Activation of G-protein-coupled receptors: a common molecular mechanism

Sadashiva S. Karnik, Camelia Gogonea, Supriya Patil, Yasser Saad and Takanobu Takezako

Department of Molecular Cardiology, Lerner Research Institute, The Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA

G-protein-coupled receptors (GPCRs) are a large family of proteins that contain a seven transmembrane helical structural motif. They mediate responses to several ligands by binding and activating intracellular heterotrimeric G proteins. Since the cloning of the first GPCR, insights gained from structure–function studies, genetics and drug development have contributed to uncovering a common mechanism that explains the activation of diverse GPCRs by their cognate agonists. This mechanism takes into consideration the conservation of the structure–function relationship in the basic seven transmembrane structural motif, and the dynamic changes in receptor conformation that are associated with activation. Combining models derived from the X-ray structure of rhodopsin with structure–function data allows a deeper understanding of the activation mechanism of GPCRs.

The G-protein-coupled receptor (GPCR) superfamily is one of the largest families of proteins in mammals. GPCR signaling is the primary mechanism by which cells sense changes in the external environment and convey this information to their interior. Abnormalities of signaling by GPCRs are at the root of disorders that affect most tissues and organs in our body, such as hyperfunctioning thyroid adenoma, precocious puberty, nephrogenic diabetes insipidus and color blindness. Targeting GPCRs for therapeutic intervention has been fruitful, with >50% of drugs on the market acting as either surrogate activators or inhibitors of the GPCRs that have defined native ligands. However, the majority of GPCRs identified (>75%) are orphan receptors, which presents a challenge for identifying their native ligands and defining their function. The common structural criteria for inclusion in the GPCR superfamily is the presence of seven stretches of 25–35 predominantly hydrophobic residues that are believed to form a seven transmembrane (7TM) α-helical bundle with helices linked by three intracellular and three extracellular (EC) polypeptide loops [1]. The enormous diversity observed in the secondary structure of GPCRs and agonist structures has prompted the question: does a common structure–function principle link members in this superfamily?

Here, we try to answer this question by examining evidence from several GPCRs that substantiates a crude but consistent picture of the movement of specific TM helices on activation. Conserved structural features in this superfamily indicate that there is a structural basis for a common mechanism of activation.

The primary structures of most GPCRs are derived from their gene sequences. The secondary structures of GPCRs (Fig. 1), modeled using hydrophobicity algorithms, yield a structure that comprises seven antiparallel TM helices [1]. The 7TM structural motif was first identified in the archaebacterial proton pump, bacteriorhodopsin (bR), which is the best understood 7TM protein [2]. The model of the secondary structure of the 7TM domain of rhodopsin was proposed, based on protein and gene sequencing. High-resolution structures of bR and rhodopsin indicate considerable precision in the inferred 7TM structures [3,4] and validate the structure-prediction methods used currently.

The functional criterion for inclusion in the GPCR superfamily is the principle that the 7TM domain in each receptor exerts a biological effect by recruiting and activating heterotrimeric G proteins, but this is not established formally in all cases [5,6]. In response to stimulation, it is assumed that GPCRs bind heterotrimeric GTP-binding proteins and activate GTP–GDP exchange, which leads to the subsequent dissociation of the GTP-bound α-subunit and βγ-dimer from the GPCR. Both Gα and Gβγ subunits can modulate several signaling pathways, including activation of phospholipases and phosphodiesterases, and modulation of adenylate cyclases and ion channels. In addition to activating G proteins, several GPCRs recruit signaling molecules, including receptor serine/threonine kinases, tyrosine kinases, protein tyrosine phosphatases and adaptor proteins. GPCRs perform a variety of vital functions, including the response to light, odor, taste, neurotransmitters and hormones. Their ligand structures are diverse, including small organic molecules, lipids, ions, hormones, short and large polypeptides, and glycoproteins [6].
On the basis of conserved sequences in the 7TM bundle homology, their primary structure contains the 7TM motif. Although these putative GPCRs have no overall sequence estimate of.

Analysis of nonredundant GPCR sequences yields an estimate of >800 GPCRs in the human genome [7–9]. Although these putative GPCRs have no overall sequence homology, their primary structure contains the 7TM motif. On the basis of conserved sequences in the 7TM bundle in several functionally related receptors, the A–F clan has been used to classify GPCRs from vertebrates, invertebrates, fungi and archaea [8]. The human GPCRs can be classified into five distinct families (Box 1): rhodopsin (701 members); glutamate receptor (15 members); adhesion receptor (24 members); frizzled/taste2 receptor

**The human GPCR superfamily**

Analysis of nonredundant GPCR sequences yields an estimate of >800 GPCRs in the human genome [7–9]. Although these putative GPCRs have no overall sequence homology, their primary structure contains the 7TM motif. On the basis of conserved sequences in the 7TM bundle in several functionally related receptors, the A–F clan has been used to classify GPCRs from vertebrates, invertebrates, fungi and archaea [8]. The human GPCRs can be classified into five distinct families (Box 1): rhodopsin (701 members); glutamate receptor (15 members); adhesion receptor (24 members); frizzled/taste2 receptor

**Box 1. G-protein-coupled receptors (GPCRs) encoded by the human genome**

Over 800 GPCR-encoding sequences have been identified in the human genome [9]. Five main families have been identified that have high boot-strap values. These receptors – glutamate, rhodopsin, adhesion, frizzled/taste2 and secretin – form the basis of the GRAFS classification system. This system is better suited to GPCRs encoded by a single mammalian genome than the previously used A–E Classification System [8], which encompassed seven transmembrane (7TM) receptors from several species.

The phylogenetic relationship of the 802 GPCRs in humans is based on the characteristics of the TM domains. The five families represent the smallest number of clusters with high boot-strap values. The rhodopsin family is further subdivided into four main groups with 13 branches.

**Family G**

Family G has 15 members that have an N-terminal domain of 280–580 residues. Examples include eight metabotropic glutamate receptors, two GABA receptors, one Ca2+ receptor and five group-1 taste receptors.

**Family R**

There are 701 members in family R, including 241 nonolfactory receptors, and there are four main groups. The N-terminal length is <100 residues with an important exception in receptors for glyco-protein hormones, LH, FSH, TSH and CG. Current estimates indicate that there are 480 unique, functional olfactory receptors. These proteins form a stable phylogenetic cluster.

The **ε-group** consists of five main branches. The prostanoid-receptor cluster has 15 members, the amine-receptor cluster has 39 members, the opioid-receptor cluster has nine members, the melanin-receptor cluster has three members, and the melanocortin-endoelin-cannabinoid-adenosin-receptor cluster has 22 members.

The **γ-group** has three main branches. The somatostatin-opioid-galanin cluster has 15 members, the melanin-concentrating hormone cluster has two members and the chemokine receptor cluster has 42 members.

The **β-group** consists of four branches. The MAS related cluster with eight members, the glycprotein-receptor cluster with eight members, the purine-receptor cluster with 42 members, and the olfactory receptors with an estimated 460 members.

**Family A**

This family consists of 24 members with either three or four branches. The N terminus contains 200–2800 residues, an EGF-like motif and a mucin-like motif. It is rich in Pro, Cys and highly glycosylated. Examples include brain-specific angiogenesis inhibitory receptors, lectomedin receptors, CD97 and GPR56.

**Family F**

Family F consists of 24 members with two distinct clusters of 10 frizzled receptors and 13 group-2 taste receptors. Their grouping is based on short conserved sequences in transmembrane helix 2 (TM2), TM5 and TM7. The N terminus is ~200 residues in frizzled receptors and it is very short in group-2 taste receptors.

**Family S**

Family S has 15 members with four subgroups that each bind large peptides with shared identity. The N-terminal length is 60–80 residues with conserved cysteine residues. Examples include vasooactive intestinal peptide receptor, calcitonin receptor, gastric inhibitory peptide receptor, corticotropin releasing hormone receptor, glucagon receptors, growth hormone releasing hormone receptor and parathyroid hormone receptor.
(24 members); and secretin receptor (15 members). At least 24 sequences that do not belong to these families are reported [9].

The 7TM topology shown in Fig. 1 is emerging as the typical structure in the GPCR superfamily. Several, highly conserved features, such as an S–S bond that links TM domain 3 (TM3) and extracellular loop-2 (ECL2), the presence of a DRY-motif in TM3, and an NPxxY-motif in TM7, that might have an essential role in the structure or function, unify the members of the rhodopsin family. Although a few members of the rhodopsin family have N termini of >100 amino acids, receptors in all other GPCR families contain long N-terminal sequences (500–600 amino acids) that are rich in cysteine residues. These receptors have no homology to the rhodopsin family, apart from the disulfide bridge between TM3 and ECL2. Although the diverse structures observed in the clan-classification system are well represented, clans D (phagomone receptors), E (cAMP receptors), F (archae-bacterial opsins) and family IV (invertebrate opsins) in clan A are not represented in the human GPCRs.

The human GPCR families demonstrate the diversity in the basic 7TM structure in a single mammalian genome. Specificity for the enormous diversity of ligands appears to be accomplished by adding modules to the N terminus of the basic 7TM structure. The conservation of the 7TM motif might indicate that the mechanism of activation and G-protein-coupling in GPCR signal transduction is preserved. This raises two questions: (1) how do diverse ligands interact and induce physical changes in the 7TM structure; and (2) are these changes similar or distinct in the different receptors?

Ligand-induced activation of GPCRs

The molecular mechanism of ligand activation is shown best for rhodopsin and related visual pigments. These contain the covalently bound light-sensing chromophore, 11-cis-retinal, which is an inverse agonist [3,5]. In general, binding of an inverse agonist causes the transition of a GPCR from the native, partially active state to an inactive state [5,6]. Indeed, the opsins form of rhodopsin is partially active and the rhodopsin form is inactive [5]. Absorption of a photon causes 11-cis-retinal to isomerize to the agonist all-trans-retinal, which induces conformational changes in rhodopsin that are similar to those identified in bR by high-resolution crystallography (Box 2) [3–5].

Disruption of a salt-bridge interaction between TM3 and TM7 is the primary trigger for conformational changes in opsins. However, binding of agonists activates all other GPCRs. In the amine receptors, agonist binding disrupts the salt bridge between TM3 and TM7, similar to opsins [10]. In the peptide-hormone receptors salt-bridge disruption is not common, but displacement of residues in TM3 leads to activation [11,12]. Thus, the mechanism of GPCR activation involves the relaxation of constraining intramolecular interactions and the formation of new interactions.

Specific movements of the TM helices are essential for activation of function [13–21]. Disulfide cross-linking between cysteine pairs introduced artificially into TM3 and TM6 prevents G-protein activation by the cytoplasmic loops. The magnitude of movement of TM6 is greater than that of TM3 and TM7. As a result, the inner faces of TM2, TM3, TM6 and TM7 become more exposed and the cytoplasmic ends of TM4 and TM5 become less exposed [5]. These changes mapped onto the high-resolution structure of the inactive state of bovine rhodopsin indicate that activation by light opens the 7TM bundle at the cytoplasmic end, which enables rhodopsin to bind and activate the G protein.

Conservation of helical movement in agonist-activated GPCRs

The X-ray-crystal structure of the inactive state of bovine rhodopsin provides a deeper understanding of functional

---

Box 2. Bacteriorhodopsin (bR), the prototype seven transmembrane (7TM) receptor

bR is the only 7TM receptor in which the ordered conformational changes responsible for transmembrane signal transduction is described at high resolution [4]. bR is a light-driven proton pump from the purple membrane of Halobacterium halobium. It contains 7TM helices linked by short extracellular loops (ECLs). Each bR contains one molecule of the chromophore, all-trans-retinal, which is covalently attached to a lysine residue in TM7 to form a protonated Schiff base. Photon absorption induces isomerization of the chromophore to 13-cis-retinal. This triggers changes in the protein conformation, which passes through several distinct spectral intermediates before returning to the ground state. These changes in protein conformation are necessary to transport a hydrogen ion out of the cell against some bulky residues in TM6. This initiates an outward movement (3–4 Å) of the cytoplasmic end of TM6 followed by the cytoplasmic end of TM7. These movements increase the accessibility of the Schiff base to the protonated Asp96 in the cytoplasmic part of the TM domain. In the final stages of the proton-pumping cycle, movement of residues involved in a hydrogen-bonded network that faces the outside of the cell facilitates the release of the proton from Asp85 to the outside of the cell. Finally, the retinal relaxes to the all-trans form, TM6, TM7 and TM3 swing back to their original position, and the next proton-pumping cycle begins.

The primary changes in protein conformation that are coupled to vectorial translocation of the proton through the protein are smaller than might be anticipated. The largest motion is observed in the cytoplasmic segments of TM6, TM7 and the loop that links TM5 and TM6, which induces an open configuration in the cytoplasmic part of the TM bundle. The agonist acts as a valve that imparts a unique direction to the coordinated movement of helices.

Although the mechanism of bR activation forms a prototype for considerably more complex 7TM proteins, differences in the detailed mechanisms are expected. For GPCRs, the seven TM helices that surround the bound ligand presumably transmit subtle conformational changes to the cytoplasmic regions to recruit a G protein. A working model of the primary conformational changes in signal transduction by GPCRs (Fig. 2) is similar to that of bR.
Diversity in coupling between EC and TM domains

Interaction between agonists and the TM domain is not always necessary to activate the conserved helical movements. Short peptide ligands for the rhodopsin family GPCRs, angiotensin II, N-formyl-Met-Leu-Phe, thyrotrpin-releasing hormone and gonadotropin-releasing hormone appear to enter the TM core as well as make contact with the ECLs [12,24–25]. The contacts in the EC domain are crucial for activation of some receptors [12,24–27]. In the secretin family of receptors for glucagon, parathyroid hormone and vasoactive intestinal peptide, agonists bind primarily to the long N-terminal tail (116–147 amino acids), but require the EC loops to activate the receptor [27–31]. With protease-activated receptors, the tethered ligand is activated by proteolysis and binds intramolecularly to the ECL2 to cause the physiological signal: platelet aggregation [32]. The glycoprotein hormone receptors contain long, N-terminal regions of ~350–400 residues that contain leucine-rich repeats. Although the N-terminal segment alone can bind the hormone, the response depends on the EC loops [33–36]. The activation signal is believed to be generated by the EC loops because mutations that cause endocrine disorders of the glycoprotein hormones, leutinizing hormone, follicle-stimulating hormone, choriongonotropin and thyrotropin result in specific structural changes in the TM domain that lead to G-protein activation and that direct ligand–TM domain interaction is not always necessary. How does TM-helical movement occur in such GPCRs?

Coupling between domains in GPCRs

Several lines of evidence indicate that coupling occurs between the EC, TM and cytoplasmic domains in GPCRs. G-protein binding to the cytoplasmic domain induces a high-affinity state in many GPCRs, which indicates that this causes an optimal reconfiguration of agonist binding groups on the TM or EC domain [44]. Mutations that increase the affinity of receptors for G proteins also induce the high affinity agonist-binding state. In most GPCRs, constitutively activating mutations can be located in any part of the receptor [6,23]. The network of intramolecular interactions that constrain ligand-free GPCRs are distributed throughout the receptor and work cooperatively. Although folding of the receptor polypeptide generates this intrinsic thermodynamic constraint in the wild-type receptor, it fails to do so in constitutively active mutants, which results in higher conformational entropy [6,23,45,46].

Detailed studies of the retinitis pigmentosa (RP) mutations in rhodopsin demonstrate that tertiary folding of the receptor polypeptide generates co-operativity between domains [45–49]. RP mutations localize to the EC, TM and cytoplasmic domains of rhodopsin and cause opsin to be either partially or completely misfolded. The misfolded fraction of the polypeptide lacks the conserved disulfide bond needed in the properly folded, functionally active fraction [47,48]. This disulfide bond is central to producing the native structure [45] and is also required for rhodopsin function [45–49]. Therefore, coupling between domains results from receptor folding.

These observations demonstrate that structural perturbation caused by gain- and loss-of-function mutations in one domain are transmitted to other domains, thus the tertiary structures of these domains are coupled to one another.

Conserved structural features in the EC domain

What is the structural basis of coupling between the GPCR domains? The interhelical contacts in the TM domain are recognized as a major factor. The interaction between cytoplasmic loops is known from structure–function studies [50,51]. It is also possible that ECL1, ECL2 and ECL3 associate with the N-terminal tail to form a compact structure, as seen in the crystal structure of rhodopsin [3]. Generation of this structure might be more important to GPCR function than recognized currently. As shown in Table 1, 52% of inactivating mutations in the interhelical loops of GPCRs are located in ECL2 and ECL3. Assuming that these ECLs do not bind agonists directly, their
mutation indicates a role in receptor structure. The crystallographic structure of rhodopsin reveals extensive interactions between ECL2, the inverse agonist ligand and the TM domain. Drugs that target GPCRs are, in many instances, affected selectively by mutations in the ECLs [51]. In the metabotrophic glutamate receptors, the frizzled family and the secretin receptor family, either deletion or mutation of the large, N-terminal, ligand-binding domain results in constitutive activity. This indicates that the activity of TM domain is constrained by the EC domain in the native receptors in absence of the ligand [25–43].

Hydrophobic contacts, salt bridges and hydrogen bonds between the N-terminal segment and ECL1, ECL2 and ECL3 are thought account for their cooperative activity. The structure of the N-terminal EC domain varies most, with an average length of 62 ± 98 residues (mean ± SD) for the entire GPCR superfamily. The variation in length in all other EC loops is less (Table 1). The similar length of the ECLs of the GPCR superfamily implies that they might mediate a conserved function.

Evolutionary conservation of the S–S bond between TM and EC domains
A disulfide bond linking TM3 and ECL2 occurs in >91% members of the GPCR superfamily (Fig. 1). In most GPCRs, mutations that disrupt this link directly affect all aspects of receptor function (Table 1). The first evidence for the involvement of a disulfide bond in GPCR activation comes from experiments by Pederson and Ross [52] and it’s role is understood best in rhodopsin and members of the rhodopsin family [44–49,53–59]. In all GPCRs, this disulfide linkage might be responsible for securing the interactions between TM helices and ECL2, similar to those observed in rhodopsin (Fig. 2). ECL2 might be the core of EC domain structure. Therefore, conservation of the disulfide bond indicates its importance to folding during receptor biogenesis and to structural coupling between TM and EC domains. Several putative, orphan GPCR families appear to lack this disulfide bond and it is possible that other types of interaction act as a substitute in these receptors.

Structural determinants of a common molecular mechanism of GPCR activation
In the GPCRs, the change from an inactive to an active conformation is necessary for G-protein activation [60].
Conformational change in the cytoplasmic domain ensures G-protein binding and activation. This is accomplished by a conserved pattern of movements of the TM helices. These are initiated when an agonist binds to either the TM domain or the EC domain (formed by interactions between the N-terminal region and three ECLs). Coupling between the EC and the TM domains is deemed crucial, and the conservation of the disulfide bond is important for this coupling (Fig. 2). Ligand activation in the GPCRs that have large ligand-binding modules at the N terminus is likely to involve extensive interaction with the ECLs. Therefore, direct contact with the TM domain is not necessary for agonists that bind to the EC-domain to initiate movement of the TM helices and signal transduction.

Concluding remarks

In summary, the 7TM motif in GPCRs reveals a conserved activation mechanism in which movements of the TM helices that occur on ligand binding cause conformational changes in the cytoplasmic domain (Fig. 2). The pattern of movements of the helices that have been established in bR and observed in other members of the rhodopsin family of receptors might hold true for the agonist-induced activation of diverse GPCRs. The similar length of the polypeptide segments and the conserved disulfide linkage in the EC domain in GPCRs indicates that there is significant evolutionary pressure for conservation and also that the EC domain has an important role in the function of this diverse superfamily of receptors. In this context, some families of GPCR receptors are active as dimers. How a mechanism based on activation by a single receptor molecule applies to these GPCRs needs additional work. Investigation of different GPCR families is needed to confirm the importance of coupling between different domains in GPCRs. This might hold promise for developing novel strategies for drug development and therapeutic intervention in diseases caused by mutations in GPCRs.

Acknowledgements

We thank Robin Lewis for assistance preparing the article. This work was supported in part by ROI Grants HL57470 and HL64845 to S.K. and a Fellowship award to Y.S. from the American Heart Association. S.K. is an Established Investigator award recipient from the American Heart Association.

References

10 Porter, J. et al. (1996) Activation of the α1B-adrenergic receptor is initiated by disruption of an interhelical salt bridge constraint. J. Biol. Chem. 271, 28318–28323
17 Boucard, A.A. et al. Constitutive activation of the AT1 receptor alters the spatial proximity of transmembrane 7 to the ligand binding pocket. J. Biol. Chem. (in press)
18 Gether, U. et al. (1997) Agonists induce conformational changes in transmembrane domains III and VI of β2-adrenergic receptor. EMBO J. 16, 6737–6747
21 Chen, S. et al. (1999) Phe310 in transmembrane VI of the α1B-adrenergic receptor is a key switch residue involved in activation and catecholamine ring aromatic bonding. J. Biol. Chem. 274, 16320–16330
27 Davison, J.S. et al. (1997) Irreversible activation of the gonadotropin-releasing hormone receptor by photoaffinity cross-linking: localization of attachment site to Cys residue in N-terminal segment. Biochemistry 36, 12881–12889
Alteration of an extracellular agonist recognition domain causes constitutive signaling. J. Biol. Chem. 271, 702–706
33 Parma, J. et al. (1995) Somatic mutations causing constitutive activity of the thyrotropin receptor are the major cause of hyperfunctioning thyroid adenomas: identification of additional mutations activating both the cyclic adenosine 3',5'-monophosphate and inositol phosphate-Ca++ cascades. Mol. Endocrinol. 9, 725–733
34 Li, R. et al. (2001) Mutations of the second extracellular loop of the human lutropin receptor emphasize the importance of receptor activation and de-emphasize the importance of receptor phosphorylation in agonist-induced internalization. J. Biol. Chem. 276, 7968–7973
35 Chen, C.R. et al. (2003) Targeted restoration of cleavage in a noncleaving thyrotropin receptor demonstrates that cleavage is insufficient to enhance ligand-independent activity. Endocrinology 144, 1324–1330
36 Vlaeminck-Guillemin, V. et al. (2002) Activation of the cAMP pathway by the TSH receptor involves switching of the ectodomain from a tethered inverse agonist to an agonist. Mol. Endocrinol. 16, 736–746
37 Kubo, Y. et al. (1998) Structural basis for a Ca++-sensing function of the metabotropic glutamate receptors. Science 279, 1722–1725
39 Litschig, S. et al. (1999) CPCCOEt, a noncompetitive metabotropic glutamate receptor 1 antagonist, inhibits receptor signaling without affecting glutamate binding. Mol. Pharmacol. 55, 453–461
51 Dixon, R.A. et al. (1987) Structural features required for ligand binding to the beta-adrenergic receptor. EMBO J. 6, 3269–3275
53 Noda, K.N. et al. (1994) The high affinity state of the β2-adrenergic receptor requires unique interaction between conserved and non-conserved extracellular loop cysteines. J. Biol. Chem. 269, 6743–6752
54 Zeng, F.Y. et al. (1999) Conserved extracellular cysteine pair in the M3 muscarinic acetylcholine receptor is necessary for proper receptor cell surface localization, but not for G protein coupling. J. Neurochem. 72, 2404–2414
56 Zhang, R. et al. (1996) Requirement of cysteine residues in exons 1-6 of the extracellular domain of the luteinizing hormone receptor for gonadotropin binding. J. Biol. Chem. 271, 5755–5760

Endeavour

the quarterly magazine for the history and philosophy of science

Sex glands, vasectomy and the quest for rejuvenation by C. Sengoopta
Global science: the eruption of Krakatau by M. Potts
Two pills, two paths: a tale of gender bias by M. Potts

Locate Endeavour in the BioMedNet Reviews collection. (http://reviews.bmn.com) or on ScienceDirect (http://www.sciencedirect.com)