inhibition of mGluR-induced Ca$^{2+}$ release from IP3 receptors in Purkinje cells (Xiao et al., 1998). Homer proteins have also been shown to be localized in the post-synaptic density of synaptic structures (Tu et al., 1998). This provides evidence that Homer 1a, which is rapidly expressed during high synaptic activity, may feed back to reduce synaptic function. A functional association of mGluR1 with ryanodine receptors has also been found in cerebellar granule cells in which activation of these receptors induced activation of L-type Ca$^{2+}$ channels (Chavis et al., 1996). The nature of such an interaction has not yet been identified, and an implication of Homer remains to be demonstrated.

**Conclusion and questions**

The GPCR saga started with one very simple question addressed over 25 years ago: how do hormones, such as glucagon and noradrenaline, activate adenyl cyclase? Although we now know the basic principles of cell–cell communication, there is no doubt that very important chapters and many questions remain to be addressed, including: how are GPCRs targeted within the cell? What is the physiological significance of their homo- and heterodimerization? What is the biological importance of their interactions with proteins other than G proteins? And finally, the crucial question, how will we succeed in resolving the GPCR structure?

**Acknowledgements**

We would like to thank Mireille Passama for her excellent iconographic techniques and Angela Madeuf for secretarial assistance. We would also like to thank Laurent Journot and Philippe Vernier for fruitful discussions.

**References**


Molecular tinkering of G protein-coupled receptors


Received November 11, 1998; revised January 12, 1999; accepted January 19, 1999