ABSTRACT—Since the first discovery of mammalian receptors for adrenaline ($\beta_2$) and acetylcholine (M$_1$) in 1986, many G protein-coupled receptors for known ligands have been cloned by protein purification, PCR (polymerase chain reaction) and low stringency hybridization, and they have been identified by expression cloning techniques. Now we are almost out of the known ligands pool. However, through the achievement of the Human Genome Project, numerous orphan receptors (whose natural ligands are not yet found) are also available for analysis. In this review, I would like to review recent achievements in the discovery of natural ligands, to describe useful orphan receptor strategies, and to predict the future of reverse pharmacology.

Keywords: Orphan G protein-coupled receptor, G protein, Ligand, Reverse pharmacology

I. Introduction

Dixon et al. reported the first cloned mammalian G protein-coupled receptor (GPCR) sequence (1). The DNA sequence from a hamster encodes the $\beta_2$-adrenergic receptor. In the same year, 1986, Kubo et al. published the muscarinic acetylcholine receptor sequence (M$_1$) from a rat (2). Both groups purified the receptor proteins and identified the DNA sequences (1, 2). These novel findings opened a new era of pharmacology, a period I would call the ‘receptor-hunting period’. Many GPCR for known ligands like histamine, serotonin, prostaglandin, and even morphine have been cloned by PCR (polymerase chain reaction), low stringency hybridization, and expression cloning techniques (3). Now, the Human Genome Project is completed and all the candidate genes are available for study (4). Almost 350 GPCRs are considered as functional receptors in human beings excluding olfactory receptors (5). Among the genes, about 200 GPCRs are matched with their own ligands. However, the rest of the GPCRs are classified as ‘orphan receptors’, because their natural ligands are not yet known (5). Considering that half or more than half of the drugs on the current market are acting on the GPCRs positively (agonists) or negatively (antagonists), the orphan receptor project is a fascinating ‘goldmine’ to the pharmaceutical companies and academic scientists. In this review, I would like to introduce the techniques useful for orphan GPCR research with several examples and prospect the future of orphan receptor research.

II. GPCR

II-1. Characteristics of GPCR

Each GPCR binds specifically to its own ligand and signals via G proteins to the inside of cells. Receptors have seven transmembrane domains. The amino terminal faces the extracellular space and the carboxy terminal is located in the cytosol (Fig. 1). Transmembrane (TM) domains are more conserved among GPCRs than extracellular or intracellular domains. There are several signature amino acid motifs which provide us their identity as GPCRs; for example, the LxxxD motif in the TM II, the DRY motif at the end of the TM III and the NPxxY motif on the TM VII. Usually, the intracellular domain III (between TM V and TM VI) and the carboxy terminal are considered to play roles in G protein coupling (6). Each receptor has its own selectivity to G proteins, like $\beta$-adrenceptors to G$_s$ and $\alpha_2$-adrenoceptors to G$_i$. However, the precise sequences activating each G protein (G$_s$, G$_i$, G$_q$, G$_{12}$, etc.) are as yet unknown, although there is a proposed theory that basic amino acids are important for G protein coupling like BBxxB sequences (B means basic amino acid) (7). Therefore, orphan receptor sequences give us no clue to help predict which G protein is involved in the receptor coupling. It is possible, however, to presume a candidate ligand from the amino acid identity of an orphan receptor to
II-2. Discovery by sequence homology

High homology of amino acid sequences with known receptors sometimes gives an idea, if the identity is high enough (like $\geq 45\%$). Sphingosine 1-phosphate receptor (S1P$_1$, formerly known as Edg-1) has a high identity to other members of the Edg orphan receptor subfamily (8). Edg-3, Edg-5, Edg-8 and Edg-6 have been characterized as S1P receptors serially and been renamed as S1P$_2$, S1P$_3$, S1P$_4$ and S1P$_5$, respectively (Fig. 2) (8, 9). Other members of the subgroup have turned out to be receptors for lyso-phosphatidic acid (LPA), the chemical structurally-related to S1P (Fig. 2) (8, 9). They have a high similarity to each other in amino acid sequences (about 45 to 55%). S1P$_1$, the first cloned orphan of the subfamily, was transfected into HEK293 cells (10). In the presence of serum, S1P$_1$-expressing HEK293 cells showed morphologic change. This phenomenon became a clue, which led to the identification of the lipid ligand from the serum (10). S1P$_2$, S1P$_3$ and S1P$_5$ were easily expected to be S1P receptors by the high identity of amino acid sequences (11, 12). S1P$_4$ has the lowest

![Fig. 1.](image1) Structural characteristics of G protein-coupled receptors. G protein-coupled receptors have heptahelical domains. Amino termini are present in the extracellular space and carboxy termini are present in the intracellular space. Highly conserved common motifs in each TM follow: TM I, GN or GxxxN; TM II, LxxxDxxxxxxx(x)P; TM III, SxxxxLxxIxxD/ERY; TM IV, Wxxxxxxx(x)P; TM V, FxxPxxxxxxxY; TM VI, FxxCxxP; and TM VII, LxxxxxxxN/DpxxY.

![Fig. 2.](image2) Dendrogram of receptors for lysolipid mediators. The dendrogram was constructed by the GCG program pileup. Highly related receptors are branched out at closer point than less related receptors. Longer distance between two receptors means lower identity. LPA receptors, S1P receptors and receptors for SPC, LPC and psychosine form three subgroups. Within a subgroup, receptors recognize the same ligand or highly-related chemicals. Receptor names for S1P and LPA followed the IUPHAR nomenclature system. hRHOD stands for human rhodopsin receptor. The underlying amino acid sequences are found associated with the following GenBank flatfiles: hS1P1, M31210; hS1P2, AF034780; hS1P3, X83864; hS1P4, AJ000479; hS1P5, AC073749; hLPA1, U78192; hLPA2, AC002306; hLPA3, AF186380; hOGR1, U48405; hGPR4, L36148; hTDAG8, U95218; hG2A, AF083955.
identity to other members. Several groups have failed to get a positive coupling of S1P\(_4\) to S1P in several well-known assays. However, a couple of reports suggested that S1P\(_4\) is another member of the S1P receptors having a low affinity (13, 14). The difference of ligand binding between S1P\(_4\) and other S1P receptors was also found in the different recognition by an immunomodulator, FTY-720, which has a sphingosine-like structure (15, 16). Another example of deorphanization discovery by sequence homology is OGR-1, a sphingosylphosphorylcholine (SPC) receptor (Fig. 2) (17). OGR-1 (ovarian G protein-coupled receptor-1) was cloned as an orphan and named by its specific expression in ovarian cancer tissue. Its identification as a receptor for SPC stimulated scientists to investigate three cousin GPCRs (GPR4, TDAG-8 and G2A). Those receptors have a high identity of amino acid sequences with each other (38 – 50%). The results are quite interesting. GPR4 was characterized as a high affinity receptor for SPC (18). Furthermore, it binds to lysophosphatidylcholine (LPC), which in glycerolipid is a counter molecule to SPC, with a low affinity (18). Another member, G2A, was characterized as a receptor for LPC and has a low affinity to SPC (18). The last member of the OGR-1 subfamily, TDAG-8, was identified as a receptor for another sphingolipid, psychosine (galactosyl sphingosine) (19). The involvement of psychosine in Krabbe’s disease and the role of psychosine receptor TDAG-8 in cytokinesis are quite interesting findings (19). However, in many cases, the identities of orphans to ligand-known receptors are too low (≤35%) to convince the ligands to recognize them.

II-3. Expression cloning techniques with known ligands
When there is a ligand and search receptor(s) for the ligand, a useful technique that has been applied is ‘expression cloning’ (3). Many receptors for the known ligands were found with this technique; for example, receptors for endothelin ET\(_\alpha\), opioid \(\delta\), platelet activating factor, angiotensin AT\(_1\), histamine H\(_1\) and others (3). Discovery of the CysLT\(_1\) leukotriene D\(_4\) (LTD\(_4\)) receptor is also a good example (20). G\(_{q}\)/11-type G proteins were supposed to be coupled to the LTD\(_4\) receptor and mediate an increase of intracellular Ca\(^{2+}\) concentration, based on the previous pharmacological studies on LTD\(_4\) action. Lynch et al. used frog oocytes (Xenopus laevis) for receptor expression, because oocytes are devoid of LTD\(_4\) responses such as responses to other GPCR ligands (oocytes have been a popular tool for GPCR research) (20). The expression of a hundred of orphan receptor mRNAs to each oocyte and challenging each oocyte with LTD\(_4\) led successfully to the identification of an orphan receptor (HG55, the former orphan name) as the CysLT\(_1\) LTD\(_4\) receptor (illustrated in Fig. 3a). A subsequent binding experiment had to be performed to fulfill the pharmacological criteria. Another interesting example is the BLT\(_1\) leukotriene B\(_4\) (LTB\(_4\)) receptor.

---

**Fig. 3.** Illustrations of expression cloning techniques. a: The expression cloning procedure, which was used for the CysLT\(_1\) leukotriene D\(_4\) (LTD\(_4\)) receptor discovery. Each mRNA of 100 orphan GPCRs was synthesized and injected into different oocytes. Each oocyte was tested with LTD\(_4\). The one giving a response to LTD\(_4\) was HG55, renamed as the CysLT\(_1\) LTD\(_4\) receptor. Other orphan receptor mRNAs did not evoke significant Cl\(^-\) conductance. b: The subtraction strategy procedure, which was used for the leukotriene B\(_4\) (LTB\(_4\)) receptor discovery. After hybridization of single-strand cDNAs of poly A\(^+\) mRNAs from retinoic acid-differentiated HL-60 cells with excess amounts of poly A\(^+\) mRNAs from undifferentiated HL-60 cells, the remaining single-strand cDNAs were isolated and cloned in a vector. Candidate genes were expressed in CHO cells and tested by LTB\(_4\) binding.
An idea occurred to Yokomizo et al. from the fact that the LTB₄-induced response is enhanced after the granulocyte-like differentiation of HL-60 leukemia by the treatment of retinoic acid (21). They subtracted cDNAs from differentiated HL-60 cells with ones from undifferentiated HL-60 cells. Among the remaining cDNAs, the BLT₁/LTB₄ receptor was found (Fig. 3b). The group established CHO cells stably-expressing BLT₁ receptor and characterized the signaling property of the receptor including the chemotactic effect. The subtraction strategy and molecular cloning work confronted an unexpected previous report. The same gene had already been published as a putative purinergic ATP receptor P2Y₁. The authors conducted a precise confirmation and denied the former report (21).

II-4. Reverse pharmacology with orphan receptors

As described in the introduction, we have about 150 orphan receptors but are out of the ligand pool for the orphan receptors. From the late 20th century, pharmacologists started to search for natural ligands with cloned orphan receptors. We call this strategy ‘reverse pharmacology’ because it starts from receptor DNAs, in contrast to the traditional pharmacology that started from drug molecules or disease studies and resulted in receptor molecule discovery at the end.

One of the orphan receptor strategies is to choose an orphan and investigate the ligand for the orphan receptor. Discovery of bioactive novel substances like norceptin/orphanin FQ, orexin/hypocretin, prolantin-releasing peptide and apelin are examples (22 – 26). An orphan receptor, APJ, was cloned by B.F. O’Dowd and his colleague in 1993 (27). APJ is similar to the angiotensin receptor. However, angiotensin was not the ligand for the APJ. A group selected APJ and established APJ-expressing CHO cells (26). They measured changes of extracellular pH around APJ-CHO cells after the addition of a variety of tissue extracts from cows. Extracts from the stomach and brain induced pH change. Then, the novel ligand, apelin, was purified from the extracts (Fig. 4a) (26). ORL1 (opioid receptor-like-1), an orphan receptor, is another example. Its high identity of amino acid sequence to opioid receptors (>65%) and lack of response to the opioid agonists led scientists to search for the ligand of the receptor. Its negative coupling to adenylyl cyclase was suggestive and also provided a research base for the screening and purification of the peptide ligand. Finally, the ligand named norceptin/orphanin FQ was identified by two separate groups (22, 23).

The other strategy is high-throughput screening on the orphan receptor library with a massive library of candidate chemicals (Fig. 4b). I would like to introduce a couple of massive experiments: First, a library of known and putative natural GPCR agonists including neuropeptides, bioactive lipids (leukotrienes, prostaglandins, etc.), steroids (aldosterone, testosterone, etc.), amines (catecholamines, etc.), cannabinoids, all commonly occurring l-amino acids and nucleotides (ATP, UTP, ADP, etc.) and chemically related substances (UDP-glucose, UDP-N-glucosamine, etc.) was collected (28). Out of over 700 chemicals, UDP-glucose...
was found to be a ligand for the orphan receptor KIAA0001 (28). To avoid endogenous responses to the testing compounds, the authors adapted yeast cells for the expression of orphan GPCRs. As a single-cell organism, yeast has only two GPCR genes. Using a pheromone receptor signaling system and its facilitation by overexpression of promiscuous G protein (αi1b) reduced background signals to zero. Another story is the matching of an orphan GPR38-A with motilin (29). A pharmaceutical company developed a high-throughput assay, which measured the increase of [Ca2+]i with a bioluminescent Ca2+-sensitive reporter protein, aequorin. In this assay, activation of a Gα protein-coupled receptor by a ligand induces IP3-mediated mobilization of intracellular calcium and concomitant calcium-induced aequorin bioluminescence. A broad set of peptide and non-peptide molecules for receptors, ion channels, and intracellular signaling enzymes (>500 substances) was tested in this assay system. Transient expression of the GPR38-A protein gave a fourfold bioluminescent response with the combination of the peptide motilin.

III. Prospective

In the above section, several deorphanizing stories were introduced that happened in the past several years. Methods have been developed from the first identification of GPCR by protein purification to the high-throughput screening with massive ligand candidates. Now receptor purification is not necessary any more. It is even easy to get GPCR DNAs by molecular biology techniques. However, still much labor is required for testing each receptor in oocytes clamping or purifying candidate ligands from tissues. The high-throughput screening technique is a semi-automated effective method, but has a limitation of source chemicals. Another benefit of high-throughput techniques is the application for pharmacodynamic studies of many drugs, that is, making dose-response curves for agonists and antagonists in a short time and in a reliable manner.

Finding the receptor molecules for the known ligands is undertaken not only to fulfill scientific curiosity (especially of pharmacologists), but also to open new pharmacological opportunities such as molecular-level study of drug-receptor interaction, receptor pharmacogenetics, and receptor structure-based drug design. The closer the pool of ligands has sunk towards the bottom, the more orphan receptors have been found by chance. From the late 20th century, many pharmaceutical companies launched orphan GPCR projects and made great contributions to the discovery of important natural ligands. All these findings are now followed by intensive studies on the physiological and pathological meanings of the novel ligands (5). Still many orphan receptors are waiting for the discovery of their natural ligands by scientists. Information of the cloned orphan receptors and putative GPCRs is available from internet-based databases (NCBI://www.ncbi.nlm.nih.org, EBI://www.ebi.ac.uk, JCRB://www.nih.go.jp). A recent review also helpfully grouped them by identity to ligand-known receptors (30). For drug discovery, however, we have to continue basic and clinical research with undiminished enthusiasm until new medicines are found. Discovery of novel natural ligands is not the final goal for reverse pharmacology. The aims are to investigate the physiological roles of newly identified ligands, to discover their implications in diseases, and to develop medicines, mimicking or antagonizing the ligands, for the cure of diseases and for the welfare of human beings. This whole process furthermore will automatically accompany the progress of sciences, including biology, physiology, pathology and clinical sciences.

REFERENCES

1 Dixon RA, Kobilka BK, Strader DJ, Benovic JL, Dohman HG, Freillie T, Bolanowski MA, Bennett CD, Rands E, Diehl RE and et al.: Cloning of the gene and cDNA for mammalian beta-adrenergic receptor and homology with rhodopsin. Nature 321, 75 – 79 (1986)
2 Kubo T, Fukuda K, Mikami A, Maeda A, Takahashi H, Mishina M, Haga T, Haga K, Ichiyama A, Kangawa K and et al.: Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. Nature 323, 411 – 416 (1986)
3 Marchese A, George SR and O’Dowd BF: Cloning of G protein-coupled receptor genes: the use of homology screening and the polymerase chain reaction. In Identification and Expression of G Protein-Coupled Receptors, Edited by Lynch KR, Vol I, pp 1 – 26, Wiley-Liss, New York (1998)
4 Myers EW, Sutton GG, Smith HO, Adams MD and Venter JC: On the sequencing and assembly of the human genome. Proc Natl Acad Sci USA 99, 4145 – 4146 (2002)
5 Civelli O, Nothacker HP, Saito Y, Wang Z, Lin SH and Reinscheid RK: Novel neurotransmitters as natural ligands of orphan G-protein-coupled receptors. Trends Neurosci 24, 230 – 237 (2001)
6 Oliveira L PA and Vriend G: A common motif in G-protein-coupled seven transmembrane helix receptors. Journal of Computer-Aided Molecular Design 7, 649 – 658 (1993)
7 Okamoto T, Murayama Y, Hayashi Y, Inagaki M, Ogata E and Nishimoto I: Identification of a Gs activator region of the beta 2-adrenergic receptor that is autoregulated via protein kinase A-dependent phosphorylation. Cell 67, 723 – 730 (1991)
8 Hla T, Lee MJ, Ancellin N, Paik JH and Kluk MJ: Lysophospholipids–receptor revelations. Science 294, 1875 – 1878 (2001)
9 Lynch KR and Im DS: Life on the edg. Trends Pharmacol Sci 20, 473 – 475 (1999)
10 Lee MJ, Van Brocklyn JR, Thangada S, Liu CH, Hand AR, Menzelev R, Spiegel S and Hla T: Sphingosine-1-phosphate as a ligand for the G protein-coupled receptor EDG-1. Science 279, 1552 – 1555 (1998)
11 Okamoto H, Takuwa N, Yatomi Y, Gonda K, Shigematsu H and Takuwa Y: EDG3 is a functional receptor specific for sphingosine 1-phosphate and sphingosylphosphorylcholine with
signaling characteristics distinct from EDG1 and AGR16. Biochem Biophys Res Commun 260, 203 – 208 (1999)

12 Im DS, Heise CE, Ancellin N, O'Dowd BF, Shi GJ, Heavens RP, Rigby MR, Hla T, Mandala S, McAllister G, George SR and Lynch KR: Characterization of a novel sphingosine 1-phosphate receptor, Edg-8. J Biol Chem 275, 14281 – 14286 (2000)

13 Yamazaki Y, Kon J, Sato K, Tomura H, Sato M, Yoneya T, Okazaki H, Okajima F and Ohta H: Edg-6 as a putative sphingosine 1-phosphate receptor coupling to Ca²⁺ signaling pathway. Biochem Biophys Res Commun 268, 583 – 589 (2000)

14 Van Broeklyn JR, Graier MH, Bernhardt G, Hobson JP, Lipp M and Spiegel S: Sphingosine-1-phosphate is a ligand for the G protein-coupled receptor EDG-6. Blood 95, 2624 – 2629 (2000)

15 Mandala S, Hajdu R, Bergstrom J, Quackenbush E, Xie J, Brinkmann V, Davis MD, Heise CE, Albert R, Cottens S, Hof Van Brocklyn JR, Graler MH, Bernhardt G, Hobson JP, Lipp M and Spiegel S: Sphingosine-1-phosphate is a ligand for the G protein-coupled receptor EDG-6. Blood 95, 2624 – 2629 (2000)

23 Meunier JC, Mollereau C, Toll L, Sautadeau C, Moisand C, Alvinerie P, Butour JL, Guillenet JC, Ferrara P, Monsarrat B and et al.: Isolation and structure of the endogenous agonist of opioid receptor-like ORL1 receptor. Nature 377, 532 – 535 (1995)

24 Sakurai T, Amemiya A, Ishii M, Matsuizaki I, Chemelli RM, Tanaka H, Williams SC, Richardson JA, Kozlowski GP, Wilson S, Arch JR, Buckingham RE, Haynes AC, Carr SA, Annan RS, McNulty DE, Liu WS, Terrett JA, Elshourbagy NA, Bergsma DJ and Yanagisawa M: Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. Cell 92, 573 – 585 (1998)

25 Hinuma S, Habata Y, Fujii R, Kawamata Y, Hosoya M, Fukusumi S, Kitada C, Masuo Y, Asano T, Matsumoto H, Sekiguchi M, Kurokawa T, Nishimura O, Onda H and Fujino M: A prolactin-releasing peptide in the brain. Nature 393, 272 – 276 (1998)

27 O'Dowd BF, Heiber M, Chan A, Heng HH, Tsui LC, Kennedy JL, Shi X, Petronis A, George SR and Nguyen T: A human gene that shows identity with the gene encoding the angiotensin receptor is located on chromosome 11. Gene 136, 355 – 360 (1993)

28 Chambers JK, Macdonald LE, Sarau HM, Ames RS, Freeman K, Foley JJ, Zhu Y, McLaughlin MM, Murdock P, McMillan L, Trill J, Swift A, Aiyar N, Taylor P, Vawter L, Naheed S, Seckeres P, Hervieu G, Scott C, Watson J, Murphy AJ, Duzic E, Klein C, Bergsma DJ, Wilson S and Livi GP: A G protein-coupled receptor for UDP-glucose. J Biol Chem 275, 10767 – 10771 (2000)

29 Feighner SD, Tan CP, McKay KK, Palyha OC, Hreniuk DL, Pong SS, Austin CP, Figueroa D, MacNeill D, Cascieri MA, Nargund R, Bakshi R, Abramovitz M, Stocco R, Kargman S, O'Neill G, Van Der Ploeg LH, Evans J, Patchett AA, Smith RG and Howard AD: Receptor for motilin identified in the human gastrointestinal system. Science 284, 2184 – 2188 (1999)

30 Lee DK, George SR, Evans JF, Lynch KR and O'Dowd BF: Orphan G protein-coupled receptors in the CNS. Curr Opin Pharmacol 1, 31 – 39 (2001)