Diversification and Evolution of the Signal Transduction System : Rapid Divergence of Tissue Specific Genes in Early Evolution of Chordates

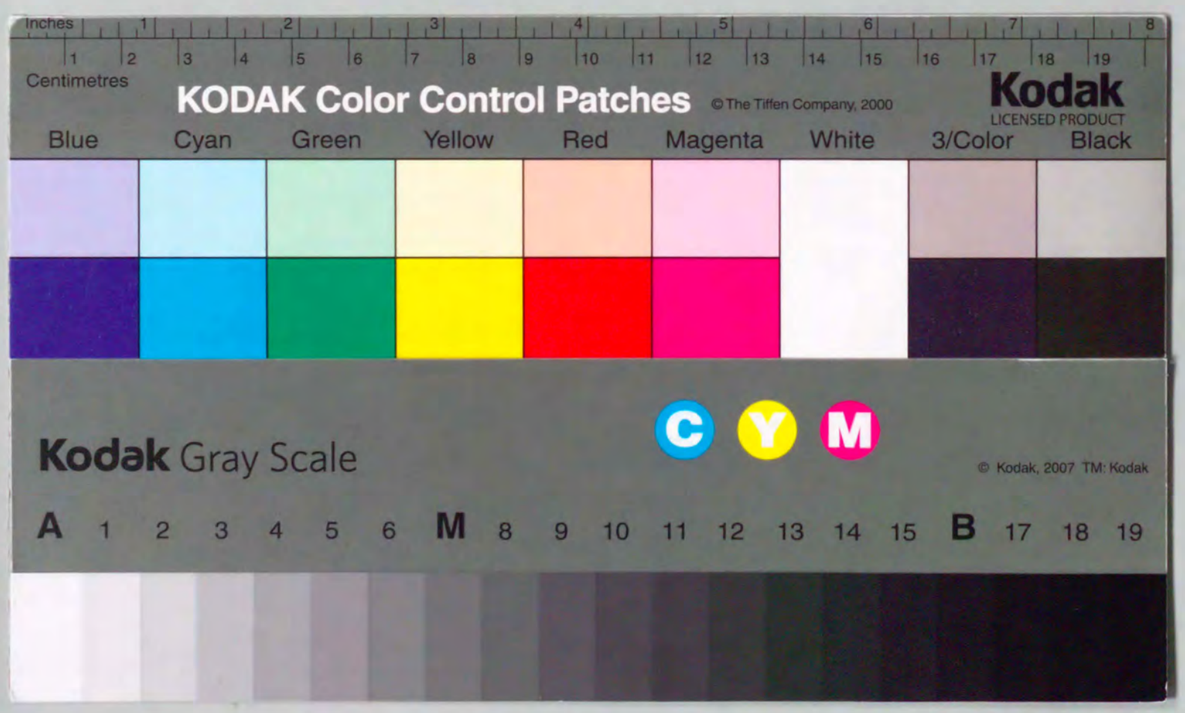
岩部, 直之
九州大学理学研究科生物学専攻

https://doi.org/10.11501/3073259

出版情報：九州大学, 1993, 博士（理学）, 課程博士
バージョン：
権利関係：

Diversification and Evolution of the Signal Transduction System: Rapid Divergence of Tissue Specific Genes in Early Evolution of Chordates

東 郁之



Diversification and Evolution of the Signal Transduction System: Rapid Divergence of Tissue Specific Genes in Early Evolution of Chordates

Naoyuki IWABE

Department of Biology
Faculty of Science
Kyushu University
Fukuoka 812, Japan

Contents

Summary

Recently, many genes of key molecules playing important roles in the signal transduction systems have been cloned and sequenced. These key components, including rhodopsin family receptors (G protein coupled receptors), G proteins, adenylyl cyclases, phosphodiesterases, and phospholipase Cs, comprise large families whose genes were derived from each ancestral precursor by gene duplications. In order to clarify how the genes in the signal transduction systems diverged during the evolution of organisms, especially of animals, molecular phylogenetic trees of the receptors, G protein α subunits, adenylyl cyclases, phosphodiesterases, and phospholipase Cs were reconstructed. On the phylogenetic tree of 124 rhodopsin family receptors, those with similar ligand binding specificities made clusters. The phylogenetic tree of 45 G protein α subunits was also inferred. On this tree, G protein α subunits regulating a specific type of amplifier also made a cluster. On both receptor and G protein α subunit trees, genes with different functions are shown to have diverged before the separation of vertebrate (or deuterostomia) and insect (or protostomia) lineages , while genes duplicated after the divergence of the two animal lineages seems to have same function but different tissue distributions. The receptors binding with same ligand not necessarily interact with similar G proteins. For example, one subtype of muscarinic acetylcholine receptor (mAChR), m1, activating G protein subtype q and then stimulating PLC β subtypes in the next signal transducing step, whereas the other mAChR subtype, m2, activating G protein subtype i and then inhibiting adenylyl cyclases. The two receptors are closely related on the phylogenetic tree but the two G protein subtypes are distantly related. The duplicated genes of receptors binding with same ligand but coupling with different types of G proteins seem to guarantee the development of various types of cells and tissues in both vertebrate and insect lineages. The possible combinations of the signal pathways mediated by the receptors, G protein subunits, and

amplifiers known at present are enormous number even though many of them have never been realized. In the case of the animal amplifiers of G proteins, it becomes clear from the molecular phylogenetic analyses that subtypes of the amplifiers regulated by G proteins, like adenylyl cyclases, cGMP phosphodiesterases, and phospholipase C β subtype, were diverged from non-G protein coupled types, like guanylyl cyclase, calmodulin-dependent phosphodiesterase, and phospholipase C γ subtype, respectively, before the two animal lineages were separated.

It is generally thought that evolutionary changes at phenotypic and molecular levels are caused by distinct mechanisms, Darwinian selection of selectively positive mutants on one hand and random fixation of selectively neutral mutants on the other hand (Kimura, 1983). It is of particular importance to understand evolution at two levels in a unified way, but no data suggesting a link between them has been provided to date. The explosive diversification of metazoa at the early Cambrian would provide a unique opportunity to find a clue for the problem: It is widely accepted that the invasion of new vacant niches followed by relaxed selective constraints at the phenotypic level is a prerequisite for the rapid evolutionary burst at high taxonomic levels (Simpson, 1967; Kimura, 1991; Valentine, 1977). If this hypothesis is correct, the liberation from selective constraints at the phenotypic level is expected to reduce functional constraints at the molecular level, giving rise to rapid accumulations of genetic variations, because molecules that are expressed tissue specifically are constrained not only from functional requirements of individual molecules, but also from higher levels like tissues or organs (Kuma et al., 1993), and a partial lack of tissue function results in the elevated rate of molecular evolution (Hendriks et al., 1987). From a phylogenetic analysis of tissue specific genes we report here that gene duplications as well as amino acid substitutions had occurred with rapid rate during the evolution of echinoderm/chordate common ancestor and the early evolution of chordates,

but these genetic changes had remarkably reduced in the later stages. Thus these results are supporting evidence for the above hypothesis from molecular data and suggests a link between evolution at phenotypic and molecular levels.

References

Hendriks, W., Leunissen, J., Nevo, E., Bloemendal, H., and de Jong, W. W. (1987). The lens protein αA-crystallin of the blind mole rat, *Spalax ehrenbergi*: Evolutionary change and functional constraints. Proc. Natl. Acad. Sci. USA *84*, 5320-5324.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.

Kimura, M. (1991). The neutral theory of molecular evolution: A review of recent evidence. Jpn. J. Genet. *66*, 367-386.

Kuma, K., Iwabe, N., and Miyata, T. (1993). Global constraint and molecular evolutionary rate: slowly evolving brain specific genes demonstrated by protein kinase and immunoglobulin supergene families. Proc. Natl. Acad. Sci. USA submitted.

Simpson, G. G. (1967). *The Meaning of Evolution. A Study of the History of Life and its Significance for Man.* Rev. ed, Yale University Press, New Heaven and London.

Valentine, J. W. (1977). General patterns of metazoan evolution. in *Patterns of Evolution as Illustrated by the Fossil Record. chapter 2. A. Hallan ed. Developments in Palaeontology and Stratigraphy, 5.* Elsevier Scientific Publishing Company, Amsterdam, Oxford, New York.

Part 1

Diversification and Evolution of the Signal Transduction System: Rhodopsin Family Receptors, G protein α Subunits, Adenylyl Cyclases, Phosphodiesterases, and Phospholipase Cs.

Introduction

Recently, the signal transduction systems converting extracellular signals through G proteins to intracellular second messengers have been enthusiastically studied in molecular level. The proteins participating in the systems were identified and classified into four groups, receptors, transducers, amplifiers, and effectors. As the primary structures of these proteins were revealed more and more, it was found that there were several gene families. Rhodopsin, a photo receptor having seven transmembrane regions, is homologous to the receptors of neurotransmitters, peptide hormones, and odorant substances (for review see Birnbaumer et al., 1990; Iyengar and Birnbaumer, 1990). Heterotrimeric guanine nucleotide binding proteins (or G proteins), consist of α, β, and γ subunits, each make gene families (for review see Iyengar and Birnbaumer, 1990; Kaziro et al., 1990; Kaziro et al., 1991; Simon et al., 1991; Wilkie et al., 1992; Hepler and Gilman, 1992; Birnbaumer, 1992). G proteins regulates amplifiers such as adenylyl cyclase (for review see Tang and Gilman, 1992), phosphodiesterase (for review see Beavo and Reifsnyder, 1990), phospholipase C (for review see Rhee and Choi, 1992), which also make their own large gene families, independently. Second messengers, like cyclic nucleotides (cAMP and cGMP), inositol 1,4,5-trisphosphate, and diacylglycerol, regulates a great number of effectors including protein kinases, channels, calmodulin, etc. Because the members of each gene family of signal transduction systems were diversified by gene duplications, the combinations of these factors became larger and larger, which ensured multicellular organisms to develop many types of cells and tissues. It is very important to study the diversification process of each factor and the interaction of one factor to the other in the cell.

On this study, we paid much attention to the diverging points of vertebrates and insects (and mollusks) on the phylogenetic trees because the divergence time of the two animal lineages seemed to be good measure to

divide the gene duplications into two types. The gene duplications before the divergence time produced the genes of different functions, whereas those after the time generated the genes of same function but different tissue expressions.

Material and Methods

Sequence Data Sources Sequence data were from GenBank Release 77.0 and NBRF (National biochemical Research Foundation) Release 36.0. Accession numbers of the data were listed on tables and figure legends. Data not from the data bases were referred.

Phylogenetic Tree Inference Amino acid sequence data of each gene family were aligned for their homologous regions (Miyata et al., 1985). On the basis of the alignments, the evolutionary distance *k* was calculated between each pair of the sequence, excluding amino acid positions where gaps existed in any one of the aligned sequences; the distance *k* was measured by calculating the amino acid difference *K* (per site) between sequences compared and correcting multiple substitutions as $k = -ln(1 - K)$ (Jukes and Cantor, 1969; Kimura, 1983). Based on the evolutionary distances, phylogenetic trees were inferred by the neighbor-joining (NJ) method (Saitou and Nei, 1987). The reliability of the inferred tree was evaluated by bootstrap probability (Felsenstein, 1985).

Calculation of Nonsynonymous and Synonymous Substitutions Corrected nucleotide differences at the nonsynonymous (k_A) and synonymous (k_S) were calculated (Miyata and Yasunaga, 1980) for the homologous regions of the gene families.

Results and Discussion

Rhodopsin Family Receptors Extracellular signals like neurotransmitters, peptide hormones, odorant molecules, and photons are primarily recognized by the receptors specifically identifying them. After complete amino acid sequences of bovine rhodopsin, the visual pigment, was determined in 1982 (Ovchinnikov et al., 1982), many receptors interacting with G proteins to send the information of the extracellular signals to intracellular second messengers has been identified. Although all these receptors have common structure of seven hydrophobic membrane spanning α-helices, several types of such receptors have no amino acid sequence homology with rhodopsin. Two groups of genes, metabotropic glutamate receptors (Masu et al., 1991) and secretin receptor (Ishihara et al., 1991) and its related genes (Lin et al., 1991; Jüppner et al., 1991; Abou-Samra et al., 1992; Thorens, 1992), are thought to be the G protein coupled receptors having no sequence homology with rhodopsin family receptors and with the members of the other group. Yeast *STE2* and *STE3* (Nakayama et al., 1985) also have seven hydrophobic membrane spanning regions and interact with yeast G protein, *GPA1* (for review see Marsh et al., 1991), but have no homology with the members of other receptor groups and with each other. Bacteriorhodopsin (Ovchinnikov et al., 1979) and halorhodopsin (Blanck and Oesterhelt, 1987) of halophilic archaebacteria (Archaea) also have seven transmembrane regions but no homology with the eukaryotic receptors. *Dictyostelium* cAMP receptor (Klein et al., 1988), which also contains seven transmembrane regions and interacts with G protein, have weak homology with rhodopsin family receptors. In any case, there is no evidence that all the receptors with seven transmembrane structures are the descendants of a common ancestral gene.

On this analysis, we aligned 124 amino acid sequences of opsins and receptors (Table 1) showing apparent homologies with rhodopsin (Fig. 1; Appendix A). *Dictyostelium* cAMP receptor was not included in this alignment

because of its weak homology with rhodopsin. From the alignment figure of consensus sequences of closely related groups (Fig. 1), highly homologous regions of the sequences were shown to be restricted to the seven transmembranes and the nearby regions.

On the basis of the alignment, every gene pair of evolutionary distances was calculated (Jukes and Cantor, 1969; Kimura, 1983) and the phylogenetic tree of the receptors (Fig. 2) was inferred by the neighbor-joining (NJ) method (Saitou and Nei, 1987). Because there was no information about the oldest gene duplication, the tree should be expressed as an unrooted one (Fig. 2(a)). From the phylogenetic tree, we found three important points.

First, receptors for the same ligand or structurally related ligands made clusters on the tree (Iwabe et al., 1989). For example, opsins (or visual pigments), olfactory receptors, and biogenic amine receptors made single clusters on the tree (Fig. 2(b)), respectively. The receptors for low molecular weight neurotransmitters including adenosine, cannabinoid, and biogenic amines are more closely related to each other than to neuropeptide receptors or peptide hormone receptors. These facts indicate that ligand candidates for so-called "orphan" receptors can be assumed when those receptors are included in the phylogenetic tree. Murphy et al. (1992) also pointed out the relationship between clusters of the receptors on their phylogenetic tree and the ligand specificities. The alignment of the consensus sequences of the receptors (Fig. 1), whose grouping of the members was based on the clusters of the phylogenetic tree (Fig. 2(b)), suggests the existence of subgroup specific amino acid sequences, which can be helpful to search the new member of the subgroup and to elucidate their functional specificities. In some cases, the receptors for same ligand consisting a cluster on the tree couple to different effector systems. For example, m1 and m2 muscarinic acetylcholine receptor subtypes make a cluster on the tree (Fig. 2(b)) but couple to different G protein subtypes, q and i, respectively. Opsins of

vertebrates couple to transducins (or G protein αt subtypes) but those of insects to αq subtype (dgq). It seems that after the gene duplications or separation of species, changing the ligand specificities of the receptors was more difficult than changing the specificities for G protein coupling.

Second, on the phylogenetic tree, gene duplications producing the receptors for different ligand specificities occurred before the divergence of vertebrates and insects (and mollusks). For example, gene duplications of muscarinic acetylcholine receptors and histamine H-1 receptor and of opsins and the other receptors preceded the divergence of the two animal lineages (Fig. 2(b)). We will call these receptors of different ligand specificities "anisoforms". Mammalian tachykinin peptide receptor case, however, is an exceptional one because the duplication of substance K, substance P, and neuromedin K receptors occurred after the divergence of vertebrates and insects (Fig. 2(b)). The fact that these receptors can be cross-activated at reduced potency by the other neuropeptide (Nakanishi, 1991) indicates that they are now on the way to accepting the rigid ligand specificities. Because neuropeptides and peptide hormones might have frequently duplicated and diverged after the separation of vertebrates and insects, it seemed appropriate that some subtypes of the receptors became to get modified ligand specificities.

Third, gene duplications producing the receptors for same ligand specificity occurred independently in each lineage of vertebrates and insects (and mollusks). The genes diverged by these duplications have their own tissue or cell specific expressions. For example, color opsins and rhodopsins of vertebrates and opsin Rh1, 2, 3, and 4 of insects diverged independently after the divergence of the two animal lineages (Fig. 2(b)), indicating the independent acquisition of color visions of vertebrates and insects (Zuker et al., 1987; Iwabe et al., 1989; Fryxell and Meyerowitz, 1991) and possibly mollusks (Iwabe et al., 1989). These opsins are expressed in different visual

cells in both vertebrates and insects (Table 1). Vertebrates' muscarinic acetylcholine receptor subtypes m1, 2, 3, 4, and 5 also diverged after the separation of the two animal lineages and the tissue expressions of those subtypes are different to each other (Table 1; Bonner et al., 1987). We will call these receptor subtypes of same ligand specificity but of different tissue (or cell) distributions "tissue specific isoforms". In some cases, like m1 and m2 muscarinic acetylcholine receptor subtypes, G protein coupling specificity of one receptor subtype is different to that of others, as previously discussed. In vertebrate lineage, many of the gene duplications of tissue specific isoforms occurred before the divergence of fish and amphibians (Iwabe et al., 1993; Part 2). In the vertebrate opsin lineage, the evolutionary rate of amino acid substitutions before the separation of lamprey and amphibians (extending 600 myrs ago to 500 myrs ago (Dickerson, 1971)) is apparently higher than that after the separation of the two lineages (Fig. 2(b)). The tendency of high evolutionary rates of the first period is not restricted to opsins but can be seen in many gene families (Iwabe et al., 1993; Part 2).

In any case, the common ancestor of vertebrates and insects possessed a variety of "anisoform" receptors, although it had only one or a few "tissue specific isoform" receptors per one ligand.

Recently, Okano et. al. reported that vertebrate rhodopsins have diverged from cone visual pigments (Okano et. al., 1992). The phylogenetic tree inferred in this study also suggest the same result (Fig. 2(b)). Previously, Iwabe et al. reported that human blue opsin was closely related with rhodopsin than red and green opsins (Iwabe et. al., 1989), while Yokoyama and Yokoyama suggested that blue and red/green opsins made a cluster (Yokoyama and Yokoyama, 1989). The genes of chicken green opsin (Okano et. al., 1992) and gecko opsin P467 (Kojima et. al., 1992) showed sufficient evidences that the genes for scotopic vision (rhodopsin) was evolved from that for photopic vision (color pigments).

Corrected nucleotide differences at the nonsynonymous and synonymous positions (Miyata and Yasunaga, 1980) of the receptor genes compared between human and mouse (or rat) were calculated for their common homologous regions (Table 2). The nonsynonymous substitution values (k_A) were ranging from 0.00 to 0.14, whereas synonymous substitution values (k_S) were from 0.24 to 0.93. The k_A values of thromboxane-A2 receptor, olfactory receptor I15, interleukin-8 receptor, thrombin receptor, and BK-2 bradykinin receptor were seemed to be larger than those of the receptors expressed mainly at brain (Table 1 and 2). Kuma et al. (1993) previously showed that the evolutionary rates of protein kinase genes expressed at immune systems are higher than those expressed at brain, which suggested that tissue specific genes are constrained against amino acid alternations not only from the structural or functional requirements of individual molecules but also from higher level as tissues or organs, where they are expressed. The fact that the receptors with higher k_A values are expressed in immune systems, blood, or peripheral nervous systems may be consistent with the results of Kuma et al. (1993) although it must be considered that the receptors with different ligand specificities have different functional constraints. On the receptors for same ligand, the k_A values varied to some extent. For example, k_A value of β3 adrenergic receptor, 0.052, is about 6 times higher than that of α2B adrenergic receptor, 0.0082. Such differences are expected as a consequence of the constraint from the tissues where receptors are expressed (Kuma et al., 1993), but the tendency that k_A values of the receptors coupling to G protein subtype i are smaller than that of the isoforms coupling to G protein subtype s or q may explain some part of the differences.

G protein α subunits Heterotrimeric guanine nucleotide binding proteins (G proteins), consisting of α, β, and γ subunits, transduce extracellular signals

received by the rhodopsin family receptors to intracelluar signal amplifiers. The α subunits, which catalyze GTP to GDP and mainly interact with the receptors and amplifiers, contain GTP binding domains homologous to those of other GTP binding proteins including ras, rab, rho, ADP ribosylation factor, polypeptide chain elongation factor Tu/1α and G/2, and initiation factor 2.

Forty-five amino acid sequences of G protein α subunits from various eukaryotes including yeast, plants, *Dictyostelium*, and animals (Table 3) were aligned for their almost whole regions (Fig. 3). On the basis of the alignment, the distant matrix was calculated and the molecular phylogenetic tree of the α subunits (Fig. 4) was inferred by NJ method (Saitou and Nei, 1987). Because there was no information about the oldest gene duplication or separation of the species, the tree should be expressed as an unrooted one (Fig. 4(a)). From the phylogenetic tree, we found three important points about animal α subunits, which is consistent with the three ideas about the rhodopsin family receptors, as is stated above. Recently, Yokoyama and Starmer (1992) also inferred phylogenetic trees of animal G protein α subunits.

First, animal α subunits interacting the same type of amplifiers made clusters on the tree. For example, αs and αolf, activating adenylyl cyclases, made a single cluster on the tree, whereas αi1, 2, and 3, inhibiting adenylyl cyclases, made another cluster (Fig. 4(b)). αq, y, 14, 15, and 16, recently probed to activate phospholipase C β subtypes (for review see Sternweis and Smrcka, 1992; Birnbaumer, 1992), also made a cluster (Fig. 4(b)). Two types of transducin α subunits, αt1 and αt2, activating phosphodiesterases on vertebrate photocells, also made a cluster on the tree (Fig. 4(b)). These facts indicates that gastducin, which thought to transduce the signal of tastes molecules to unknown effectors on taste receptor cells (McLaughlin et al., 1992), activates some type of phosphodiesterase because it belongs to the cluster of transducins (Fig. 4(b)). As *Caenorhabditis gpa* genes make a single independent cluster on the tree (Fig. 4(b)), there is a possibility that they

interact with an unknown amplifier protein. The alignment of the consensus sequences of the α subunits (Fig. 3), whose grouping of the members was based on the clusters of the phylogenetic tree (Fig. 4(b)), suggests the existence of specific amino acid sequences for the subgroups which can be helpful to search the new member of the subgroup and to elucidate their functional specificities. There is no relationship between the clusters of the α subunits and the relatedness of the receptors that interact them, which will be discussed later.

Second, gene duplications of the α subunits of different amplifier interactions occurred before the divergence of vertebrates and insects (and mollusks). For example, gene duplications making αi, o, z, and t subtypes preceded the separation of the two animal lineages (Fig. 4(b)). The αs and αq subtypes also duplicated before the divergence. Thus these α subunit subtypes of different amplifier interactions are "anisoforms", a term defined in the above discussion of the receptors. Although αq and 16, both activating phospholipase C β subtypes (Sternweis and Smrcka, 1992; Birnbaumer, 1992), made a cluster on the tree, the divergence of the two genes was before the separation of vertebrates and insects(Fig. 4(b)). This suggests that the two α subtypes comprise some different functions.

Third, gene duplications of the α subunits interacting the same amplifier occurred in vertebrate lineage and each duplicated gene has different tissue or cell specific expression. For example, the duplication of αs, expressed in many tissues, and αolf, expressed in olfactory sensory neurons, occurred after the divergence of vertebrates and insects (and mollusks) (Fig. 4(b)). Thus these α subunits are "tissue specific isoforms". Furthermore, in vertebrate lineage, the gene duplications of tissue specific isoforms, like αs and αolf or αi1, 2, and 3 (Fig. 4(b)), occurred before the divergence of amphibians and mammals. In the vertebrate αs lineage, the evolutionary rate of amino acid substitutions before the separation of amphibians and reptiles (extending 600

myrs ago (Dickerson, 1971) to 350 myrs ago (Dayhoff, 1978)) is apparently higher than that after the separation of the two lineages (Fig. 4(b)). The tendency of higher evolutionary rates before the divergence of fish and amphibians than those after it can be seen also in many other genes (Iwabe et al., 1993; Part 2).

In the case of the G protein α subunits, the common ancestor of vertebrates and insects possessed a variety of "anisoforms" but had only one (or a few) "tissue specific isoforms" (Fig. 4(b)), like the case of the receptors (Fig. 2(b)).

The mammalian gene of G protein α subunit subtype o (αo) produces two alternatively spliced mRNA encoding αo1 and αo2 that contain different last two exons (exons 7 and 8) (Tsukamoto et al., 1991). As the amino acid sequences of the exons 7 and 8 regions of αo1 show strong homology with those of αo2 and to other α subunit subtypes, the sequences of these α subunits were aligned (Appendix B'(a)). The phylogenetic tree of the αo subunit exons 7 and 8 regions inferred on the basis of the alignment (Appendix B'(b)) indicated that duplication of the two exons of αo1 and αo2 was occurred before the divergence of vertebrates and insects (and mollusks). Because the exons 7 and 8 of αo1 were paralogous to the other animal αo subunits, we included human αo2 in the phylogenetic tree (Fig. 4(b)). Yokoyama and Starmer (1992) included both αo1 and αo2 in their phylogenetic trees, which misled their results. There is a possibility that some functions of αo1 and αo2 are different because the duplication was before the separation of vertebrates and insects.

Although the root of the phylogenetic tree was not determined, no G protein α subunit of fungi, plants, and *Dictyostelium* was closely related to animal αs, q, i, o, z, or t (Fig. 4(b)). There is a possibility that α subunit "anisoforms" regulating adenylyl cyclases, phosphodiesterases, and

phospholipase Cs diverged only in animal lineages. The evolutionary relationships of G proteins and the amplifiers will be discussed later.

Corrected nucleotide differences at the nonsynonymous and synonymous positions (Miyata and Yasunaga, 1980) of the G protein α subunit genes compared between human and mouse (or rat) were calculated for their common homologous regions (Table 4). The k_A values were ranging from 0.00 to 0.065, whereas k_S values were from 0.29 to 0.97. Apparently the k_A value of α16, expressed specifically in myeloid cells (Table 3; Hepler and Gilman, 1992), was larger than those of the other α subunit genes. This may be consistent with the results of Kuma et al. (1993) that the evolutionary rates of protein kinase genes expressed at immune systems are higher than those expressed at brain.

Amplifiers The extracellular signals received by the receptors are finally converted into the intracellular second messengers by signal amplifiers. There has been identified and sequenced several types of amplifiers: Adenylyl cyclase (AC), producing cAMP and regulating many effector proteins like A kinases or cAMP-gated channels; phosphodiesterase (PDE), hydrolyzing cGMP to GMP and closing the cGMP-gated channels; phospholipase C (PLC), producing two types of second messengers, inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol (DG) and regulating calmodulins and C-kinases. We inferred the molecular phylogenetic trees of these three amplifier families (Fig. 5, 6, and 7) by NJ method (Saitou and Nei, 1987).

Animal ACs comprise common homologous structure of two repeats of [6 transmembrane regions plus one catalytic domain]-structure and are regulated by G proteins. The phylogenetic tree of animal ACs (Fig. 5) was inferred by using *Dictyostelium* AC-A as an outgroup, which have overall homology with them. The guanylyl cyclases and yeast, *Trypanosoma*, and *Rhizobium* ACs also contain the homologous catalytic domains but are not

regulated by G proteins (for review see Tang and Gilman, 1992). They were not included in the tree (Fig. 5) because their structures except for the catalytic domains are completely different from those of animal ACs. The branching patterns of the phylogenetic tree of animal ACs, which was also inferred by maximum likelihood method (data not shown) (Kishino et al., 1990), indicates that the common ancestor of vertebrates and insects possessed at least four types of ACs, I and rutabaga, II and IV, III, and V and VI (Fig. 5). The results of rhodopsin family receptors and G protein α subunits that the gene duplications before the divergence of vertebrates and insects (and mollusks) usually produced "anisoforms" suggest the functional differences of those four types of ACs. The fact that the effects of G protein $\beta\gamma$ subunit complex and Ca^{2+}-calmodulin to these animal four types are different to each other (Tang and Gilman, 1992) is consistent with the above suggestion.

Figure 6 is the phylogenetic tree of PDE catalytic subunits. Mouse photocell cGMP PDE α and β subunits and bovine photocell cGMP PDE are regulated by G protein αt subunits (or transducins), whereas other PDEs are regulated by other factors like cGMP or calmodulin (for review see Beavo and Reifsnyder, 1990). These PDEs regulated by different factors seems to have generated by gene duplications before the divergence of vertebrates and insects (Fig. 6). They can be defined as "anisoforms". On the other hand, four types of mammalian cAMP PDEs, whose functions are same but expressions are different to each other (Swinnen et al., 1989), had generated by gene duplications after the divergence of vertebrates and insects (Fig. 6). They can be defined as "tissue specific isoforms".

Three types of animal PLC, β, γ, and δ, apparently diverged before vertebrates and insects separated (Fig. 7). PLCβ, γ, and δ can be defined as "anisoforms" because they are regulated by G proteins, receptor type protein kinases, and unknown factors, respectively (for review see Rhee and Choi, 1992). Furthermore, because three PLC β subtypes, [mammalian β1, 2, and

3], [*Drosophila plc-21*], and [bovine retina β and *Drosophila norpA*], also diverged before the separation of the two animal lineages (Fig. 7), some functions of the three subtypes are possibly different to each other, like the case of animal ACs.

There are four points that can be seen from the three phylogenetic trees of the amplifiers (Fig. 5, 6, and 7). First, genes with same or similar functions made clusters on the trees. Second, gene duplications generating "anisoforms", gene products of different functions, occurred before the divergence of vertebrates and insects. Third, gene duplications producing "tissue specific isoforms" occurred after the separation of the two animal lineages. Fourth, G protein coupled types of amplifiers are possibly generated by gene duplications of G protein non-coupled types before the separation of the two animal lineages (Fig. 6 and 7). Although the root of the trees should be determined to clarify the relationships of the G protein coupled and non-coupled types of amplifiers, the phylogenetic trees of G protein α subunits (Fig. 4) and the amplifiers (Fig. 6 and 7) strongly suggest that the G protein coupled type amplifiers and G protein α subunits regulating them were generated by the gene duplications in early animal lineages. If plants and fungi also have G protein coupled type amplifiers, it is possible that they produced the subtypes by gene duplications in their own lineages. In any case, to understand when these systems were established, the G protein α subunit and amplifier genes of primitive animals like sponge, jellyfish, and planarian and of many plants and fungi will be very important.

Diversification of the Signal Transduction System Figure 8 is the schematic phylogenetic relationships of the receptors and G protein α subunits extracted from the trees of Figures 2(b) and 4(b). The figure indicates that there is little relationship between the phylogeny of receptors and that of G protein α subunits from the view point of the coupling of the two factors.

Closely related receptors not necessarily couple to the same or closely related G protein α subunits. Furthermore, in some cases, a receptor couples to more than two G proteins (Ross, 1989; Birnbaumer et al., 1990). After gene duplication, one of the duplicated receptors changed its coupling affinity to the G protein, which guaranteed the development of complex tissues or cell types that differently respond to the same ligand.

Gene duplications not only of the receptors but also of G protein α, β, and γ subunits, adenylyl cyclases, phosphodiesterases, and phospholipases contributed the diversification of the signaling pathways (Ross, 1989; Birnbaumer et al., 1990). The possible combination of the signal pathways of the receptors, G protein subunits, and amplifiers known at present is more than hundred thousands although many of them have never been realized. Some signal pathways are listed on Figure 9. Many ligands and the receptors had been acquired when the unicellular animal ancestor became multicellular organism with many cell types, whereas the second messengers like cAMP (and/or cGMP), IP_3, and DG were already used by the common ancestor of animals, fungi, and plants (Fig. 6 and 7). The G proteins combined the newly synthesized ligand perception systems with the preexisting second messenger regulating systems. The G proteins also directly regulate many ion channels, which were not discussed in this study (for review see Hepler and Gilman, 1992; Birnbaumer, 1992). The diversification of signal pathways by combinations of the duplicated gene products can be also seen in the protein kinase, phosphatase, and other systems.

Concluding Remarks As previously discussed, it seems very informative to pay attention to the divergence of vertebrates and insects (or mollusks) on the phylogenetic trees of signal transducing factor families. Gene duplications producing the proteins of different functions occurred before the divergence of the two animal lineages, whereas those producing the proteins of same

functions but different tissue distributions did after the divergence. This tendency is also seen in house keeping genes like peptide elongation factor EF-1α and aldolase. Gene duplications of the latter case in the vertebrate lineage often happened before the divergence of fish and amphibians. During this period (from about 600 myrs ago to 400 myrs ago), molecular evolutionary rates not only of the genes of signal transduction system but also of EF-1α and aldolase were higher than those after this time (from 400 myrs ago to now) (Iwabe et al., 1993; Part 2). This was probably related to the development of complex tissues of early chordates, and possibly of arthropods and other animal lineages, at the early Cambrian period (Iwabe et al., 1993; Part 2).

To clarify the detail processes of the diversification of signal transduction systems of eukaryotes, gene data from primitive animals, fungi, plants, and protista are very important.

References

Abou-Samra, A.-B., Jüppner, H., Force, T., Freeman, M. W., Kong, X. F., Schipani, E., Urena, P., Richards, J., Bonventre, J. V., Potts, Jr. J. T., Kronenberg, H. M., and Segre, G. V. (1992). Expression cloning of a common receptor for parathyroid hormone and parathyroid hormone-related peptide from rat osteoblast-like cells: A single receptor stimulates intracellular accumulation of both cAMP and inositol triphosphates and increases intracellular free calcium. Proc. Natl. Acad. Sci. USA *89*, 2732-2736.

Albert, P. R., Zhou, Q.-Y., Van To, H. H. M., Bunzow, J. R., and Civelli, O. (1990). Cloning, functional expression, and mRNA tissue distribution of the rat 5-hydroxytryptamine 1A receptor gene. J. Biol. Chem. *265*, 5825-5832.

Beavo, J. A. and Reifsnyder, D. H. (1990). Primary sequence of cyclic nucleotide phosphodiesterase isozymes and the design of selective inhibitors. Trends in Pharmacol. Sci. *11*, 150-155.

Birnbaumer, L. (1992). Receptor-to-effector signaling through G proteins: roles for βρ dimers as well as α subunits. Cell *71*,1069-1072.

Birnbaumer, L., Abramowitz, J., and Buckley, N. J. (1990). Receptor effector coupling by G-proteins. Biochem. Biophys. Acta *1031*, 163-224.

Blanck, A. and Ostehelt, D. (1987). The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. EMBO J. *6*, 265-273.

Bonner, T. I., Buckley, N. J., Young, A. C., and Brann, M. R. (1987). Identification of a family of muscarinic acetylcholine receptor genes. Science *237*, 527-532.

Dayhoff, M. O. (1978). in *Atlas of protein sequence and structure*. volume 5, supplement 3, p. 3.

Dickerson, R. E. (1971). The structure of cytochrome *c* and the rates of molecular evolution. J. Mol. Evol. *1*, 26-45.

Drayer, A. L. and van Haastert, P. J. M. (1992) Molecular cloning and expression of]a phosphoinositide-specific phospholipase C of *Dictyostelium discoideum*. J. Biol. Chem. *267*, 18387-18392.

Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution *39*, 783-791.

Fryxell, K. J. and Meyerowitz, E. M. (1991). The evolution of rhodopsins and neurotransmitter receptors. J. Mol. Evol. *33*,367-378.

Gantz, I., Schäffer, M., DelValle, J., Logsdon, C., Campbell, V., Uhler, M., and Yamada, T. (1991). Molecular cloning of a gene encoding the histamine H2 receptor. Proc. Natl. Acad. Sci. USA *88*, 429-433.

Hepler, J. R. and Gilman, A. G. (1992). G protein. Trends in Biochem. Sci. *17*, 383-387.

Hershey, A. D., Dykema, P. E., and Krause, J. E. (1991). Organization, structure, and expression of the gene encoding the rat substance P receptor. J. Biol. Chem. *266*, 4366-4374.

Hirata, M., Hayashi, Y., Ushikubo, F., Yokota, Y., Kageyama, R., Nakanishi, S., and Narumiya, S. (1991). Cloning and expression of cDNA for a human thromboxane A2 receptor. Nature *349*, 617-620.

Honda, Z., Nakamura, M., Miki, I., Minami, M., Watanabe, T., Seyama, Y., Okado, H., Toh, H., Ito, K., Miyamoto, T., and Shimizu, T. (1991). Cloning by functional expression of platelet-activating factor receptor from guinea-pig lung. Nature *349*, 342-346.

Ishihara, T., Nakamura, S., Kaziro, Y., Takahashi, T., Takahashi, K., and Nagata, S. (1991). Molecular cloning and expression of a cDNA encoding the secretin receptor. EMBO J. *10*, 1635-1641.

Iwabe, N., Kuma, K., Nikoh, N., and Miyata, T. (1993). Rapid Divergence of Tissue Specific Genes in Early Evolution of Chordates. Nature submitted

Iwabe, N., Kuma, K., Saitou, N., Tsuda, M., and Miyata, T. (1989). Evolution of rhodopsin supergene family. Independent divergence of visual pigments in vertebrates and insects and possibly in mollusks. Proc. Japan Acad. *65(B)*, 195-198.

Iyengar, R. and Birnbaumer, L. (1990). Overview. pp.1-14. in R. Iyengar and L. Birnbaumer eds. *G proteins*. Academic Press, San Diego, California.

Jones, D. T. and Reed, R. R. (1989). G_{olf}: an olfactory neuron specific-G protein involved in odorant signal transduction. Science *244*, 790-795.

Jukes, T. F. and Cantor, C. R. (1969). Evolution of protein molecules. pp.21-132. in *Mammalian Protein Metabolism III* ed. H. N. Munro, Academic Press, New York.

Julius, D., MacDermott, A. B., Axel, R., and Jessell, T. M. (1988). Molecular characterization of a functional cDNA encoding the serotonin 1c receptor. Science *241*, 558-564.

Jüppner, H., Abou-Samra, A.-B., Freeman, M. , Kong, X. F., Schipani, E., Richards, J., Kplakowski, Jr. L. F. Hock, J., Potts, Jr. J. T., Kronenberg, H. M., and Segre, G. V. (1991). A G-protein-linked receptor for parathroid hormone and parathroid hormone-related peptide. Science *254*, 1024-1026.

Kaziro, Y., Itoh, H., and Nakafuku, M. (1990). Organization of genes coding for G-protein α subunits in higher and lower eukaryotes. pp.63-80. in R. Iyengar and L. Birnbaumer eds. *G proteins*. Academic Press, San Diego, California.

Kaziro, Y., Itoh, H., Kozasa, T., Nakafuku, M., and Satoh, T. (1991). Structure and function of signal transducing GTP-binding proteins. Annu. Rev. Biochem. *60*, 349-400.

Kimura, M. (1983). pp.55-97. in *The Neutral Theory of Molecular Evolution*, Cambridge Univ. Press, Cambridge, England.

Kishino, H., Miyata, T., and Hasegawa, M. (1990). Maximum likelihood inference of protein phylogeny, and the origin of chloroplasts. J. Mol. Evol. *31*, 151-160.

Klein, P. S., Sun, T. J., Saxe III, C. L., Kimmel, A. R., Johnson, R. L., and Devreotes, P. N. (1988). A chemoattractant receptor controls development in Dictyostelium discoideum. Science *241*, 1467-1472.

Kojima, D., Okano, T., Fukada, Y., Shichida, Y., Yoshizawa, T., and Ebrey, T. G. (1992). Cone visual pigments are present in gecko rod cells. Proc. Natl. Acad. Sci. USA *89*,6841-6845.

Kubo, T., Fukuda, K., Mikami, A., Maeda, A., Takahashi, H., Mishina, M., Haga, T., Haga, K., Ichiyama, A., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T., and Numa, S. (1986). Cloning, sequencing and expression of complementary DNA encoding the muscarinic acetylcholine receptor. Nature *323*, 411-416.

Kuma, K., Iwabe, N., and Miyata, T. (1993). Global constraint and molecular evolutionary rate: slowly evolving brain specific genes demonstrated by protein kinase and immunoglobulin supergene families. Proc. Natl. Acad. Sci. USA submitted.

Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic characters of a protein. J. Mol. Biol. *157*, 105-132.

Libert, F., Parmentier, M., Lefort, A., Dinsart, C., Van Sande, J., Maenhaut, C., Simons, M.-J., Dumont, J. E., and Vassart, G. (1989). Selective amplification ad cloning of four new members of the G protein-coupling receptor family. Science *244*, 569-572.

Lin, H. Y., Harris, T. L., Flannery, M. S., Aruffo, A., Kaji, E. H., Gron, A., Kolakowski, Jr. L. F., Londish, H. F., and Goldring, S. R. (1991).

Expression cloning of an adenylate cyclase-coupled calcitonin receptor. Science *254*, 1022-1024.

Lochrie, M. A., Mendel, J. E., Sternberg, P. W., and Simon, M. I. (1991). Homologous and unique G protein alpha subunits in the nematode Caenorhabditis elegans. Cell Regul. *2*,135-154.

Lomasney, J. W., Cotecchia, S., Lorenz, W., Leung, W.-Y., Schwinn, D. A., Yang-Feng, T. L., Brownstein, M., Lefkowitz, R. J., and Caron, M. G. (1991). Molecular cloning and expression of the cDNA for the alpha 1A-adrenergic receptor. J. Boli. Chem. *266*, 6365-6369.

Loosfelt, H., Misrahi, M., Atgr, M., Salesse, R., Thi, M. T. V. H.-L., Jolivet, A., Guiochon-Mantel, A., Sar, S., Jallal, B., Garnier, J., and Milgrom, E. (1989). Cloning and sequencing of porcine LH-hCG receptor cDNA: variants lacking transmembrane domain. Science *245*, 525-528.

Marsh, L., Neiman, A. M., and Herskowitz, I. (1991). Signal transduction during pheromone response in yeast. Annu. Rev. Cell Biol. *7*, 699-728.

Masu, M., Tanabe, Y., Tsuchida, K., Shigemoto, R., and Nakanishi, S. (1991). Sequence and expression of a metabotropic glutamate receptor. Nature *349*, 760-765.

McFarland, K. C., Sprengel, R., Phillips, H. S., Köhler, M., Rosemblit, N., Nikolics, K., Segaloff, D. L., and Seeburg, P. H. (1989). Lutropin-choriogonandotropin receptor: an unusual member of the G protein-coupled receptor family. Science *245*, 494-499.

McLaughlin, S. K., McKinnon, P. J., and Margolskee, R. F. (1992). Gustducin is a taste-cell-specific G protein closely related to the transducins. Nature *357*, 563-569.

Miyata, T. and Yasunaga, T. (1980). Molecular evolution of mRNA: A method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. J. Mol. Evol. *16*, 23-36.

Miyata, T., Hayashida, H., Kikuno, R., Toh, H., and Kawade, Y. (1985). Evolution of interferon genes. pp.1-30. in *Interferon 6*, I. Grosser ed. Academic Press, London.

Montell, C., Jones, K., Zuker, C., and Rubin, G. (1987). A second opsin gene expressed in the ultraviolet-sensitive R7 photoreceptor cells of *Drosophila melanogaster*. J. Neurosci. *7*, 1558-1566.

Murphy, P. M., Özçelik, T., Kenney, R. T., Tiffany, H. L., McDermott, D., and Francke, U. (1992). A structural homologue of N-formyl peptide receptor. J. Biol. Chem. *267*, 7637-7643.

Mussin, P., Revelli, J.-P., Kuhne, F., Gocayne, J. D., McCombie, W. R., Venter, J. C., Giacobino, J.-P., and Fraset, C. M. (1991). An adipose tissue-specific beta-adrenergic receptor. Molecular cloning and down-regulation in obesity. J. Biol. Chem. *266*, 24053-24058.

Nakanishi, S. (1991). Mammalian tachykinin receptors. Annu. Rev. Neurosci. *14*, 123-136.

Nakayama, N., Miyajima, A., and Arai, K. (1985). Nucleotide sequences of *STE2* and *STE3*, cell type-specific sterile genes from *Saccharomyces cerevisiae*. EMBO J. *10*, 2643-2648.

Okano, T., Kojima, D., Fukada, Y., Shichida, Y., and Yoshizawa, T. (1992). Primary structures of chicken cone visual pigments: Vertebrate rhodopsins have evolved out of cone visual pigments. Proc. Natl. Acad. Sci. USA *89*,5932-5936.

Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Artamonov, I. D., Zolotereve, A. S., Kostina, M. B., Bogachuk, A. S., Miroshinikov, A. I., Martinov, V. I., and Kudelin, A. B. (1982). The complete amino acid sequence of visual rhodopsin. Bioorg. Khim. *8*, 1011-1014.

Ovchinnikov, Y. A., Abdulaev, N. G., Feigina, M. Y., Kiselev, A. V., and Lobanov, N. A. (1979). The structural basis of the functioning of bacteriorhodopsin: An overview. FEBS Lett. *100*, 219-224.

Rhee, S. G. and Choi, K. D. (1992). Regulation of inositol phospholipid-specific phospholipase C isozymes. J. Biol. Chem. *267*,12393-12396.

Ross, E. M. (1989). Signal sorting and amplification through G protein-coupled receptors. Neuron *3*, 141-152.

Saitou, N. and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic tree. Mol. Biol. Evol. *4*, 406-425.

Simon, M. I., Strathmann M. P., and Gautam, N. (1991). Diversity of G proteins in signal transduction. Science *252*, 802-808.

Sprengel, R., Braun, T., Nikolics, K., Segaloff, D. L., and Seeburg, P. H. (1990). The testicular receptor for follicle stimulating hormone: structure and functional expression of cloned cDNA. Mol. Endocrinol. *4*, 525-530.

Sternweis, P. C. and Smrcka, A. V. (1992). Regulation of phospholipase C by G proteins. Trends Biochem. Sci. *17*, 502-506.

Swinnen, J. V., Joseph, D. R., and Conti, M. (1989). Molecular cloning of rat homologues of the *Drosophila melanogaster* dunce cAMP phosphodiesterase: evidence for a family of genes. Proc. Natl. Acad. Sci. USA *86*, 5325-5329.

Tang, W.-J. and Gilman, A. G. (1992). Adenylyl cyclases. Cell *70*, 869-872.

Thorens, B. (1992). Expression cloning of the pancreatic β cell receptor for the gluco-incretin hormone glucagon-like peptide 1. Proc. Natl. Acad. Sci. USA *89*, 8641-8645.

Tsukamoto, T., Toyama, R., Itoh, H., Kozasa, T., Matsuoka, M., and Kaziro, Y. (1991). Structure of the human gene and two rat cDNAs encoding the a chain of GTP-binding regulatory protein Go: two different mRNAs are generated by alternative splicing. Proc. Natl. Acad. Sci. USA *88*, 2974-2978.

Wada, E., Way, J., Shapira, H., Kusano, K., Lebacq-Verheyden, A. M., Coy, D., Jensen, R., and Battey, J. (1991). cDNA cloning, characterization, and

brain region-specific expression of a neuromedin-B-preferring bombesin receptor. Neuron *6*, 421-430.

Wilkie, T. M., Gilbert, D. J., Olsen, A. S., Chen, X.-N., Amatruda, T. T., Korenberg, J. R., Trask, B. J., de Jong, P., Reed, R. R., Simon, M. I., Jenkins, N. A., and Copland, N. G. (1992). Evolution of the mammalian G protein α subunit multigene family. Nature Genet. *1*, 85-91.

Yokoyama, S. and Starmer, W. T. (1992). Phylogeny and evolutionary rates of G protein a subunit genes. J. Mol. Evol. *35*, 230-238.

Yokoyama, S. and Yokoyama, R. (1989). Molecular evolution of human visual pigment genes. Mol. Biol. Evol. *6*,186-197.

Young, D., O'Neill, K., Jessell, T., and Wigler, M. (1988). Characterization of the rat mas oncogene and its high-level expression in the hippocampus and cerebral cortex of rat brain. Proc. Natl. Acad. Sci. USA *85*, 5339-5342.

Yu, L., Nguyen, H., Le, H., Bloem, L. J., Kozak, C. A., Hoffman, B. J., Snutch, T. P., Lester, H. A., Davidson, N., and Lübbert, H. (1991). The mouse 5-HT 1C receptor contains eight hydrophobic domains and is X-linked. Mol. Brain Res. *11*, 143-149.

Zucker, C. S., Montell, C., Jones, K., Laverty, T., and Rubin, G. M. (1987). Arhodopsin gene expressed in photoreceptor cell R7 of the *Drosophila* eye: homologies with other signal-transducing molecules. J. Neurosci. *7*, 1550-1557.

Table 1

Expressions and Accession Numbers of Rhodopsin Family Receptors

No.	Gene	Expression	Accession No.
(A) opsin			
1.	human rhodopsin	rod cell	K02281
2.	chicken rhodopsin	rod cell	D00702
3.	lamprey rhodopsin	short photoreceptor cell	M63632
4.	chicken green	cone cell	M92038, M88178
5.	*Gecko* P467	rod cell	M92035
6.	chicken blue	cone cell	M92037
7.	human blue	cone cell	M13299
8.	chicken violet	cone cell	M92039
9.	human red	cone cell	M13305
10.	human green	cone cell	K03494
11.	chicken iodopsin (red)	cone cell	X57490
12.	*Astyanax* R007	(genomic)	M38630
13.	*Astyanax* G101	(genomic)	M38624
14.	*Astyanax* G103	(genomic)	M60945
15.	*Gecko* P521	rod cell	M92036
16.	octopus rhodopsin	visual cell	X07797
17.	*Loligo* rhodopsin	visual cell	S14332 [a]
18.	*Drosophila* Rh1	photoreceptor cell R1-6	K02315
19.	*Calliphora* Rh1	photoreceptor cell R1-6	M58334
20.	*Drosophila* Rh2	photoreceptor cell R8	M12896
21.	*Drosophila* Rh3	photoreceptor cell R7	M17718, Y00043
22.	*Drosophila* Rh4 (ultraviolet)	photoreceptor cell R7	[b]

Table 1 (continued)

No.	Gene	Expression	Accession No.
(B) olfactory (odorant) receptors (putative)			
23.	rat I3	olfactory sensory cell	M64385
24.	rat I8	olfactory sensory cell	M64387
25.	rat 14	olfactory sensory cell	M64391
26.	rat I9	olfactory sensory cell	M64388
27.	rat I15	olfactory sensory cell	M64392
28.	rat F5	olfactory sensory cell	M64377
29.	rat F12	olfactory sensory cell	M64381
30.	rat F3	olfactory sensory cell	M64376
31.	rat I7	olfactory sensory cell	M64386
32.	rat F6	olfactory sensory cell	M64378
33.	human HGMP07I (ligand unknown)	testis	X64994
34.	human HGMP07J (ligand unknown)	testis	X64995
35.	dog DTMT (ligand unknown)	testis (spermatocytes)	X64996
(C) pituitary glycoprotein hormone receptors			
36.	human luteinizing hormone-choriogonadotropic hormone (LH-CG) R.	testis, ovary [c, d]	M63108
37.	human follicle-stimulating hormone (FSH) R.	ovary	M65085
38.	human thyroid stimulatory hormone (TSH; thyrotropin) R.	thyroid	M32215

Table 1 (continued)

No.	Gene	Expression	Accession No.
(D) tachykinin peptide receptors			
(D-1) substance P receptor and others			
39.	human substance P R. (neurokinin 1 R.)	brain, alimentary canal [e], lung, spinal cord	M74290, M81797, M84425, X65177
40.	human substance K R. (neurokinin 2 R.)	lung	M57414, M60284
41.	human neuromedin K R. (neurokinin 3 R.)	brain, kidney	M89473
42.	human putative opioid R. (ligand unknown)	placenta, brain	M84605
43.	*Drosophila* DTKR (tachykinin-like peptide R.)	head	X62711
44.	*Drosophila* NKD (tachykinin R.)	head > body	M77168
(D-2) glucocorticoid-induced receptor and PR4			
45.	mouse glucocorticoid-induced (GI) R. (ligand unknown)	T-cell	M80610
46.	*Drosophila* PR4 (neuropeptide R.)	head, body	M81490
(D-3) neuropeptide Y receptor			
47.	human neuropeptide Y R.	brain	M84755
(E) gastrointestinal hormone receptors			
48.	rat cholecystokinin A (CCK-A) R.	pancreas	M88096
49.	rat cholecystokinin B (CCK-B) R.	brain	M99418
50.	dog gastrin R.	gastric parietal cells, pancreas, brain	M87834

Table 1 (continued)

No.	Gene	Expression	Accession No.
(F) endotheline receptors			
51.	human endotheline-1 (ET-1) R.	arota, lung, atrium, colon, placenta	D90348
52.	human endotheline-B (ET-B) R.	brain > placenta, lung, kidney, adrenal, colon, duodenum	D90402, M74921
(G) bombesin-like peptide receptors			
53.	human gastrin-releasing peptide (GRP) R.	brain, colon, pancreas	M73481
54.	human neuromedin B R.	brain, esophagus [f]	M73482
(H) neurotensin receptor			
55.	rat neurotensin R.	brain > heart, duodenum, intestine, liver	JH0164 [a]
(I) posterior pituitary hormone receptors			
56.	human vasopressin V2 R. (antidiuretic hormone R.)	kidney	Z11687
57.	rat V1a arginine vasopressin R.	liver, kidney, spleen	Z11690
58.	human oxytocin R.	myometrium, mammary gland, non- pregnant endometrium, ovary	X64878
59.	mouse gonadotropin-releasing hormone (GNRH) R.	pituitary gland	M93108
(J) thyrotropin-releasing hormone receptor			
60.	mouse thyrotropin-releasing hormone (TRH) R.	pituitary gland	M94384, M59811

Table 1 (continued)

No.	Gene	Expression	Accession No.
(K) cannabinoid receptor related			
61.	human cannabinoid R.	brain	X54937
62.	human edg-1 (ligand unknown)	endothelial cell	M31210
63.	rat R334 (ligand unknown)	brain, testis	X61496
(L) adenosine receptors			
64.	rat A1 adenosine R.	brain, thyroid [g]	M64299, M69045
65.	human A2 adenosine R.	brain	M97370
66.	human A2b adenosine R.	brain	M97759
67.	rat A3 adenosine R.	testis > kidney, lung, heart > brain	M94152
(M) amine receptors			
(M-1) muscarinic acetylcholine receptors (mAChRs)			
68.	human m1 mAChR	brain [h]	X52068, M35128
69.	human m3 mAChR (hM4)	brain, pancreas	X15266
70.	human m5 mAChR	brain	M80333
71.	human m2 mAChR	brain, heart	M16404, X15264
72.	chicken m2 mAChR	brain, heart	M73217
73.	human m4 mAChR (hM3)	brain	X15265
74.	chicken m4 mAChR	brain, heart	J05218
75.	*Xenopus* m4 mAChR	ovary	X65865
76.	*Drosophila* mAChR	head	M23412
(M-2) histamine H1 receptor			
77.	bovine histamine H1 R.	lung, small intestine > adrenal medulla, uterus > brain, spleen	D10197

Table 1 (continued)

No.	Gene	Expression	Accession No.
(M-3) histamine H2 receptor			
78.	human histamine H2 R.	gastric fundus > brain [i]	M64799
(M-4) β adrenergic receptors			
79.	human β1 adrenergic R.	brain, adipose tissue,heart, lung [j]	J03019
80.	turkey β1 adrenergic R.	erythrocyte	M14379
81.	human β2 adrenergic R.	lung > brain, heart, adipose tissue [k]	J02960
82.	human β3 adrenergic R.	adipocytes, liver, soleus muscle, ileum	M29932
(M-5) α1 adrenergic receptors			
83.	human α1A adrenergic R.	vas deferens, brain, heart, spleen [l]	M76446
84.	rat α1B adrenergic R.	liver	M60655
85.	bovine novel α1 adrenergic R.	brain	J05426
(M-6) α2 adrenergic receptors			
86.	human α2A (2C10) adrenergic R.	platelet	M18415
87.	human α2B (2C4) adrenergic R.	brain, kidney	J03853
88.	human α2C2 adrenergic R.	liver, kidney	M34041
(M-7) tyramine receptor			
89.	*Drosophila* tyramine R. (= octopamine R.)	head	X54794, M60789

Table 1 (continued)

No.	Gene	Expression	Accession No.
	(M-8) serotonine (5HT) receptors (1A, 1B, 1D, 1E, dro, dro2A, dro2B)		
90.	human 5HT1A R. (G21)	brain [m]	X13556
91.	human 5HT1B R.	brain	M75128, M83180, M81590, D10995
92.	human 5HT1D R.	brain	M81589, M89955
93.	human 5HT1E R.	brain	M91467, Z11166
94.	*Drosophila* 5HT-dro R.	head	M55533
95.	*Drosophila* 5HT-dro2A R.	head	Z11489
96.	*Drosophila* 5HT-dro2B R.	head	Z11490
	(M-9) serotonine (5HT) receptors (1C, 2, SRL)		
97.	human 5HT1C R.	brain [n]	M81778
98.	human 5HT2 R.	brain	M86841
99.	rat 5HT R. (SRL)	stomach fundus	X66842
	(M-10) dopamine receptors (D1, D5)		
100.	human D1 dopamine R.	brain	X55760
101.	human D5 dopamine R.	brain	M67439, X58454
	(M-11) dopamine receptors (D2, D3, D4)		
102.	human D2 dopamine R.	brain, retina	M30625, X51362; M29066
103.	*Xenopus* D2 dopamine R.	brain	X59500
104.	rat D3 dopamine R.	brain	X57764
105.	human D4 dopamine R.	brain	X58497

Table 1 (continued)

No.	Gene	Expression	Accession No.
	(N) mas oncogene related receptors		
106.	human mas oncogene (angiotensin R.)	brain [o]	M13150
107.	human mrg (ligand unknown)	not detected	A39485 [a]
108.	rat RTA (ligand unknown)	brain, vas deferens, uterus, intestine, stomach, arota	M35297
	(O) arachidonic acid derivertive receptors		
109.	human thromboxane-A2 (TXA2) R.	placenta, lung	S13647 [a]
110.	mouse prostaglandin E (PGE) R. EP3 subtype	kidney, uterus >brain, thymus, lung, heart, stomach, spleen	D10204
	(P) chemoattractant peptide receptors		
111.	human FPRL1 (formylpeptide R. like)	granulocyte	M84562, X63819, M88107
112.	human N-formylpeptide R.	myeloid cell	M60627
113.	human RMLP-related R.I	leukocyte	M76673
114.	human C5a anaphylatoxin R.	myeloid cell series	M62505
115.	human vasoactiveintestinal peptide (VIP) R.	brain, colon, heart, kidney, lung, spleen, small intestine	M64749
116.	human angiotensin II-1 R.	heart, placenta, lung, liver, skeletal muscle, adrenal	M87290, M93394, M91464, Z11162
117.	human interleukin-8 (IL-8) R. (high affinity)	neutrophil	M68932
118.	human interleukin-8 (IL-8) R. (low affinity)	neutrophil	M73969
119.	human neuropeptide Y3 R.	spleen	M99293
120.	human platelet activating factor (PAF) R.	leukocyte > spleen, lung, kidney [p]	M76674
121.	human thrombin R.	platelet, vascular endothelial cell	M62424

Table 1 (continued)

No.	Gene	Expression	Accession No.
122.	human BK-2 bradykinin R.	kidney, uterus, lung > testis, pancreas,brain, heart	M88714
123.	human somatostatin 1 (SST1) R.	jejunum, stomach,colon, kidney	M81829
124.	human somatostatin 2 (SST2) R.	brain, kidney > jejunum, colon, liver	M81830

Note. - Sequence data are from GenBank release 77.0 and NBRF release 36.0 (data not from the banks are referred). Expression data were taken from the papers listed on the data banks. If not available, the expression data were from the papers of other mammalian receptors. > indicates that tissue expressions of the right are weaker than those of the left.

a data from NBRF release 36.0

b data from Montell et. al. (1987).

c expressions of rat LH-CG receptor (McFarland et al., 1989).

d expressions of pig LH-CG receptor (Loosfelt et al., 1989).

e expressions of rat substance P receptor (Hershey et al., 1991).

f expressions of rat neuromedin B receptor (Wada et al., 1991).

g expressions of dog A1 adenosine receptor (Libert et al., 1989).

h expressions of pig m1 mAChR (Kubo et al., 1986).

i expressions of dog histamine H2 receptor (Gantz et al., 1991).

j expressions of rat β1 adrenergic receptor (Mussin et al., 1991).

k expressions of rat β2 adrenergic receptor (Mussin et al., 1991).

l expressions of rat α1A adrenergic receptor (Lomasney et al., 1991).

m expressions of rat 5HT1A receptor (Albert et al., 1990).

n expressions of rat 5HT1C receptor (Julius et al., 1988).

o expressions of rat mas oncogene (Young et al., 1988).

p expressions of guinea-pig platelet activating factor receptor (Honda et al., 1991).

Table 2

Corrected nucleotide differences at the nonsynonymous and synonymous positions of rhodopsin family receptors that are compared between human and mouse (or rat)

Gene	k_A	k_S	Accession No. Human	Mouse (or Rat)
thromboxane-A2 R.	0.14 ± 0.02	0.65 ± 0.06	a	D10849
interleukin-8 R.(low affinity)	0.13 ± 0.02	0.62 ± 0.06	M94582	L13239
odorant R. I15[b]	0.096 ± 0.014	0.49 ± 0.05	X64994	M64392[c]
thrombin R.	0.094 ± 0.014	0.85 ± 0.09	M62424	L03529
BK-2 bradykinin R.	0.082 ± 0.012	0.56 ± 0.06	M88714	M59967[c]
β3 adrenergic R.	0.052 ± 0.010	0.49 ± 0.05	M29932	M74716[c]
D4 dopamine R.	0.051 ± 0.010	0.35 ± 0.04	X58497	M84009[c]
thyrotropin R.	0.046 ± 0.009	0.57 ± 0.06	M32215	M34842[c]
A2b adenosine R.	0.045 ± 0.009	0.56 ± 0.06	M97759	M91466[c]
histamine H2 R.	0.044 ± 0.009	0.49 ± 0.05	M64799	S57565[c]
LH-CG R.	0.041 ± 0.009	0.72 ± 0.07	M63108	M81310
substance K R.	0.038 ± 0.008	0.58 ± 0.06	M57414	X62933
antidiuretic hormone R.	0.037 ± 0.008	0.60 ± 0.06	Z11687	Z11932[c]
A2 adenosine R.	0.036 ± 0.008	0.55 ± 0.06	M97370	M91214[c]
neuromedin B R.	0.036 ± 0.008	0.62 ± 0.06	M73482	c,d
β2 adrenergic R.	0.036 ± 0.008	0.62 ± 0.06	J02960	X15643
D5 dopamine R.	0.035 ± 0.008	0.65 ± 0.06	X58454	M69118[c]
serotonin 1D R.	0.032 ± 0.008	0.61 ± 0.06	M89955	M89953[c]
rhodopsin	0.029 ± 0.007	0.55 ± 0.06	K02281	M36695-9
α2A adrenergic R.	0.028 ± 0.007	0.51 ± 0.05	M18415	M99377
angiotensin II-1 R.	0.022 ± 0.006	0.84 ± 0.09	M87290	X62295[c]
D1 dopamine R.	0.021 ± 0.006	0.68 ± 0.07	X55760	M35077[c]
follicle-stimulating hormone R.	0.020 ± 0.006	0.56 ± 0.06	M65085	c,e
gastrin-releasing peptide R.	0.019 ± 0.006	0.46 ± 0.05	M73481	M57922
serotonin 1B R.	0.018 ± 0.006	0.39 ± 0.04	D10995	M85151
serotonin 1C R.	0.018 ± 0.006	0.55 ± 0.06	M81778	f
endotheline-B R.	0.017 ± 0.006	0.52 ± 0.05	D90402	X57764[c]
m1 mAChR	0.015 ± 0.005	0.39 ± 0.04	X52068	J04192
somatostatin R. 2	0.015 ± 0.005	0.58 ± 0.06	M81830	M81832
neuromedin K R.	0.015 ± 0.005	0.62 ± 0.06	M89473	J05189[c]
neuropeptide Y R.	0.013 ± 0.005	0.93 ± 0.10	M84755	Z11504[c]
m3 mAChR (hM4)	0.013 ± 0.005	0.53 ± 0.06	X15266	M16407[c]

Table 2 (continued)

Gene	*kA*	*ks*	Accession No. Human	Mouse (or Rat)
β1 adrenergic R.	0.012 ± 0.005	0.24 ± 0.03	J03019	J05561[c]
substance P R.	0.012 ± 0.005	0.39 ± 0.04	M84425	X62934
D2 dopamine R.	0.012 ± 0.005	0.42 ± 0.04	X51362	X55674
α1A adrenergic R.	0.012 ± 0.005	0.62 ± 0.06	M76446	M60654[c]
α2C2 adrenergic R.	0.011 ± 0.005	0.66 ± 0.06	M34041	L00979
angiotensin R. (mas)	0.011 ± 0.005	0.83 ± 0.09	M13150	J03823[c]
m5 mAChR	0.011 ± 0.004	0.39 ± 0.04	M80333	M22926[c]
serotonin 1A R.	0.0091 ± 0.0041	0.76 ± 0.08	X13556	J05276[c]
serotonin 2 R.	0.0085 ± 0.0039	0.52 ± 0.05	M86841	X13971[c]
α2B adrenergic R.	0.0082 ± 0.0039	0.54 ± 0.05	J03853	M99376
m4 mAChR (hM3)	0.0073 ± 0.0036	0.47 ± 0.05	X15265	M16409[c]
cannabinoid R.	0.0054 ± 0.0031	0.45 ± 0.05	X54937	X55812[c]
endotheline-1 R.	0.0033 ± 0.0023	0.93 ± 0.10	D90348	M60786[c]
m2 mAChR	0.0018 ± 0.0018	0.60 ± 0.06	M16404	J03025[c]
somatostatin R. 1	0.0	0.59 ± 0.06	M81829	M81831

Note - Corrected nucleotide differences at the nonsynonymous (*ka*) and synonymous (*ks*) were calculated (Miyata and Yasunaga, 1980) for the homologous regions of the sequences corresponding to the aligned sites of figure 1. The short regions between transmembrane 1 and 2, 2 and 3, 3 and 4, and 6 and 7 were included for the difference calculations. Abbreviations: R. = receptor, LH-CG = luteinizing hormone-choriogonadotropic hormone, mAChR = muscarinic acetylcholine receptor.

a Sequence data from Hirata et al. (1991).
b human HGMP07I.
c rat data.
d Sequence data from Wada et al. (1991).
e Sequence data from Sprengel et al. (1990).
f Sequence data from Yu et al. (1991).

Table 3

Expressions and Accession Numbers of G-protein α subunits

No.	Gene	Expression	Accession No.
(A) Gs			
1.	human αs-1	ubiquitous[a, b]	M21142, X04408
2.	*Xenopus* αs	NI	X56091
3.	rat αolf	olfactory neuro-epithelium[a, b]	c
4.	*Drosophila* αs-S	head > body brain[b]	M33998, M23233
5.	*Lymnaea* αs	NI	Z15096
6.	*Schistosoma* αs (SG12)	NI	M81085
(B) Gi			
7.	human αi1	nearly ubiquitous[a] brain[b]	M17219
8.	*Xenopus* αi1	NI	X56089
9.	human αi3	nearly ubiquitous[a, b]	J03005, M20604, J03198; M27543, J03238
10.	*Xenopus laevis* αi3	NI	X56090
11.	human αi2	ubiquitous[a, b]	M20593, X04828
12.	*Asterina* α (SG)	NI	X66378
13.	*Lymnaea* αi	NI	Z15095
14.	*Drosophila* DGα1	embryo, pupa > adult head	M23094

Table 3 (continued)

No.	Gene	Expression	Accession No.
(C) Go			
15.	human αo2	brain[a, b], others[a]	M60165
16.	*Xenopus laevis* αo	NI	X14636
17.	*Drosophila* αo (DGo2)	head > body	M86660, M29731, M30151, M29602,
18.	*Lymnaea* αo	NI	Z15094
19.	*Caenorhabditis* αo (*goa-1*)	NI	M38251
(D) Gz			
20.	human αz	brain[a, b], adrenal, platelets[a]	J03260
(E) Gt			
21.	human αt1	rod cell[a, b]	X15088
22.	human αt2	cone cell[a, b]	D10384
23.	rat αg (gustducin, gnat-3)	taste bud[a]	X65747
(F) Gq			
24.	mouse αq	nearly ubiquitous[a]	M55412
25.	human αy	nearly ubiquitous[a]	M69013
26.	mouse α14	bone marrow adherent cell lines (lung, kidney, liver)[a]	M80631
27.	*Drosophila dgq*	adult retina, ocellus	M58016

Table 3 (continued)

No.	Gene	Expression	Accession No.
(G) G16			
28.	human α16	T cell, myeloid cell[a]	M63904
29.	mouse α15	B cell, myeloid cell[a]	M80632
(H) G12			
30.	mouse α12	ubiquitous[a]	M63659
31.	mouse α13	ubiquitous[a]	M63660
32.	*Drosophila cta*	oocyte, nurse cell embryo, larva, adult	M63651, M94285
(I) others			
33.	*Dictyostelium* α4	early mound stage >> vegetative growth, early development	A40990 [d]
34.	*Dictyostelium* α1	loose aggregate formation >vegetative cell	M25060
35.	*Dictyostelium* α2	during aggregation >> vegetative cell	M25061
36.	*Caenorhabditis gpa-1*	NI	e
37.	*Caenorhabditis gpa-2*	NI	X53156
38.	*Caenorhabditis gpa-3*	NI	M38250
39.	*Schizosaccharomyces gpa1*	NI	M64286
40.	*Coprinus CGP1*	NI	X68031
41.	*Saccharomyces GPA2*	NI	J03609
42.	*Saccharomyces GPA1*	NI	M15867
43.	*Candida CAG1*	mycelial and yeast cells	M88113
44.	*Arabidopsis GPA1*	vegetative plant tissue	M32887
45.	tomato *TGA1*	NI	M74419

Table 3 (continued)

Note. - Sequence data are from GenBank release 77.0 and NBRF release 36.0 (data not from the banks are referred). Expression data were taken from the papers listed on the data banks and (or) from Kaziro et al. (1991) and Hepler and Gilman (1992). > and >> indicate that tissue expressions of the right are weaker and much weaker than those of the left, respectively.

NI = not identified.

a Expression data from Hepler and Gilman (1992).

b Expression data from Kaziro et al. (1991).

c Sequence data from Jones and Reed (1989).

d Data from NBRF release 36.0.

e Sequence data from Lochrie et al. (1991).

Table 4

Corrected nucleotide differences at the nonsynonymous and synonymous positions of G-protein α subunit that are compared between human and mouse (or rat)

Gene	*kA*	*ks*	Accession No. Human	Mouse (or Rat)
α16[a]	0.065 ± 0.010	0.97 ± 0.10	M63904	M80632
αy[b]	0.019 ± 0.005	0.67 ± 0.07	M69013	M55411
αt1	0.011 ± 0.004	0.81 ± 0.08	X15088	M25506-13
αz	0.0077 ± 0.0034	0.46 ± 0.05	J03260	J03773[c]
αi2	0.0076 ± 0.0034	0.48 ± 0.05	X04828	M13963
αo2	0.0061 ± 0.0031	0.41 ± 0.04	M60156-62	M36778
αi3	0.0060 ± 0.0030	0.45 ± 0.05	M27543	M20713[c]
αs-1	0.0	0.29 ± 0.03	X04408	Y00703
αi1	0.0	0.95 ± 0.10	M17219	M17527[c]

Note - Corrected nucleotide differences at the nonsynonymous (*ka*) and synonymous (*ks*) were calculated (Miyata and Yasunaga, 1980) for the homologous regions of the sequences coresponding to the aligned sites of figure 3.

a mouse α15.

b mouse α11.

c rat data.

Figure Legends

Figure 1. Alignment of the consensus sequences of 35 groups of the rhodopsin family receptors. The gene numbers shown at the left of the sequences correspond to those of Table 1. In the parentheses, the shortest and longest amino acid intervals of the members of each group between the aligned segments and the shortest and longest amino acid lengths of the unaligned N- and C-terminal regions are shown. Above the alignment, the seven transmembrane regions (TM-1 to TM-7) assumed by the hydropathy profiles (Kyte and Doolittle 1982) using the average hydrophilicity score of every aligned site are indicated by bars. - indicates not identical sites. Short region just after transmembrane 4 indicated by a double headed arrow under the alignment was not included for the construction of phylogenetic tree (Fig. 2). Complete alignment of the receptors are on Appendix A.

Figure 2. Phylogenetic tree of the rhodopsin family receptors. The tree was constructed by NJ method (Saitou and Nei, 1987) on the basis of the alignment of Fig. 1. Alignment length for tree construction was 195 amino acids. (a) Schematic unrooted tree of the receptors. The shaded polygons on the tree indicate the clusters of the receptors. Arrows point the separation of vertebrates and insects (or mollusks). (b) Rooted tree of the receptors. The root of the tree does not indicate the oldest time because there is no information on the tree about the oldest gene duplication or oldest separation of the organisms. The tree is essentially unrooted as (a). Filled circles point the separation of vertebrates and insects (or mollusks). The number at each node is a bootstrap probability that the members diverged from the node make a cluster, which has been estimated by a standard procedure with 1000 times resamplings (Felsenstein, 1985). A - P represent the group names of the receptors shown in Table 1. Accession numbers of the sequences are also listed on Table 1. Abbreviations: LHCG = luteinizing hormone-

choriogonadotropic hormone; FSH = follicle-stimulating hormone; TSH = thyroid stimulating hormone; TXA2 = thromboxane-A2; PGE = prostaglandine E; GRP = gastrin-releasing peptide; ET = endotheline; GNRH = gonadotropin-releasing hormone; VIP = vasoactiveintestinal peptide; IL-8 = interleukin-8; PAF = platelet activating factor; SST = somatostatin; GI = glucocorticoid-induced; CCK = cholecystokinin; TRH = thyrotropin-releasing hormone; mACh = muscarinic acetylcholine; adr. = adrenergic; dopa. = dopamine.

Figure 3. Alignment of the consensus sequences of 16 groups of the G-protein α subunits. The gene numbers shown at the left of the sequences correspond to those of Table 3. In the parentheses, the shortest and longest amino acid intervals of the members of each group between the aligned segments and the shortest and longest amino acid lengths of the unaligned N- and C-terminal regions are shown. - indicates not identical site. Abbreviations: *D.d.* = *Dictyostelium discoideum*; *C.e.* = *Caenorhabditis elegans*; *S.p.* = *Schizosaccharomyces pombe*; *C.c.* = *Coprinus congreatus; S.c.* = *Saccharomyces cerevisiae*; *C.a.* = *Candida albicans*. Complete alignment of the G-protein α subunits are on Appendix B.

Figure 4. Phylogenetic tree of the G-protein α subunits. The tree was constructed by NJ method (Saitou and Nei, 1987) on the basis of the alignment of figure 3. Alignment length for tree construction was 277 amino acids. (a) Schematic unrooted tree of the α subunits. The shaded polygons on the tree indicate the clusters of the α subunits. Arrows point the separation of vertebrates and insects (or mollusks). Abbreviations: *D. d.* = *Dictyostelium discoideum*; *C. e.* = *Caenorhabditis elegans*. (b) Rooted tree of the α subunits. The root of the tree does not indicate the oldest time because there is no information on the tree about the oldest gene duplication or oldest separation of the organisms. The tree is essentially unrooted as (a). Filled circles indicate

the separation of vertebrates and insects (or mollusks). The number at each node is a bootstrap probability that the members diverged from the node make a cluster, which has been estimated by a standard procedure with 1000 times resamplings (Felsenstein, 1985). Accession numbers of the sequences are listed on Table 3.

Figure 5. Phylogenetic tree of adenylyl cyclases. The tree was constructed by NJ method (Saitou and Nei, 1987) for the conserved regions (377 amino acid sites). Filled circle indicates the separation of vertebrates and insects. The number at each node is a bootstrap probability that the members diverged from the node make a cluster, which has been estimated by a standard procedure with 1000 times resamplings (Felsenstein, 1985). *Dictyostelium* AC-A was used as the outgroup of the tree (not shown on the tree). Sequence data are from GenBank Release 77.0. Accession numbers of the sequences are as follows: bovine I (M25579); rat II (M80550); rat III (M55075); rat IV (M80633); rat V (M96159); rat VI (M96160); *Drosophila* rutabaga (M81887); *Dictyostelium* AC-A (M87279). Data list and alignment of the adenylyl cyclases are on Appendix C1 and C2, respectively.

Figure 6. Phylogenetic tree of the phosphodiesterases. The tree was constructed by NJ method (Saitou and Nei, 1987) for the conserved regions (168 amino acid sites). Filled circle indicates the separation of vertebrates and insects. The number at each node is a bootstrap probability that the members diverged from the node make a cluster, which has been estimated by a standard procedure with 1000 times resamplings (Felsenstein, 1985). Abbreviations: cGS PDE = cGMP-stimulated phosphodiesterase; cGI PDE = cGMP-inhibited cAMP phosphodiesterase; CaM-PDE. = calmodulin-dependent phosphodiesterase. Sequence data are from GenBank Release 77.0 except for bovine CaM-PDE 61kDa which is from NBRF release 36.0.

Accession numbers of the sequences are as follows: mouse cGMP PDEα (X60664); mouse cGMP PDEβ (X55968); bovine cGMP PDE (M37838); bovine cGS PDE (M73512); rat CaM-PDE (M94537); bovine CaM-PDE 61kDa (NBRF, A40282); human cGI PDE (M91667); rat cAMP PDE1 (M25347); human cAMP PDE2 (M37744); rat cAMP PDE3 (M25349); rat cAMP PDE4 (J04563); *Drosophila dunce* (X55167); *Saccharomyces* cAMP PDE (M14563). Data list and alignment of the phosphodiesterases are on Appendix D1 and D2, respectively.

Figure 7. Phylogenetic tree of phospholipase Cs. The tree was constructed by NJ-method (Saitou and Nei, 1987) for the conserved regions (331 amino acid sites). Filled circle indicates the separation of vertebrates and insects. The number at each node is a bootstrap probability that the members diverged from the node make a cluster, which has been estimated by a standard procedure with 1000 times resamplings (Felsenstein, 1985). Sequence data are from GenBank Release 77.0 except for bovine δ2 (NBRF release 36.0). Accession numbers: human γ1 (M34667); human γ2 (X14034, M37238); rat δ1 (M20637); bovine δ2 (NBRF, S14113); rat β1 (M20636); human β2 (M95678); rat β3 (M99567); *Drosophila plc-21* (M60452), bovine retina β (L13935); *Drosophila norpA* (J03138); *Saccharomyces PLC1* (D12738) *Dictyostelium* DdPLC (M95783). Data list and alignment of the phospholipase Cs are on Appendix E1 and E2, respectively.

Figure 8. Schematic phylogenetic trees of rhodopsin family receptors and G-protein α subunits. Major couplings of the receptors and the G-proteins are indicated by arrows. In the G protein α subunit tree, mammalian q subtype and the insect counterpart, dgq, are denoted as a branch of q. Tissue specific subtypes of G protein α subunits are also represented as the branches of "anisoforms".

Figure 9. Signal pathways from receptors to effectors. Family members of rhodopsin family receptors, G protein α subunits, adenylyl cyclases, phosphodiesterases, phospholipase Cs, cyclic nucleotide gated channels, and calcium channels are surrounded by squares, respectively. Abbreviations: R = receptor; G = G protein; AC = adenylyl cyclase; PDE = phosphodiesterase; PLC = phospholipase C; cAMP = cyclic AMP; cGMP = cyclic GMP; IP_3 = inositol 1,4,5-trisphosphate; DG = diacylglycerol; [+] = activation or increase; [-] = inhibition or decrease.

	Gene No.		TM-1		TM-2
opsin	1- 8	(32- 44)	----------G-P-N------T------L	(1)	--LNYIL-N---------- ----------
opsin	9- 15	(48- 53)	-------M--VV-AS--TN-LV--A--KFKKL	(1)	HPLNWILVNLA--DL-ET--AS TIS--NQ--
opsin	16- 22	(34- 58)	----------GNG-V---F---KSL	(1)	TP-N-----NLA--D---------P--------
odorant R.	23- 35	(22- 35)	------L--Y-------N--I---------L	(1)	-PMY-FL---S---------- ---L------
pituitary glycoprotein hormone R.	36- 38	(359-414)	--LR---W-----LA--GN--VL--L-TS-YK	(1)	-VPRFLMCNL-FAD-C-G-YLL LIASVD--T
tachikinin peptide R.	39- 44	(32-101)	--W----------GN--V-WI-------M	(1)	TVTN-F--NL----------N- --N------
tachikinin peptide R.	45- 46	(71- 90)	------Y------F-L-GN--VC-------RM	(1)	--T--FI--LA--DI-------- P--------
tachikinin peptide R.	47	(40)	FTLALAYGAVIILGVSGNLALIIIILKQKEM	(1)	NVTNILIVNLSFSDLLVAIMCL PLTFVYTLM
gastrointestinal hormone R.	48- 50	(54- 56)	A----LY--IFL-SV-GN-L-I-VL-----R-	(1)	TVTN-FLLSLAVSDL-L---CM PF-L-PNL-
endotheline R.	51- 52	(80-101)	YINTV-SC--F--G--GN-TLLRIIY-NKCM	(1)	NGPN-LIASLALGDL---VID- PINV-KLLA
bombesin-like peptid R.	53- 54	(40- 43)	-VIP--Y--II--GL-GNI-L-KIF-T---M	(1)	-VPN-FIS-LA-GDLLLL-TC- PVDASRY--
neurotensin R.	55	(64)	VLVTAIYLALFVVGTVGNSVTAFTLARKKSL	(4)	STVHYHLGSLALSDLLILLLAM PVELYNFIW
posterior pituitary hormone R.	56- 58	(37- 51)	--E-A-L----------N--VL-AL------	(1-4)	-----F--HL--ADL-VA-FQV LPQL-W---
posterior pituitary hormone R.	59	(35)	KIRVTVTFFLFLLSTAFNASFLLKLQK WTQ	(8)	SRMKVLLKHLTLANLLETLIVM PLDGMWNIT
thyrotropin-releasing hormone R.	60	(25)	VVTILLVVIICGLGIVGNIMVVLVVMRTKHM	(1)	TPTNCYLVSLAVADLMVLVAAG LPNITDSIY
cannabinoid R.and the related genes	61- 63	(44-116)	----------EN--V---I------	(1-2)	-P----IG-LA---DLL------- ----------
adenosine R.	64- 67	(3- 14)	--Y---E--I------GN-LV---V-----L	(1)	--T--F-VSLA-AD-AVG---I P-AI-----
amine R. muscarinic acetylcholine R.	68- 76	(23-104)	----------T--GN---V--S-K----L	(1)	T---NY-L-SLA-AD--IG--SM -L-------
amine R. histamine H1 R.	77	(28)	TPLVVVLSTISLVTVGLNLLVLYAVRSERKL	(1)	TVGNLYIVSLSVADLIVGVVVM PMNILYLLM
amine R. histamine H2 R.	78	(18)	ITITVVLAVLILITVAGNVVVCLAVGLNRRL	(1)	NLTNCFIVSLAITDLLLGLLVL PFSAIYQLS
amine R. β adrenergic R.	79- 82	(33- 58)	--------L-VL--V-GN-LVI-AI----RL	(1)	T-TN-F--SLA-ADLVMGL-VV P--A-----
amine R. α1 adrenergic R.	83- 85	(26- 54)	---G--L---IL----GN-LVILSVAC-RHL	(1)	--T-Y-IVNLA-ADLLL--TVL PFSA---E-L
amine R. α2 adrenergic R.	86- 88	(12- 51)	---------L---T--GN-LV--AV-TSR-L	(1)	APQNLFLVSLA-ADILVATL-- PFSLANE--
amine R. tyramine R.	89	(109)	LLTALVLSVIIVLTIIGNILVILSVFTYKPL	(1)	IVQNFFIVSLAVADLTVALLVL PFNVAYSIL
amine R. serotonine R.	90- 96	(22-226)	-----L----------N---V---------L	(1)	----NYL--SLA--DL-V--LV- P-----Y---
amine R. serotonine R.	97- 99	(53- 74)	-W-AL------II-TI-GNILVI-AVS-EK-L	(1)	-ATNYFLMSLA-AD-L-G---VM P---L-I--
amine R. dopamine R.	100-101	(23- 40)	--TAC-L-LLI--TLLGN-LVCAA--R-RHL	(2)	--TN-F--SLAVSDL-VA-LVM PWKAVAE-A
amine R. dopamine R.	102-105	(28- 34)	---------L-------GN-LVC--V--E-AL	(1)	T-TN----VSLA-ADLL-A-LV- P---VY-EV-
mas oncogene and the related genes	106-108	(32- 75)	----------G---NG---W----- ---	(0-1)	-P---Y---HL--AD---L------ ----------
arachidonic acid derivative R.	109-110	(24- 25)	--S--F------V-G---N-LA----------	(5-6)	-SFL------L-LTD--G-L-T- --V-------
chemoattractant peptide R.	111-114	(26- 37)	I--------V-F--GVLGN-LV-WV--F ----	(1)	T---I--LNLA-ADF-----LP ----------
chemoattractant peptide R.	115-119	(28- 44)	------Y---F-------N---V---V-------	(1)	--------L-L--ADL----T-P -W--------
chemoattractant peptide R.	120-121	(11-102)	---P-VY---FV------N-------VF------	(1-3)	-------M--L--AD-LF---LP --I-YY----
chemoattractant peptide R.	122	(57)	TIQPPFLWVLFVLATLENIFVLSVFCLHKSS	(1)	TVAEIYLGNLAAADLILACGLP FWAITISNN
chemoattractant peptide R.	123-124	(43- 58)	----FIY-VVC--GLCGN--VIYVILR YAK	(2)	T-TNIYILNLAIADEL-ML--P FL------L-
	consensus		N		

Figure 1

Gene No.	TM-3		TM-4		TM-5	
1- 8	(6)	——C--E-F—T--G———WSL——A-ER—V-CK	(10)	HA———-W———PP—	GWSR-	(20)	—E———-F—-F--P——I-F-Y--L——
9- 15	(6)	HP-C--EG—V--CGI--LWSL—ISWERW-VVCK	(10)	-A—GI-F-W-W——W—PPIF	GWSRY	(20)	G--S-M--L—TCC—PL—I--CY--V—AI
16- 22	(6)	——C———G---G————I--DR—VI--	(10-11)	———I———-W———W——		(18)	—R————P——I—-Y—I—-V
23- 35	(6-10)	—-C--Q-———————LL--M-YDR--A-C-	(11)	————————————	——L	(26-27)	————————————SY———-——
36- 38	(13)	G-GC--AGFFTVFASELSVYTLT-ITLERW—IT-	(11)	HA—-M—GW———A--P--	G-S-Y	(13)	-L—-Y———L-LN—AF—-C-CY—IY--V
39- 44	(6)	—--C———————S———-A—-DRY-AI—	(9)	————-IW———P——Y—--		(17-26)	——Y--———L———P————Y—-G—-
45- 46	(6- 7)	-—CH-----Q--S—VSA-TL-AI—DR—-IM-	(9)	-----IA—W--A----LP—I	—-L-	(23)	—-Y--L--F-L———PL-----YAR———
47	(6)	EAMCKLNPFVQCVSITVSIFSLVLIAVERHQLIIN	(9)	HAYVGIAVIWVLAVASSLPFLI	YQVMT	(26)	HRLSYTTLLLVLQYFGPLCFIFICYFKIYIRL
48- 50	(6)	—--CK—--Y-MG-SVSVST--LVAI-LERY-AICR	(11-14)	——————————Y--	———	(17-18)	--Q-W—-LLL-LF--PG-V--VAYGLIS-EL
51- 52	(6-11)	—--CKL-PF-QK-SVGITVL-LCALS-DRYRAVAS	(11)	-TA-EIV-IW--S--LA-PEAI	GF——	(23-25)	Y——KDWWLF-FYFC-PL—TA-FYTLMTCEM
53- 54	(6)	--GCKLIP-IQLTSVGVSVFTLTALSADRY-AIV-	(11)	--C-KA—IW--S-LLA-PEAV	FS——	(22-23)	HPKIHS——FLV——IPL-IIS-YYY-IAK-L
55	(8)	DAGCRGYYFLRDACTYATALNVASLSVERYLAICH	(11)	RTKKFISAIWLASALLAIPMLF	TMGLQ	(21)	TVKVVIQVNTFMSFLFPMLVISILNTVIANKL
56- 58	(5- 6)	D-LCR-VK-LQ——M-AS-Y——-M—DR—A-C-	(10-12)	————————S————————	———	(13-21)	G---Y-TW——-V——P———-C—-I—-I
59	(6)	EFLCKVLSYLKLFSMYAPAFMMVVISLDRSLAITQ	(9)	LEQSMISLAWILSIVFAGPQLY	IFRMI	(24)	HQAFYNFFTFGCLFIIPLLIMLICNAKIIFAL
60	(6)	YVGCLCITYLQYLGINASSCSITAFTIERYIAICH	(11)	RAKKIIIFVWAFTSIYCMLWFF	LLDLN	(19)	YYSPIYLMDFGVFYVVPMILATVLYGFIARIL
61- 63	(2- 5)	————————G———————ASV-SL———-—RY———	(10-11)	————————W———————L——	———	(11-13)	————————————————L————————————
64- 67	(4)	—--CL—--C--L--T--SI--LLA-A-DRY———	(11)	R————————W———————GLTP--	GWN--	(20-30)	—-YMV-F-F-———L-PL—M——Y—-F——
68- 76	(6)	—--CD-WLA-DY—SNASV-NLL-IS--R-F—T-	(11)	-A—MI--AW--S—LW-P-I-	-W——	(16-17)	———-TFGTA-AAFY-PV—M—LY—I-—--
77	(6)	RPLCLFWLSMDYVASTASIFSVFILCIDRYRSVQQ	(11)	RASITILAAWFLSFLWIIPIL	GWRHF	(17)	NVTWFKVMTAIINFYLPTLLMLWFYAKIYKAV
78	(6)	KVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMD	(11)	RVAISLVLIWVISITLSFLSIHLGWNSR		(16)	VNEVYGLVDGLVTFYLPLLIMCITYYRIFKVA
79- 82	(6)	—--CE-WTS-DVLCVTASIETLC--A-DRY-A-T-	(11)	-A------VW--S——SF-PI——W-R-		(18-19)	-N—Y——SS—SFY-PL—M-FVY-RV—-A
83- 85	(6)	—FC—WAAVDVLCCTASI—LC-IS-DRY-GV—	(11)	———L——W———V-S-GPL-	GW--P	(11)	EE—Y—FS---SFY-P—-I-VMYCRVY-VA
86- 88	(6)	--WC--YLALDVLFCTSSIVHLCAISLDRYW———	(11)	R-K--I—-W-I-AVIS-PPL-	———	(12-16)	——WY—-S-IGSFFAPCLIM-LVY-RIY—A
89	(6)	IHLCKLWLTCDVLCCTSSILNLCAIALDRYWAITD	(11)	RVLLLISGVWLLSLLISSPPLI	GWNDW	(13)	SQRGYVIYSSLGSFFIPLAIMTIVYIEIFVAT
90- 96	(6)	—--C-———D--CCT-SIL-L--I—DRY—-T-	(10-11)	————————W———————S--P————————		(12-14)	——Y——T---FY-P————Y--I—-A
97- 99	(7)	—LC—W—LDVLFSTASIMHLCAISLDRY-AI—	(11)	-A—KI—VW-IS-G—-P-P—G—--		(11-14)	——F-L-GS---FF-PLTIM--TY-LTI—L
100-101	(6)	- FC—WVAFDIMCSTASILNLCVISVDRYWAIS-	(11)	-A———-AWTLS-LISFIPVQL-WH—		(23-37)	L-RTYAISSS-ISFYIPVAIMIVTYTRIYRIA
102-105	(6- 7)	—--CD———DVM-CTASI-NLCAIS-DR—AV--	(11-14)	R——I—-W-L———P-L-	G-N—	(7-10)	———-V-YSS--SF--P———L-Y———L
106-108	(6- 7)	————————————————LL-AIS-ERC—V--	(11)	-S—VC-L-W-L————	———	(9-10)	————IL————————-L————
109-110	(10)	-RLC-F-G—M—FGLS-LL——AMA-ER-L-I--	(10-11)	RA———-VW——LA--LLP-L	GVGRY	(14-29)	A———A—-L————————L———VS--
111-114	(6)	—--C————————N——S--L—-I--DR——V--	(11)	LA--———-W-———LT-P————		(18-29)	———-I———GF--P————Y—-I-----
115-119	(4- 6)	—--CK————————N————L——S-DRYL-I—	(10-11)	————————W————P——	———	(16-21)	——————————G——P————Y————
120-121	(6)	--LC———F—N-Y-S———VI——RF-AV--	(11)	R———L-IW————————	—T-	(21-22)	————F——FF——I---C—-IIR-L
122	(6)	ETLCRVVNAIISMNLYSSICFLMLVSIDRYLALVK	(11)	WAKLYSLVIWGCTLLLSSPMLV	FRTMK	(20)	WEVFTNMLLNVVGFLLPLSVITFCTMQIMQVL
123-124	(5)	—--CR-V—VD--N-FTSI-CLTV-S-DRY-AVVH	(11)	-AK———VW—SLLVILPI—	———	(19)	W--GF—YTF—GFL-P——ICLCY--II-K-
consensus		R					

Figure 1 (Continued)

Gene No.	TM-6		TM-7	
1- 8	(15)	A—EV—MV—MV—F———PY——A-----	(9)	——P--F-KS---YNP-IY——NK	(37- 44)
9- 15	(15)	AE-EV-RMVVVM--A——CWGPY—F--F-A-	(9)	-AA-PA-FAKSATIYNP-IYVFMNR	(17- 39)
16- 22	(27- 29)	AE——AK————————W-PY————F	(8- 9)	————————K—A——P—Y—SHP	(39-135)
23- 35	(6)	-———F-TC-SHL--V---Y----------	(7)	----------P--NP—Y-LRN-	(13- 38)
36- 38	(9)	-DT-IAK-MA-LIFTDF-CMAPISF-A-SA—	(9)	-K-LLVLF-P-NSCANPFLYAIFTK	(64- 81)
39- 44	(18- 19)	-K——V———-V—FAICWLP-H—F-----	(12)	-YL---WLAMS--M-NP-IY——N-	(86-136)
45- 46	(19)	-K—--VKM——VV—F——W-P-N——LLL—	(8-11)	--FAFHW-AMS--CYNP-IYC—N-	(73- 76)
47	(18)	ETKRINIMLLSIVVAFAVCWLPLTIFNTVFDW	(12)	LFLLCHLTAMISTCVNPIFYGFLNK	(59)
48- 50	(71- 86)	AKKRV-RML-VIV-LFFLCW-P--S-N-WRA-	(12)	PISFI-LLSY-S-CVNP—YCFM--	(52- 53)
51- 52	(16- 17)	QRREVAKTVFCLV--FALCW-PLHLSRILK-T	(18)	-DYIGIN-A—NSCINPIALY-VSK	(51- 53)
53- 54	(20)	-RKRLAK-VLVFVG-F-FCW-PNH—Y-YRS-	(13)	----AR-L-F-NSCVNPFALYLLS-	(57- 61)
55	(37)	ALRHGVLVLRAVVIAFVVCWLPYHVRRLMFCY	(16)	FYMLTNALFYVSSAINPILYNLVSA	(50)
56- 58	(32- 45)	A——TV-MT--IV-———CW-PFF-VQ-W—W	(9-12)	------LLASLNSC-NPWIY--F--	(36- 55)
59	(22)	ARLRTLKMTVAFATSFVVCWTPYYVLGIWYWF	(11)	VNHFFFLFAFLNPCFDPLIYGYFSL	(0)
60	(41)	SRKQVTKMLAVVVILFALLWMPYRTLVVVNSF	(9)	FLLFCRICIYLNSAINPVIYNLMSQ	(68)
61- 63	(16- 35)	———-T———L———CW-P————	(6-10)	----------NS--NP-IY-----	(15- 70)
64- 67	(18- 22)	-E---AKSL———FAL-WLP——N———F	(7-11)	-----I-L-H-NS--NP—YA——-	(30-119)
68- 76	(143-398)	-E-K---T—AILL-FI-TW-PYN--VL——-	(8-10)	-W——Y-LCY-NST-NP-CYALCN-	(21- 41)
77	(194)	RERKAAKQLGFIMAAFIICWIPYFIFFMVIAF	(8)	VHMFTIWLGYINSTLNPLIYPLCNE	(14)
78	(18)	REHKATVTLAAVMGAFIICWFPYFTAFVYRGL	(9)	LEAIVLWLGYANSALNPILYAALNR	(66)
79- 82	(40- 66)	-E--AL-TLG-IMG-FTLCWLPFF—N-----	(8- 9)	-----NW-GY-NS-FNP-IY CRSP	(52-137)
83- 85	(54- 60)	REKKAAKTL-IVVG-F-LCW-PFF——P-GS-	(9)	VFK—FWLGY-NSC-NP-IYPCSS-	(135-162)
86- 88	(137-165)	-EKRFTFVLAVV-GVFV-CWFPFFF-Y-L——	(7-10)	LF-FFFW-GYCNSSLNPVIYT-FN-	(19- 20)
89	(226-226)	KERRAARTLGIIMGVFVICWLPFFLMYVILPF	(9)	FKNFITWLGYINSGLNPVIYTIFNL	(13)
90- 96	(70-331)	-E-K———L--I---F——WLPFF——L———	(8-10)	——————WLGY-NS—NP--Y———-	(16- 52)
97- 99	(55- 73)	NE--A-KVLGIVF—F--MWCPFFITN———-	(11-12)	LL--FVW-GY—S--NPL-YTLFNK	(85- 95)
100-101	(37- 44)	-ETKVLKTLSVIMGVFVCCWLPFFILNC--PF	(14-18)	TFDVFVWFGWANSSLNP-IYA FNA	(111-114)
102-105	(88-155)	-E-KA———V-G-F--CW-PFF--H———	(8- 9)	L—A-TWLGYVNSA-NP-IYT-FN-	(12- 14)
106-108	(0- 6)	————————————————————L---	(6- 7)	------L——INSSA-P--YF--G-	(40- 58)
109-110	(15- 18)	——E---QL-GIM-V-SVCW-PLL————	(18-19)	--LI--R-A--NQILDPWVY-L-R-	(30- 39)
111-114	(6)	-S——L-V——V-A-FF--W-P————-	(13-16)	-————A--N-C-NP—YV--G-	(44- 46)
115-119	(6- 11)	————————————F——W-P———D--	(17-18)	——T———-C-NP—Y-F---	(40- 52)
120-121	(9- 12)	-K-RAL———V---FIICF-P--V————	(13-16)	A———C—S———DP-IY———	(44- 49)
122	(12)	TERRATVLVLVVLLLFIICWLPFQISTFLDTL	(17)	ITQIASFMAYSNSCLNPLVYVIVGK	(54)
123-124	(14)	SE-K-T-MV—VV-VF—CW-PFY———V-	(7-11)	-----V-L-YANSCANPILY-FLSD	(52- 63)
consensus				P	

Figure 1 (Continued)

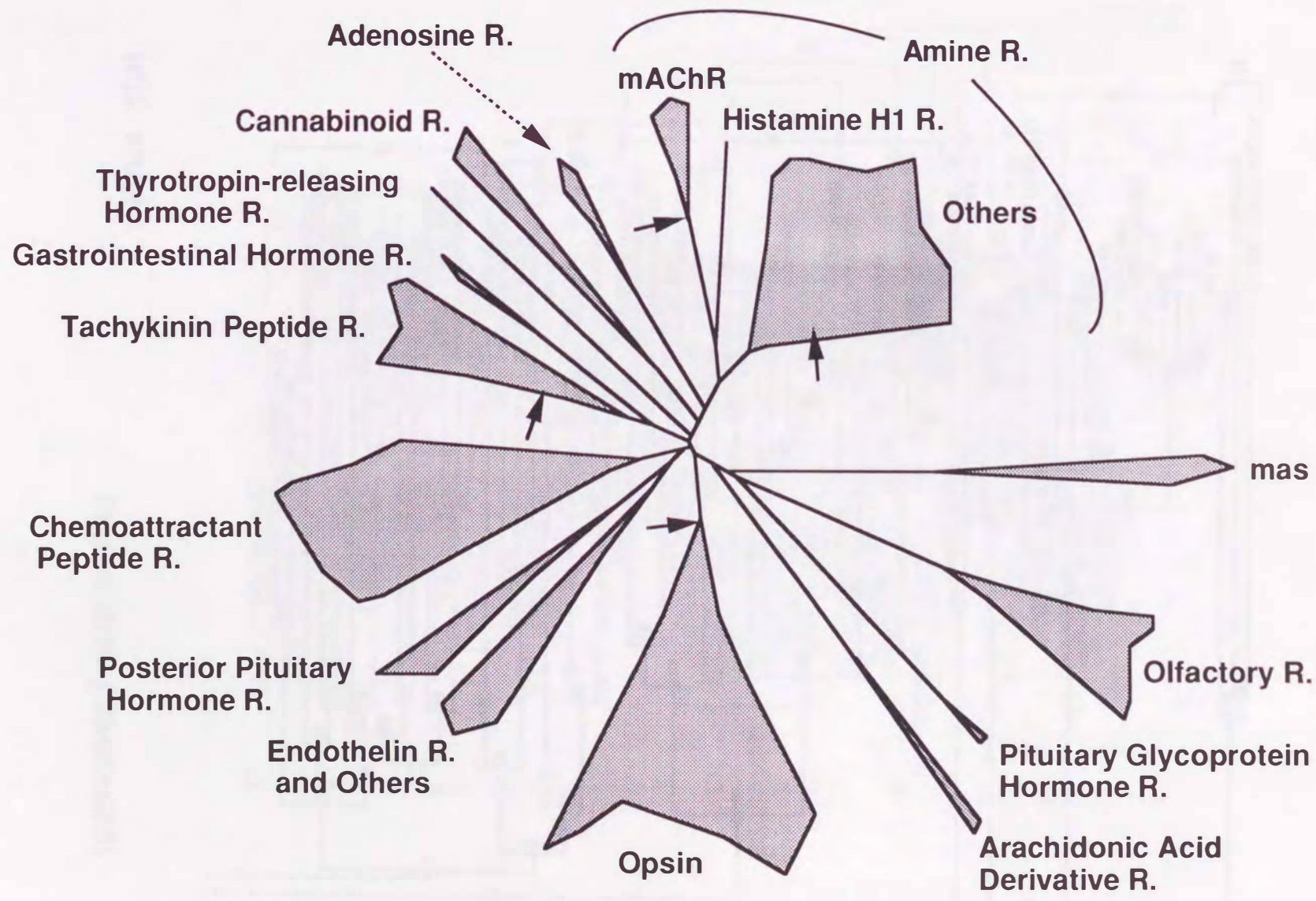

Adenosine R.
Amine R.
mAChR
Cannabinoid R.
Histamine H1 R.
Thyrotropin-releasing Hormone R.
Others
Gastrointestinal Hormone R.
Tachykinin Peptide R.
mas
Chemoattractant Peptide R.
Olfactory R.
Posterior Pituitary Hormone R.
Endothelin R. and Others
Pituitary Glycoprotein Hormone R.
Opsin
Arachidonic Acid Derivative R.

Figure 2(a)

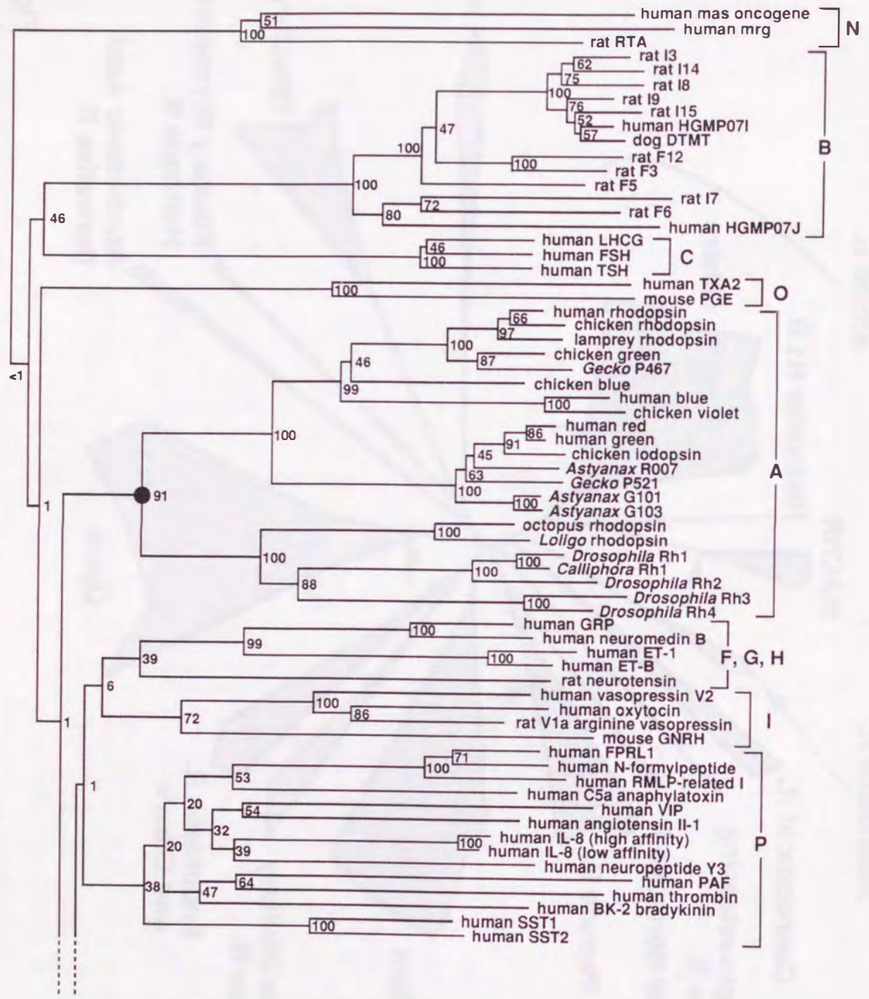


Figure 2(b)

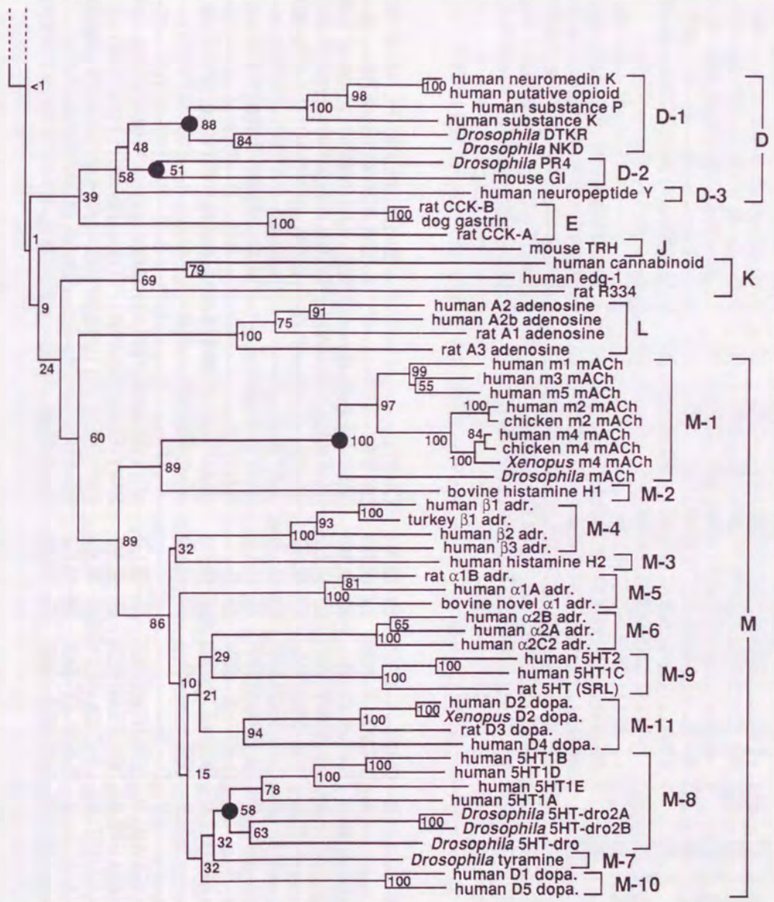


Figure 2(b) (Continued)

```
Gene No.
Gs      1- 6 ( 35- 41) -THRLLLLGAGESGKS-IVKQMRILH- -GF--- (0-15) E---K---I--N---AI--I--AM---
Gi      7-14 ( 22- 32) -EVKLLLLGAGESGKSTIVKQMKIIH- -GYS-- (0   ) -C--Y--VV-SN--QS---I--AMG-L
Go     15-19 ( 31     ) KD-KLLLLGAGESGKSTIVKQMKIIHE -GF--E (0   ) D-KQY-PVVYSNT-QSL-AI-RAM--L
Gz     20    ( 31     ) REIKLLLLGTSNSGKSTIVKQMKIIHS GGFNLE (0   ) ACKEYKPLIIYNAIDSLTRIIRALAAL
Gt     21-23 ( 27- 31) -TVKLLLLGAGESGKSTIVKQMKIIH- -GYS-- (0   ) EC-EF-A--Y-N-LQSILAI--AMTTL
Gq     24-27 ( 31- 37) RELKLLLLGTGESGKSTFIKQMRIIHG -GYS-E (0   ) D--G--KLV-QNIF-AMQ-MI-AM--L
G16    28-29 ( 40     ) -ELKLLLLGPGESGKSTFIKQMRIIHG -GYSEE (0   ) -R--FR-L-YQNIFVSM-AMI-AM-RL
G12    30-32 ( 46-130) R-VK-LLLGAGESGKSTFLKQMRIIHG --FD-- (0   ) ---E----I--N---G--VL-DAR-KL
D.d. α4 33    ( 29     ) KDVKLLLLGPGESGKSTIFKQMKIIQEDGGYSVE (0   ) ELLEYRAFVYSNCISQMEALLTASAKL
D.d. α1,2 34-35 ( 29- 34) -E-KLLLLGAGESGKSTI-KQMKIIH- -G---E (0   ) E----K-II--N----MRVL------L
C.e. gpa 36-38 ( 31     ) ---KLLLLGAGE-GKSTVLKQM----- -----E (0   ) E--------V--N---M------M---
S.p. gpa1 39   ( 72     ) NDIKVLLLGAGDSGKTTIMKQMRLLYS PGFSQV (0   ) VRKQYRVMIFENIISSLCLLLEAMDNS
C.c. CGP1 40   ( 32     ) NEIKMLLLGAGESGKSTVLKQMKLIHH GGYSDQ (0   ) EKDSYKEIIFSNTVQSMRAILDALPAL
S.c. GPA2 41   (121     ) KELKVLLLGAGESGKSTVLQQLKILHQ NGFSEQ (0   ) EIKEYIPLIYQNLLEIGRNLIQARTRF
S.c. GPA1, C.a.CAG1 42-43 ( 39 ) ---KLLLLGAGESGKSTVLKQLKLLH- GGF--Q (0   ) ER-QY--VIW-D-IQSMK-LIIQARKL
Plants 44-45 ( 36- 37) HI-KLLLLGAG-SGKSTIFKQIKLLFQ TGFDE- (0   ) ELK-Y-PVIHANVYQT-K-LHDG-KE-
       consensus             LLLG    GK      Q

Gene No.
 1- 6 (32    ) F------LW- D-G----ERSN  EY QLID   CA-YFL----------Y-P--QD-LRCRVLTSGIFET-F-V DKVN
 7-14 (31- 32) -------LW- D-GVQ--F-RSR  EY QLND   SA-YYL-----R-----YIPT-QDVLRTRVKTTGI-ETHF-- K-L-
15-19 (32    ) LL--M-RLW- D-G-Q-C--R--  EY QLND   SAKY-LD-L-R-G---Y-PTEQ-ILRTRVKTTGIVE-HF-F KNL-
20    (32    ) LLGVMRRLWA DPGAQACFSRSS  EY HLED   NAAYYLNDLERIAAADYIPTVEDILRSRDMTTGIVENKFTF KELT
21-23 (31    ) ----I-RLW- D-G-QACFERA-  EY QLND   SA-YYL--L-R---P-Y-P-EQDVL-SRVKTTGIIET-FS- KDLN
24-27 (30    ) ---AIK-LW- D-GIQECYDRRR  EY QL-D   S-KYYL-D--R-------PT-QD-LR-RVPTTGI-EYPFDL ----
28-29 (30    ) YA-AMQ-LWR DAGIRACYERRR  EF HLLD   SAVYYLSHLERI-E--Y-PTAQDVLRSRMPTTGINEYCFSV -KT-
30-32 (31- 38) Y-P----LW- D-GI--A--RR-  EF Q---   SV-YFLD---------Y-P---DIL--R--TKG--E----- ---P
33    (27    ) LAADIKHLWE DKGIKETYAQKD KHF QLND   SAAYFFDNIDRYMREDFVPNEQDVLRCRVRTTGIQESEFTF DKIR
34-35 (30- 36) ----IKALW- DPG------R--  EF--L-D   SA-Y-FDSIDR---P-Y-P---D-L--R--T-G--ET-FEI ----
36-38 (31- 32) -------LW- D--V-------  -- ----------F-----R-----Y-P---D-L--R--T-G--EV-F-- K---
39    (30    ) IYEAVHALTL DTKLRTVQSCGT  NL SLLD   NFYYYQDHIDRIFDPQYIPSDQDILHCRIKTTGISEETFLL NRHH
40    (30    ) IADAIRQLWA DPGLKEAVRRSR  EF QLND   SAVYYFNSIDRMSAPGYLPTDQDILRSRVKTTGITETTFKV GELT
41    (34    ) IAGVISTLWA LPSTQDLVNGPNASKF YLMD   STPYFMENFTRITSPNYRPTQQDILRSRQMTSGIFDTVIDMGSDIK
42-43 (98-141) IA-AI--LW--D-GIK-CF-RSN  EF QLEG   SA-YYFDN---FA--NY--TD-DILKGRIKTTGITET-F-I -S--
44-45 (38    ) ----IE-LWK DPAIQET--RGN  EL QVPD   C--Y-MENL-R-SD--YIPTKEDVL-AR-RTTGVVEIQFSPVGENK
consensus            L                                                L  R  T G
```

Figure 3

```
Gene No.
 1- 6 (0) FHMFDVGGQR-ERRKWIQCFN DVTAII-V-A-S-YNMV-RED---NR--E-L-L--SIWNNRWLR-IS-ILFLNKQD-L-EK--A
 7-14 (0) FK-FDVGGQRSERKKWIHCFE GVTAIIF-VA-S-YDL-LAED-EMNRM-ES-KLFDSICN-KWF--TSIILFLNKKDLFEEKI
15-19 (0) F-LFDVGGQRSERKKW-HCFE DVTAIIFCVA---YDQ-LHEDETTNRM-ESLKLFDSICNNKWFT-TSIILFLNKKD-F-EKI
20    (0) FKMVDVGGQRSERKKWIHCFE GVTAIIFCVELSGYDLKLYEDNQTSRMAESLRLFDSICNNNWFINTSLILFLNKKDLLAEKI
21-23 (0) FRMFDVGGQRSE-KKWIHCFE GVTCIIF-AAL-AYDMVLVED-EVNRMHESLHLFNSICNH--FA-TSIVLFLNKKD-F-EK-
24-27 (0) FRMVDVGGQRSERRKWIHCFE -VTSI-FLVALSEYDQ-L-E-DNENRMEESKALFRTIITYPWF-NSSVILFLNKKDLLE-KI
28-29 (0) LRIVDVGGQ-SER-KWIHCFE NVIALIYLASLSEYDQCLEEN-QENRM-ESLALF-TILELPWFKSTSVILFLNKTDILE-KI
30-32 (0) F--VDVGGQR--R--W--CFD---TSI-F-VSSSE-DQVL-EDR-TNRL-ES-NIF-TIVNN--F---SIILFLNK-DLL--KV—
33    (0) LKIVDVGGQRSQRRKWIHCFD CVTAVIFVAAMSDYDQVLREDESVNRTRESLALFKEIVNCDYFKETPIVLFLNKKDLFKEKL
34-35 (0) FR-VDVGGQRSERKKW--CF- -VTAV-FCVALSEYDL-LYED--TNRM-ESL--F---CN--WF-NT--ILFLNK-D-F-EKI
36-38 (0) FRV-DVGGQRS-RKKWIHCF- D--A-I--A--SEY--VL-ED-TTNRM-ES--LF-----N---F-NT--ILFLNK-DLF-EKI--
39    (0) YRFFDVGGQRSERRKWIHCFE NVTALLFLVSLAGYDQCLVEDNSGNQMQEALLLWDSICNSSWFSESAMILFLNKLDLFKRKG
40    (0) YKLFDVGGQRSERKKWIHCFE NVTALVFLVSLSEYDQMLYEDESVNRMQEALTLFDSICNSRWFVKTSIILFLNKIDLFAEKL
41    (0) MHIYDVGGQRSERKKWIHCFD NVTLVIFCVSLSEYDQTLMEDKNQNRFQESLVLFDNIVNSRWFARTSVVLFLNKIDLFAEKL
42-43 (0) FKVLDAGGQRS-RKKWIHCFE -ITAVLFVLA-SEYDQ-LFEDERVNRMHESI-LFD-L-NSKWF--TPFILFLNKID-FE-K-
44-45 (5) YRLFDVGGQRNERRKWIHLFE GVTAVIFCAAISEYDQTLFEDE-KNRMMETKELF-WVLKQPCFEKTSFMLFLNKFDIFE-KV
consensus     D GGQ      W   F                      E       E                    LFLNK D    K

Gene No.
 1- 6 GKSK--  -YF-----Y--P (18-21) F-RD-FL- ( 0) ------G-HYCYPH-TCAVDTENI-RVF-DCRDIIQRMHL (6)
 7-14 —SPLT   IC-PEY-G-N—  ( 5    ) YI---FE- ( 0) LN--KD-KEIY-  H-TCATDT-N--FVFDAVTDVIIK-NL (6)
15-19 --SPLT  ICFPEY-G----  ( 5    ) --Q-Q-E- ( 0) -NKS-- KEIY-  H-TCATDT-NIQFVFDAVTDVIIA-NL (6)
20    RRIPLT  ICFPEYKGQNTY  ( 5    ) YIQRQFED ( 0) LNRNKETKEIYS  HFTCATDTSNIQFVFDAVTDVIIQNNL (6)
21-23 -K-HLS  ICFP-Y-G-N--  ( 5    ) YIK-QFL- ( 0) LN-----KEIYS  HMTCATDTQNVKF-FDAVTDIIIKENL (6)
24-27 -YSHL-  -YFPE--GP---  ( 6    ) F-LK----  ( 0) -NPD--  ---YS  HFT-ATDTENI--VF-AVKDTI-Q--L (6)
28-29 -TSHLA  TYFPSFQGP--D  ( 6    ) FILDMY-R (12) -KG-R- RR-F-  H-TCATDTQ--R-VFKDVRDSVLARYL (6)
30-32 ----I-  -----F-G-PH-  ( 6    ) -----F--  ( 0) -RR--------H  HFTTAI-T-NI--VF--VKDTIL--NL (6)
33    KRVPLQ  SCFSDYTGPNKY  ( 5    ) FIQSQYLA  ( 0) QGPSP  RTIYT  HATCAVDTENIKFVFRAVRQTILSQAL (3)
34-35 ------  --F-EY-G---Y  ( 5    ) -IK--F--  ( 0) -N----  K-IY-  H-TCATDTNNI-VVF-AVKDI------ (7)
36-38 K-----  -AF----G----  ( 5- 6) ----K---  ( 0) ---N--  K—Y-  H-TCATDT-QVQ—LD-V---I----L (6)
39    SHFPIQ  KHFPDYQEVGST (20    ) YFYLKFES  ( 0) LNRIAS RSCYC  HFTTATDTSLLQRVMVSVQDTIMSNNL (5)
40    PARRS   TYFPDFTGGDNY  ( 5    ) YLLHRFVS  ( 0) LNQSAATKQIYA  HYTCATDTQQIKFVLSAIQDILLQLHL (6)
41    RKVPME  NYFPDYTGGSDI  ( 5    ) YILWRFVQ  ( 0) LNRAN  LSIYP  HVTQATDTSNIRLVFAAIKETILENTL (7)
42-43 K--P--  -YFPDY-G---D  ( 6    ) -FE--FL-  ( 0) -N-TN  KPIYV  -RTCATD---MKFVLSAVTD-I-QQNL (6)
44-45 --VPLN-CEWF-DYQ-VS-G (13    ) KFEE-Y-Q  ( 0) -TAPDRVDRVFKI YRTTALDQKLVKKTFKLVDETLRRRNL (6)
consensus                                                          T A
```

Figure 3 (Continued)

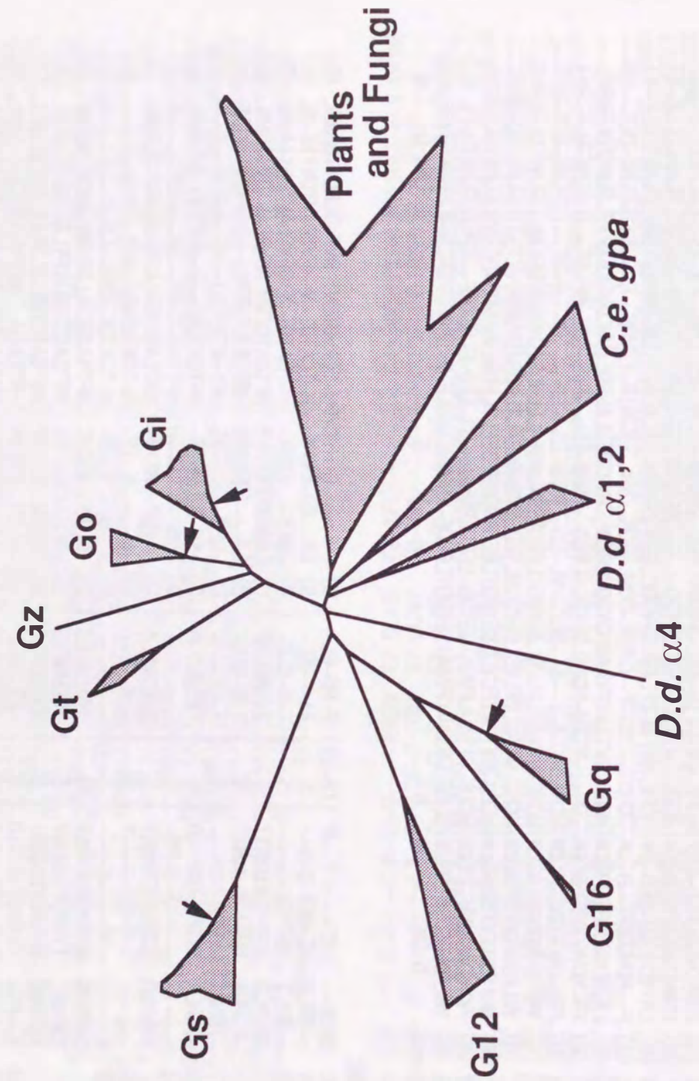


Figure 4(a)

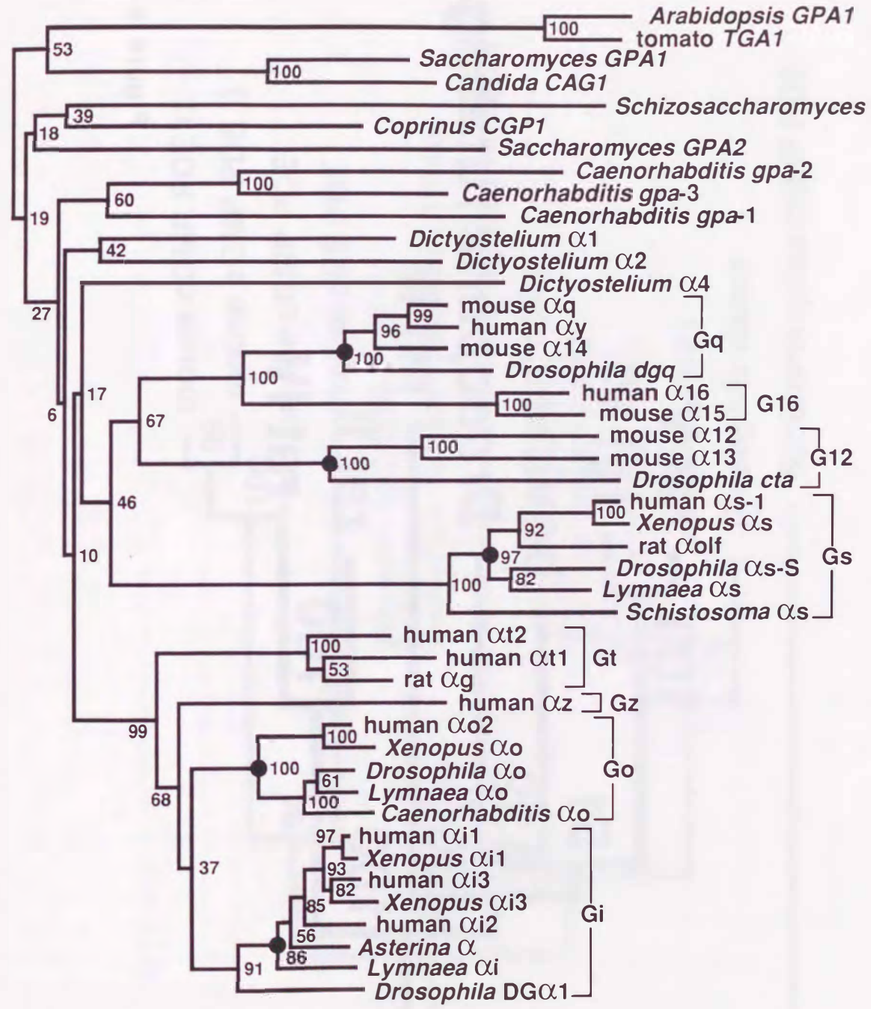


Figure 4(b)

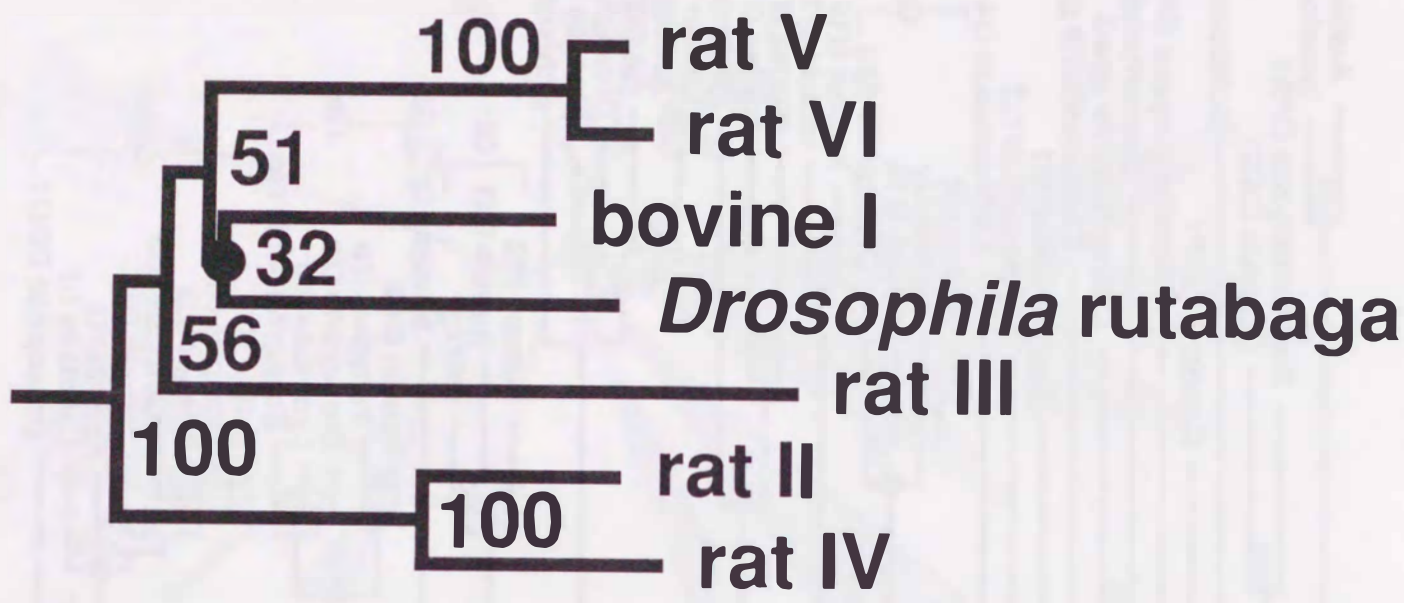


Figure 5

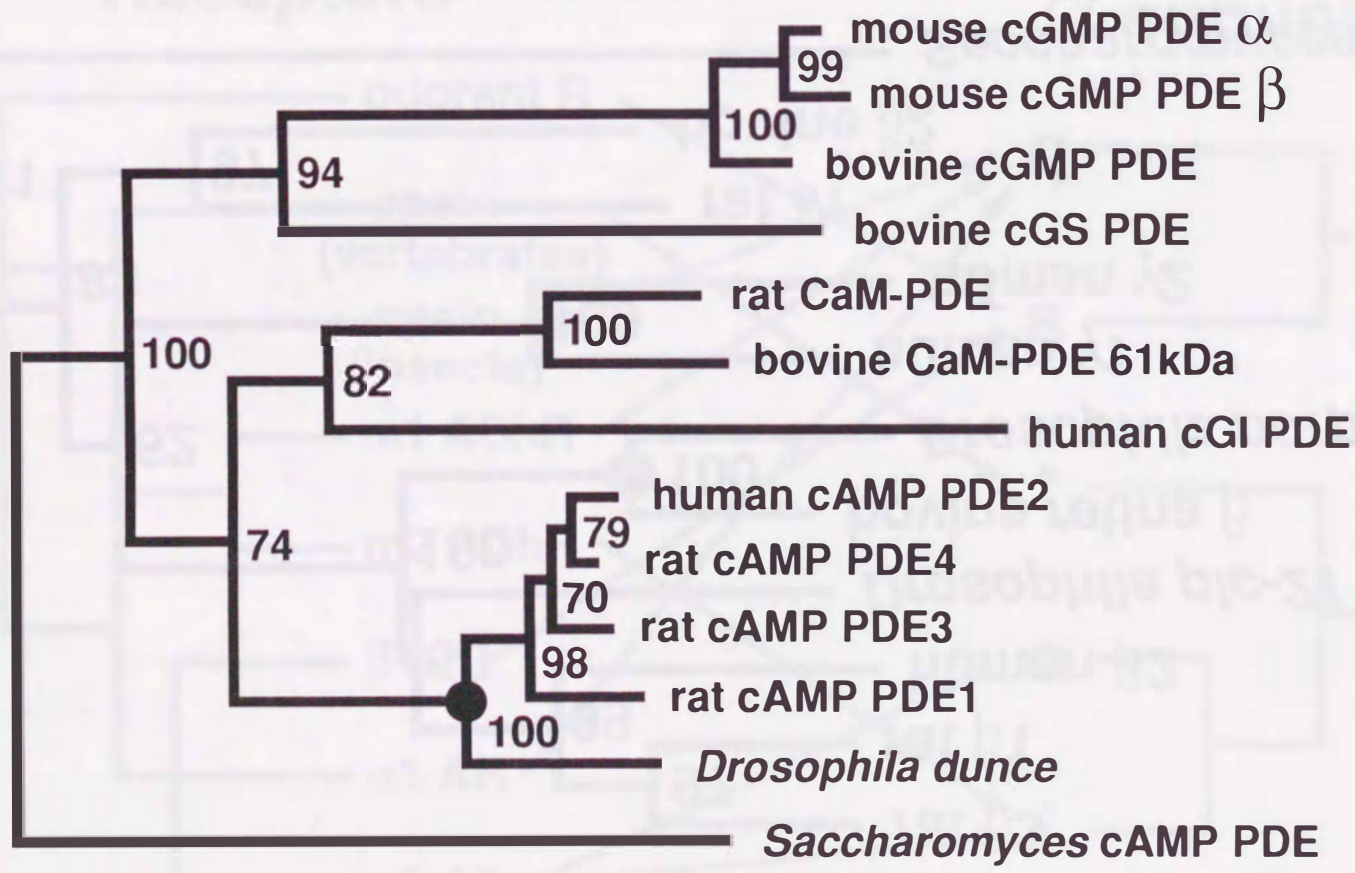


Figure 6

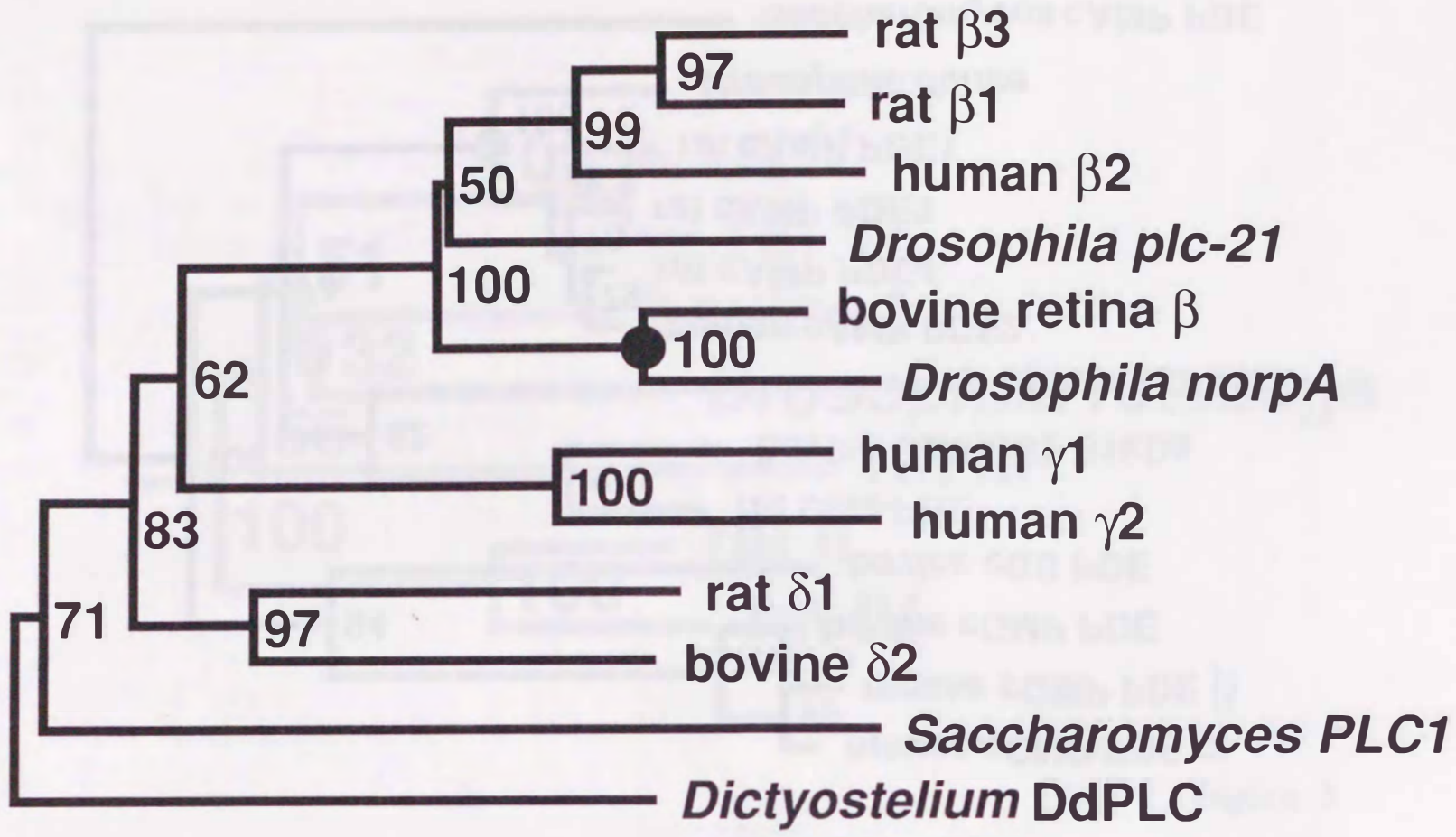


Figure 7

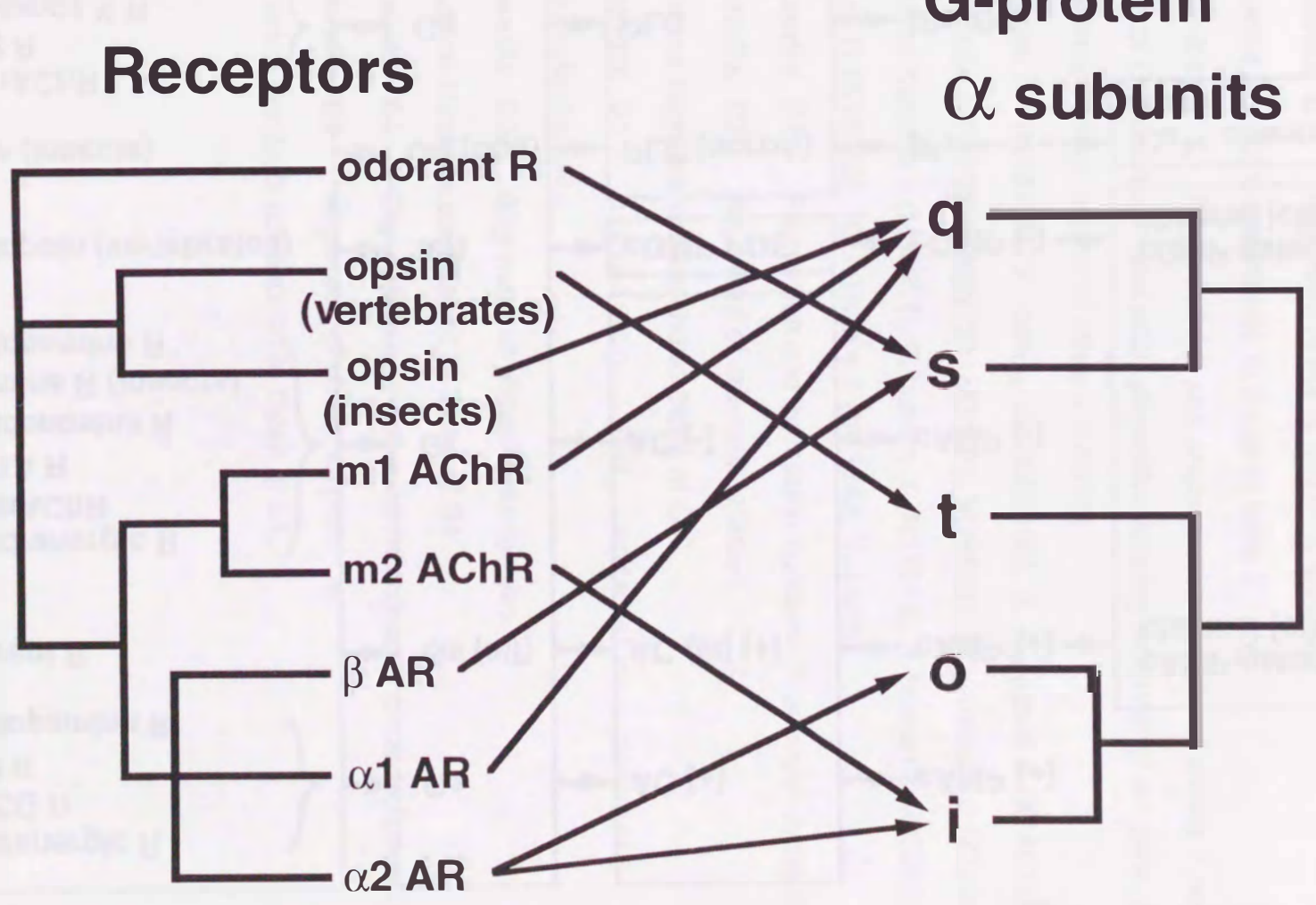


Figure 8

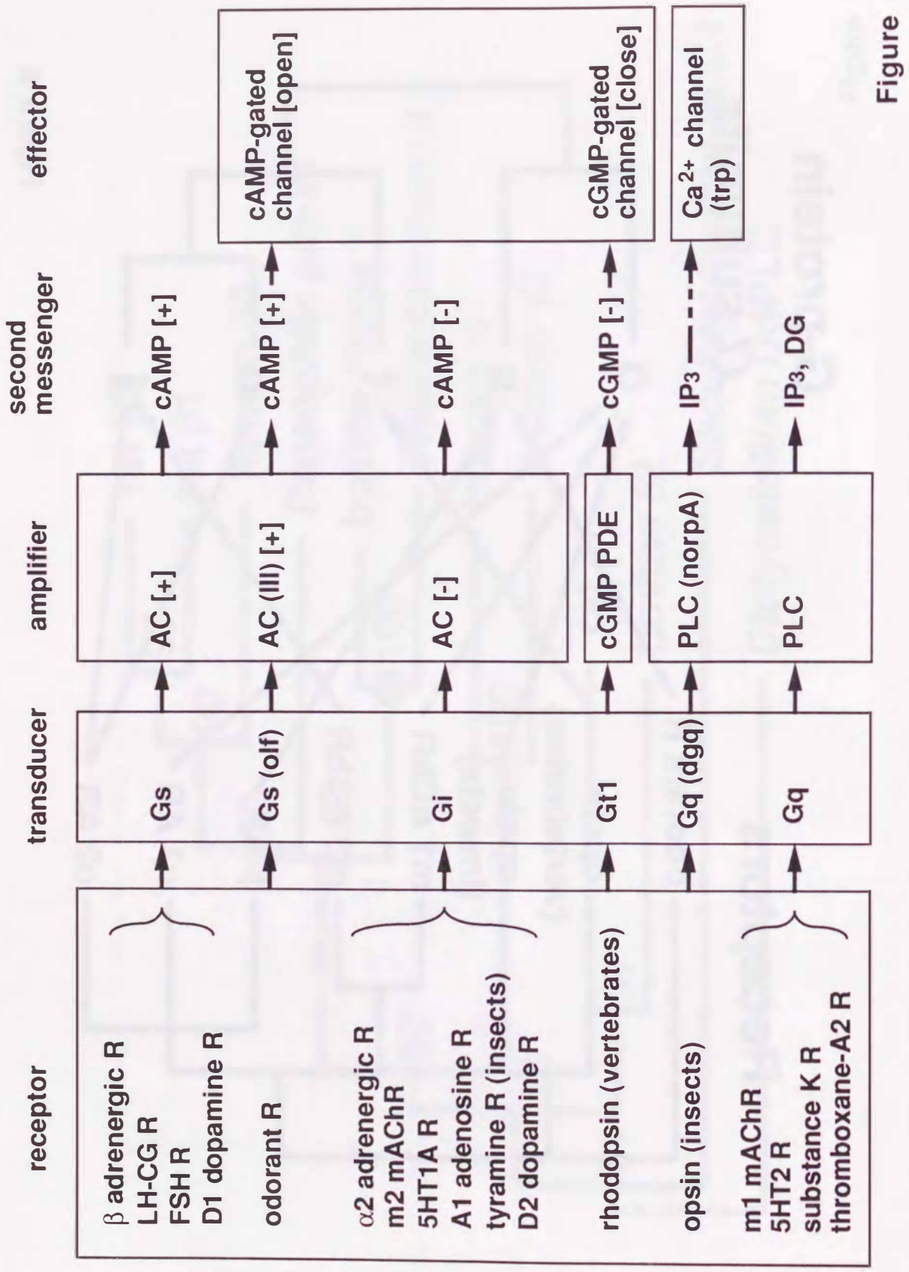


Figure 9

Appendix

Appendix A Alignment of the rhodopsin family receptors.

Sequence numbers correspond to those of table 1.

Appendix B Alignment of the G protein α subunits.

Sequence numbers correspond to those of table 3.

Appendix B' Alignment (a) and phylogenetic tree (b) of exons 7 and 8 region of G protein αo1 and αo2, alternative splicing products of αo gene, and the corresponding regions of other α subunits .

1. human o1 (M60165), 2. human o2 (M60165), 3. *Xenopus* o (X14636), 4. *Drosophila* o (M86660, M29731, M30151, M29602), 5. *Lymnaea* o (Z15094), 6. *Caenorhabditis* o (M38251), 7. human i1 (M17219), 8. human z (J03260), 9. human t1 (X15088), 10. human s-1 (M21142, X04408).

Appendix C1 Expression and accession numbers of adenylyl cyclases.

Appendix C2 Alignment of adenylyl cyclases.

Sequence numbers correspond to those of C1.

Appendix D1 Expression and accession numbers of phosphodiesterases.

Appendix D2 Alignment of phosphodiesterases.

Sequence numbers correspond to those of D1.

Appendix E1 Expression and accession numbers of phospholipase Cs.

Appendix E2 Alignment of phospholipase Cs.

Sequence numbers correspond to those of E1.

rhodopsin family receptor (1)

```
 1. ( 37) SMLAAYMFLLIVLGFPINFLTLYVTVQHKKL (1) TPLNYILLNLAVADLFMVLGGF TSTLYTSLH ( 6) PTGCNLEGFFATLGGEIALWSLVVLAIERYVVVCK (10) HAIMGVAFTWVMALACAAPPLA GWSRY
 2. ( 37) SALAAYMFMLILLGFPVNFLTLYVTIQHKKL (1) TPLNYILLNLVVADLFMVFGGF TTTMYTSMN ( 6) VTGCYIEGFFATLGGEIALWSLVVLAVERYVVVCK (10) HAIMGVAFSWIMAMACAAPPLF GWSRY
 3. ( 37) SALAAYMFFLILVGFPVNFLTLFVTVQHKKL (1) TPLNYILLNLAMANLFMVLFGF TVTMYTSMN ( 6) PTMCSIEGFFATLGGEVALWSLVVLAIERYIVICK (10) HAIMGVAFTWIMALACAAPPLV GWSRY
 4. ( 37) RLVCCYIFFLISTGLPINLLTLLVTFKHKKL (1) QPLNYILVNLAVADLFMACFGF TVTFYTAWN ( 6) PVGCAVEGFFATLGGQVALWSLVVLAIERYIVVCK (10) HAMMGIAFTWVMAFSCAAPPLF GWSRY
 5. ( 37) KVLSFYMFFLIAAGMPLNGLTLFVTFQHKKL (1) QPLNYILVNLAAANLVTVCCGF TVTFYASWY ( 6) PIGCAIEGFFATIGGQVALWSLVVLAIERYIVICK (10) HAIMGIAFTWFMALACAGPPLF GWSRF
 6. ( 44) RAMAAFMFLLIALGVPINTLTIFCTARFRKL (1) SHLNYILVNLALANLLVILVGS TTACYSFSQ ( 6) PTACKIEGFAATLGGMVSLWSLAVVAFERFLVICK (10) HAVLGCVATWVLGFVASAPPLF GWSRY
 7. ( 34) YLQAAFMGTVFLIGFPLNAMVLVATLRYKKL (1) QPLNYILVNVSFGGFLLCIFSV FPVFVASCN ( 6) RHVCALEGFLGTVAGLVTGWSLAFLAFERYIVICK (10) HALTVVLATWTIGIGVSIPPFF GWSRF
 8. ( 32) YLQTAFMGIVFAVGTPLNAVVLWVTVRYKRL (1) QPLNYILVNISASGFVSCVLSV FVVFVASAR ( 6) KRVCELEAFVGTHGGLVTGWSLAFLAFERYIVICK (10) HALLVVVATWLIGVGVGLPPFF GWSRY
 9. ( 53) HLTSVWMIFVVTASVFTNGLVLAATMKFKKL (1) HPLNWILVNLAVADLAETVIAS TISIVNQVS ( 6) HPMCVLEGYTVSLCGITGLWSLAIISWERWLVVCK (10) LAIVGIAFSWIWSAVWTAPPIF GWSRY
10. ( 53) HLTSVWMIFVVIASVFTNGLVLAATMKFKKL (1) HPLNWILVNLAVADLAETVIAS TISVVNQVY ( 6) HPMCVLEGYTVSLCGITGLWSLAIISWERWMVVCK (10) LAIVGIAFSWIWAAVWTAPPIF GWSRY
11. ( 50) NLTSLWMIFVVAASVFTNGLVLVATWKFKKL (1) HPLNWILVNLAVADLGETVIAS TISVINQIS ( 6) HPMCVVEGYTVSACGITALWSLAIISWERWFVVCK (10) LAVAGILFSWLWSCAWTAPPIF GWSRY
12. ( 50) NLATCWMFFVVVASTVTNGLVLVASAKFKKL (1) HPLNWILVNLAIADLLETLLAS TISVCNQFF ( 6) HPMCVFEGFTVATCGIAGLWSLTVISWERWVVVCK (10) MATAGIVFTWVWSAVWCAPPIF GWSRY
13. ( 48) NVSSLWMIFVVIASVFTNGLVIVATAKFKKL (1) HPLNWILVNLAIADLGETVLAS TISVINQIF ( 6) HPMCVFEGWTVSVCGITALWSLTIISWERWVVVCK (10) WAAGGIIFSWVWAIIWCTPPIF GWSRY
14. ( 50) NLASLWMIIVVIASIFTNSLVIVATAKFKKL (1) HPLNWILVNLAIADLGETVLAS TISVFNQVF ( 6) HPMCIFEGWTVSVCGITALWSLTIISWERWVVVCK (10) WAAGGIIFAWTWAIIWCTPPIF GWSRY
15. ( 52) NLVSFFMIIVVIASCFTNGLVLVATAKFKKL (1) HPLNWILVNLAFVDLVETLVAS TISVFNQIF ( 6) HPLCVIEGYVVSSCGITGLWSLAIISWERWFVVCK (10) LAIIGIVFSWVWAWGWSAPPIF GWSRY
16. ( 35) YSVGIFIGVVGIIGILGNGVVIYLFSKTKSL (1) TPANMFIINLAMSDLSFSAINGFPLKTISAFM ( 6) KVACQLYGLLGGIFGFMSINTMAMISIDRYNVIGR (11) RAFLMIIFVWMWSIVWSVGPVF NWGAY
17. ( 34) YSLGIFIAICGIIGCVGNGVVIYLFTKTKSL (1) TPANMFIINLAFSDFTFSLVNGFPLMTISCFM ( 6) NAACKVYGLIGGIFGLMSIMTMTMISIDRYNVIGR (11) KAFIMIIFVWIWSTIWAIGPIF GWGAY
18. ( 50) KILTAYMIMIGMISWCGNGVVIYIFATTKSL (1) TPANLLVINLAISDFGIMITNT PMMGINLYF ( 6) PMMCDIYAGLGSAFGCSSIWSMCMISLDRYQVIVK (10) LALGKIAYIWFMSSIWCLAPAF GWSRY
19. ( 48) KFLAAYMVLIATISWCGNGVVIYIFSTTKSL (1) TPANLLVINLAISDFGIMITNT PMMGINLFY ( 6) PLMCDIYGGLGSAFGCSSILSMCMISLDRYNVIVK (10) LAIMKIALIWFMASIWTLAPVF GWSRY
20. ( 57) KILGLFTLAIMIISCCGNGVVVYIFGGTKSL (1) TPANLLVLNLAFSDFCMMASQS PVMIINFYY ( 6) PLWCDIYAGCGSLFGCVSIWSMCMIAFDRYNVIVK (10) TSIMKILFIWMMAVFWTVMPLI GWSAY
21. ( 58) YLLGTLYIFFTLMSMLGNGLVIWVFSAAKSL (1) TPSNILVINLAFCDFMMMVKT  PIFIYNSFH ( 6) HLGCQIFGIIGSYTGIAAGATNAFIAYDRFNVITR (10) AIAM IIFIYMYATPWVVACYTETWGRF
22. ( 54) YMLGVFYIFLFCASTVGNGMVIWIFSTSKSL (1) TPSNMFVLNLAVFDLIMCLKA  PIFIYNSFH ( 6) NTWCQIFASIGSYSGIGAGMTNAAIGYDRYNVITK (10) AVIMNII IWLYCTPWVVLPLTQFWDRF
23. ( 22) HLFYALFLVMYLTTILGNLLIIVLVQLDSQL (1) TPMYLFLSNLSFSDLCFSSVTM PKLLQNMRS ( 6) YGGCLAQTYFFMVFGDMESFLLVAMAYDRYVAICF (11) LCTCLVLLLWMLTTSHAMMHTL LAARL
24. ( 22) QLFFALFLIMYLTTFLGNLLIVVLVQLDSHL (1) TPMYLFLSNLSFSDLCFSSVTM LKLLQNIQS ( 6) YAGCLTQIFFFLLFGYLGNFLLVAMAYDRYVAICF (11) LCTCLLLVFWIMTSSHAMMHTL LAARL
25. ( 24) LLFYALFLAMYLTIILGNLLIIVLVRLDSHL (1) MPMYLFLSNLSFSDLCFSSVTM PKLLQNMQS ( 6) YTGCLTQLYFFMVFGDMESFLLVVMAYDRYVAICF (11) FCASLVLLLWMLTMTHALLHTL LIARL
26. ( 24) HLFYALFLAMYLTTLLGNLIIIILILLDSHL (1) TPMYLFLSNLSFADLCFSSVTM PKLLQNMQS ( 6) YAGCLAQIYFFLFFGDLGNFLLVAMAYDRYVAICF (11) LCVSLVVLSWVLTTFHAMLHTL LMARL
27. ( 24) HVFYALFLSMYLTTVLGNLIIIILIHLDSHL (1) TPMYLFLSNLSFSDLCFSSVTM PKLLQNMQS ( 6) FAGCLTQLYFYLYFADLESFLLVAMAYDRYVAICF (11) LCVSLVVLSWVLTTFHAMLHTL LMARL
28. ( 24) QLLFLLFLIMYLATVLGNLLIILAIGTDSRL (1) TPMYFFLSNLSFVDVCFSSTTV PKVLANHIL ( 6) FSGCLTQLYFLAVFGNMDNFLLAVMSYDRFVAICH (11) LCVLLVVGSWVVANMNCLLHIL LMARL
29. ( 25) FLIFALFLSMYLVTVLGNLLIIMAIITQSHL (1) TPMYFFLANLSFVDICFTSTTI PKMLVNIYT ( 6) YEDCISQMCVFLVFAELGNFLLAVMAYDRYVAXCH (11) LCILLLLLSWVISIFHAFIQSL IVLQL
30. ( 24) PLIYGLFLSMYLVTVIGNISIIVAIISDPCL (1) TPMYFFLSNLSFVDICFISTTV PKMLVNIQT ( 6) YAGCITQIYFFLLFVELDNFLLTIMAYDRYVAICH (11) LCGFLVLVSWIVSVLHALFQSL MMLAL
31. ( 25) VLLFFLSLLXYVLVLTENMLIIIAIRNHPTL (1) KPMYFFLANMSFLEIWYVTVTI PKMLAGFIG (10) FEACMTQLYFFLGLGCTECVLLAVMAYDRYVAICH (11) LCVQMAAGSWAGGFGISMVKVF LISRL
32. ( 27) IGLFLLFLVMYLLTVVGNLAIISLVGAHRCL (1) TPMYFFLCNLSFLEIWFTTACV PKTLATFAP ( 6) LAGCATQMYFVFSLGCTEYFLLAVMAYDRYLAICL (11) LAMRLALGSWLCGFSAITVPAT LIARL
33. ( 24) NLCYALFLAMYLLTTLLGNLLIIVLIRLDSHL (1) TPMYLFLSNLSFSDLCFSSVTI PKLLQNMQN ( 6) YADCLTQMYFFLLFGDLESFLLVAMAYDRYVAICF (11) LCLALVALSWVLTTFHAMLHTL LMARL
34. ( 35) ITLFGVFLALYILTLAGNIIIVTIIRIDLHL (1) TPMYFFLSMLSTSETVYTLVIL PRMLSSLVG ( 6) LAGCATQMFFFVTFGITNCFLLTAMGYDRYVAICN (11) LRIQLVLGACSIGLIVAITQVT SVFRL
35. ( 24) DLFYALFLAMYVTTILGNLLIIVLIQLDSHL (1) TPMYLFLSNLSFSDLCFSSVTM PKLLQNMQS ( 6) YAGCLTQMYFFLFFGDLESFLLVAMAYDRYVAICF (11) LCFSLLVLSWVLTMFHAVLHTL LMARL
36. (359) DFLRVLIWLINILAIMGNMTVLFVLLTSRYK (1) TVPRFLMCNLSFADFCMGLYLL LIASVDSQT (13) GSGCSTAGFFTVFASELSVYTLTVITLERWHTITY (11) HAILIMLGGWLFSSLIAMLPLV GVSNY
37. (362) NILRVLIWFISILAITGNIIVLVILTTSQYK (1) TVPRFLMCNLAFADLCIGIYLL LIASVDIHT (13) GAGCDAAGFFTVFASELSVYTLTAITLERWHTITH (11) HAASVMVMGWIFAFAAALFPIF GISSY
38. (414) KFLRIVVWFVSLLALLGNVFVLLILLTSHYK (1) NVPRFLMCNLAFADFCMGMYLL LIASVDLYT (13) GPGCNTAGFFTVFASELSVYTLTVITLERWYAITF (11) HACAIMVGGWVCCFLLALLPLV GISSY
39. ( 32) VLWAAAYTVIVVTSVVGNVVVMWIILAHKRM (1) TVTNYFLVNLAFAEASMAAFNT VVNFTYAVH ( 6) LFYCKFHNFFPIAAVFASIYSMTAVAFDRYMAIIH ( 9) ATKVVICVIWVLALLLAFPQGY YSTTE
40. ( 33) ALWAPAYLALVLVAVTGNAIVIWIILAHRRM (1) TVTNYFIVNLALADLCMAAFNA AFNFVYASH ( 6) RAFCYFQNLFPITAMFVSIYSMTAIAADRYMAIVH ( 9) STKAVIAGIWLVALALASPQCF YSTVT
41. ( 85) ALWSLAYGVVVAVAVLGNLIVIWIILAHKRM (1) TVTNYFLVNLAFSDASMAAFNT LVNFIYALH ( 6) ANYCRFQNFFPITAVFASIYSMTAIAVDRYMAIID ( 9) ATKIVIGSIWILAFLLAFPQCL YSKTK
42. ( 60) ALWSLAYGAVVAVAVLGNLVVIWIVLAHKRM (1) TVTNSFLVNLAFADAAMAALNA LVNFIYALH ( 6) ANYCRFQNFFPITAVFASIYSMTAIAVDRYMAIID ( 9) ATRIVIGSIWILAFLLAFPQCL YSKIK
43. (101) VLWSILFGGMVIVATGGNLIVVWIVMTTKRM (1) TVTNYFIVNLSIADAMVSSLNV TFNYYYMLD ( 6) EFYCKLSQFIAMLSICASVFTLMAISIDRYVAIIR ( 9) CNLAIAAVIWLASTLISCPMMIIYRTEE
44. ( 83) TIWAIIFGLMMFVAIAGNGIVLWIVTGHRSM (1) TVTNYFLLNLSIADLLMSSLNC VFNFIFMLN ( 6) SIYCTINNFVANVTVSTSVFTLVAISFDRYIAIVD ( 9) KVRIILVLIWALSCVLSAPCLL YSSIM
45. ( 71) ALLIVAYSFTIVFSLFGNVLVCHVIFKNQRM (1) SATSLFIVNLAVADIMITLLNT PFTLVRFVN ( 6) KGMCHVSRFAQYCSLHVSALTLTAIAVDRHQVIMH ( 9) KGVIYIAVIWVMATFFSLPHAI CQKLF
46. ( 90) IIVYMLYIPIFIFALIGNGTVCYIVYSTPRM (1) TVTNYFIASLAIGDILMSFFCE PSSFISLFI ( 7) LALCHFVNYSQAVSVLVSAYTLVAISIDRYIAIMW ( 9) YATFIIAGVWFIALATALPIPI VSGLD
47. ( 40) FTLALAYGAVIILGVSGNLALIIIILKQKEM (1) NVTNILIVNLSFSDLLVAIMCL PLTFVYTLM ( 6) EAMCKLNPFVQCVSITVSIFSLVLIAVERHQLIIN ( 9) HAYVGIAVIWVLAVASSLPFLI YQVMT
48. ( 56) ALQILLYSIIFLLSVLGNTLVITVLIRNKRM (1) TVTNIFLLSLAVSDLMLCLFCM PFNLIPNLL ( 6) SAVCKTTTYFMGTSVSVSTFNLVAISLERYGAICR (14) KVIAATWCLSFTIMTPYPIYSN LVPFT
49. ( 54) AIRITLYAVIFLMSVGGNVLIIVVLGLSRRL (1) TVTNAFLLSLAVSDLLLAVACM PFTLLPNLM ( 6) TVICKAISYLMGVSVSVSTLNLVAIALERYSAICR (11) HAARVILATWLLSGLLMVPYPV YTMVQ
50. ( 54) AIRVTLYAVIFLMSVGGNVLIIVVLGLSRRL (1) TVTNAFLLSLAVSDLLLAVACM PFTLLPNLM ( 6) TVVCKAVSYLMGVSVSVSTLSLVAIALERYSAICR (11) HAARVIIATWMLSGLLMVPYPV YTAVQ
51. ( 80) YINTVISCTIFIVGMVGNATLLRIIYQNKCM (1) NGPNALIASLALGDLIYVVIDL PINVFKLLA (11) VFLCKLFPFLQKSSVGITVLNLCALSVDRYRAVAS (11) VTAIEIVSIWILSFILAIPEAI GFVMV
52. (101) YINTVVSCLVFVLGIIGNSTLLRIIYKNKCM (1) NGPNILIASLALGDLLHIVIDI PINVYKLLA ( 6) AEMCKLVPFIQKASVGITVLSLCALSIDRYRAVAS (11) WTAVEIVLIWVVSVVLAVPEAI GFDII
53. ( 40) YVIPAVYGVIILIGLIGNITLIKIFCTVKSM (1) NVPNLFISSLALGDLLLLITCA PVDASRYLA ( 6) RIGCKLIPFIQLTSVGVSVFTLTALSADRYKAIVR (11) KICLKAAFIWIISMLLAIPEAV FSDLH
54. ( 43) CVIPSLYLLIITVGLLGNIMLVKIFITNSAM (1) SVPNIFISNLAAGDLLLLLTCV PVDASRYFF ( 6) KVGCKLIPVIQLTSVGVSVFTLTALSADRYRAIVN (11) RTCVKAMGIWVVSVLLAVPEAV FSEVA
55. ( 64) VLVTAIYLALFVVGTVGNSVTAFTLARKKSL (4) STVHYHLGSLALSDLLILLLAM PVELYNFIW ( 8) DAGCRGYYFLRDACTYATALNVASLSVERYLAICH (11) RTKKFISAIWLASALLAIPMLF TMGLQ
56. ( 37) RAELALLSIVFVAVALSNGLVLAALAR RGR (4) APIHVFIGHLCLADLAVALFQV LPQLAWKAT ( 6) DALCRAVKYLQMVGMYASSYMILAMTLDRHRAICR (11) HWNRPVLVAWAFSLLLSLPQLF IFAQR
57. ( 51) KLEIAVLAVIFVVAVLGNSSVLLALHRTPRK (1) SRMHLFIRHLSLADLAVAFFQV LPQLCWTSP ( 5) DWLCRVVKHLQVFAMFASAYMLVVMTADRYIAVCH (10) RSRLMIATSWVLSFILSTPQYF IFSVI
58. ( 39) RVEVAVLCLILLLALSGNACVLLALRTTRQK (1) SRLFFFMKHLSIADLVVAVFQV LPQLLWDIT ( 6) DLLCRLVKYLQVVGMFASTYLLLLMSLDRCLAICQ (12) AVLATWLGCLVASAPQVHIFSL REVAD
59. ( 35) KIRVTVTFFLFLLSTAFNASFLLKLQK WTQ (8) SRMKVLLKHLTLANLLETLIVM PLDGMWNIT ( 6) EFLCKVLSYLKLFSMYAPAFMMVVISLDRSLAITQ ( 9) LEQSMISLAWILSIVFAGPQLY IFRMI
60. ( 25) VVTILLVVIICGLGIVGNIMVVLVVMRTKHM (1) TPTNCYLVSLAVADLMVLVAAG LPNITDSIY ( 6) YVGCLCITYLQYLGINASSCSITAFTIERYIAICH (11) RAKKIIIFVWAFTSIYCMLWFF LLDLN
61. (116) LAIAVLSLTLGTFTVLENLLVLCVILHSRSL (2) RPSYHFIGSLAVADLLGSVIFV YSFIDFHVF ( 5) RNVFLFKLGGVTASFTASVGSLFLTAIDRYISIHR (11) KAVVAFCLMWTIAIVIAVLPLL GWNCE
62. ( 45) KLTSVVFILICCFIILENIFVLLTIWKTKKF (1) RPMYYFIGNLALSDLLAGVAYT ANLLLSGAT ( 5) PAQWFLREGSMFVALSASVFSLLAIAIERYITMLK (10) RLFLLISACWVISLILGGLPIM GWNCI
```

Appendix A

rhodopsin family receptor (2)

```
 63. ( 44) NPWDIVLCSSGTLICCENAVVVLIIFHSPSL (1) APMFLLIGSLALADLLAGLGLI INFVFAYLL ( 2) EATKLVTIGLIVASFSASVCSLLAITVDRYLSLYY (11) LYLCMLVMLWGTSTCLGLLLYG LELPE
 64. (  9) AAYIGIEVLIALVSVPGNVLVIWAVKVNQAL (1) DATFCFIVSLAVADVAVGALVI PLAILINIG ( 4) FHTCLMVACPVLILTQSSILALLAIAVDRYLRVKI (11) RAAVAIAGCWILSLVVGLTPMF GWNNL
 65. (  3) SVYITVELAIAVLAILGNVLVCWAVWLNSNL (1) NVTNYFVVSLAAADIAVGVLAI PFAITISTG ( 4) CHGCLFIACFVLVLTQSSIFSLLAIAIDRYIAIRI (11) RAKGIIAICWVLSFAIGLTPML GWNNC
 66. (  7) ALYVALELVIAALSVAGNVLVCAAVGTANTL (1) TPTNYFLVSLAAADVAVGLFAI PFAITISLG ( 4) FYGCLFLACFVLVLTQSSIFSLLAVAVDRYLAICV (11) RARGVIAVLWVLAFGIGLTPFL GWNSK
 67. ( 14) ITYVTMEAAIGLCAVVGNMLVIWVVKLNRTL (1) TTTFYFIVSLALADIAVGVLVI PLAIAVSLE ( 4) FYACLFMSCVLLVFTHASIMSLLAIAVDRYLRVKL (11) RIWLFLGLCWLVSFLVGLTPMF GWNRK
 68. ( 25) AFIGITTGLLSLATVTGNLLVLISFKVNTEL (1) TVNNYFLLSLACADLIIGTFSM NLYTTYLLM ( 6) TLACDLWLALDYVASNASVMNLLLISFDRYFSVTR (11) RAALMIGLAWLVSFVLWAPAIL FWQYL
 69. ( 68) VFIAFLTGILALVTIIGNILVIVSFKVNKQL (1) TVNNYFLLSLACADLIIGVISM NLFTTYIIM ( 6) NLACDLWLAIDYVASNASVMNLLVISFDRYFSITR (11) RAGVMIGLAWVISFVLWAPAIL FWQYF
 70. ( 30) ITIAAVTAVVSLITIVGNVLVMISFKVNSQL (1) TVNNYYLLSLACADLIIGIFSM NLYTTYILM ( 6) SLACDLWLALDYVASNASVMNLLVISFDRYFSITR (11) RAGIMIGLAWLISFILWAPAIL CWQYL
 71. ( 23) VFIVLVAGSLSLVTIIGNILVMVSIKVNRHL (1) TVNNYFLFSLACADLIIGVFSM NLYTLYTVI ( 6) PVVCDLWLALDYVVSNASVMNLLIISFDRYFCVTK (11) MAGMMIAAAWVLSFILWAPAIL FWQFI
 72. ( 26) VFIVLVAGSLSLVTIIGNILVMVSIKVNRHL (1) TVNNYFLFSLACADLIIGIFSM NLYTLYTVI ( 6) PVVCDLWLALDYVVSNASVMNLLIISFDRYFCVTK (11) MAGMMIAAAWVLSFILWAPAIL FWQFI
 73. ( 32) VFIATVTGSLSLVTVVGNILVMLSIKVNRQL (1) TVNNYFLFSLACADLIIGAFSM NLYTVYIIK ( 6) AVVCDLWLALDYVVSNASVMNLLIISFDRYFCVTK (11) MAGLMIAAAWVLSFVLWAPAIL FWQFV
 74. ( 42) VFIATVTGSLSLVTVVGNILVMLSIKVNRQL (1) TVNNYFLFSLACADLIIGVFSM NLYTVYIIK ( 6) AVVCDLWLALDYVVSNASVMNLLIISFDRYFCVTK (11) MAGLMIAAAWILSFILWAPAIL FWQFI
 75. ( 33) IFIATVTGSLSLVTVVGNILVMLSIKVNRQL (1) TVNNYFLFSLACADLIIGVFSM NLYSLYIIK ( 6) PIVCDLWLALDYVVSNASVMNLLIISLERXFCVTK (11) MAGLMIAAAWLLSFELWAPAIL FWQFI
 76. (104) VVMGFVAAILSTVTVAGNVMVMISFKIDKQL (1) TISNYFLFSLAIADFAIGAISM PLFAVTTIL ( 6) PIVCDTWLALDYLASNASVLNLLIISFDRYFSVTR (11) RAAVMIGAAWGISLLLWPPWIY SWPYI
 77. ( 28) TPLVVVLSTISLVTVGLNLLVLYAVRSERKL (1) TVGNLYIVSLSVADLIVGVVVM PMNILYLLM ( 6) RPLCLFWLSMDYVASTASIFSVFILCIDRYRSVQQ (11) RASITILAAWFLSFLWIIPIL  GWRHF
 78. ( 18) ITITVVLAVLILITVAGNVVVCLAVGLNRRL (1) NLTNCFIVSLAITDLLLGLLVL PFSAIYQLS ( 6) KVFCNIYTSLDVMLCTASILNLFMISLDRYCAVMD (11) RVAISLVLIWVISITLSFLSIHLGWNSR
 79. ( 58) AGMGLLMALIVLLIVAGNVLVIVAIAKTPRL (1) TLTNLFIMSLASADLVMGLLVV PFGATIVVW ( 6) SFFCELWTSVDVLCVTASIETLCVIALDRYLAITS (11) RARGLVCTVWAISALVSFLPILMHWWRA
 80. ( 41) AGMSLLMALVVLLIVAGNVLVIAAIGRTQRL (1) TLTNLFITSLACADLVMGLLVV PFGATLVVR ( 6) SFLCECWTSLDVLCVTASIETLCVIAIDRYLAITS (11) RAKVIICTVWAISALVSFLPIMMHWWRD
 81. ( 33) VGMGIVMSLIVLAIVFGNVLVITAIAKFERL (1) TVTNYFITSLACADLVMGLAVV PFGAAHILM ( 6) NFWCEFWTSIDVLCVTASIETLCVIAVDRYFAITS (11) KARVIILMVWIVSGLTSFLPIQMHWYRA
 82. ( 37) ALAGALLALAVLATVGGNLLVIVAIAWTPRL (1) TMTNVFVTSLAAADLVMGLLVV PPAATLALT ( 6) ATGCELWTSVDVLCVTASIETLCALAVDRYLAVTN (11) CARTAVVLVWVVSAAVSFAPIMSQWWRV
 83. ( 54) VGVGVFLAAFILMAVAGNLLVILSVACNRHL (1) TVTNYFIVNLAVADLLLSATVL PFSATMEVL ( 6) RAFCDVWAAVDVLCCTASILSLCTISVDRYVGVRH (11) KAAAILALLWVVALVVSVGPLL GWKEP
 84. ( 45) ISVGLVLGAFILFAIVGNILVILSVACNRHL (1) TPTNYFIVNLAIADLLLSFTVL PFSATLEVL ( 6) SFFCDIWAAVDVLCCTASILSLCAISIDRYIGVRY (11) KAILALLSVWVLSTVISIGPLL GWKEP
 85. ( 26) ILLGVILGGLILFGVLGNILVILSVACHRHL (1) SVTHYYIVNLAVADLLLTSTVL PFSAIFEIL ( 6) RVFCNVWAAVDVLCCTASIMGLCIISIDRYIGVSY (11) RGLMALLCVWALSLVISIGPLF GWRQP
 86. ( 33) LTLVCLAGLLMLLTVFGNVLVIIAVFTSRAL (1) APQNLFLVSLASADILVATLVI PFSLANEVM ( 6) KTWCEIYLALDVLFCTSSIVHLCAISLDRYWSITQ (11) RIKAIIITCWVISAVISFPPLI SIEKK
 87. ( 51) AGLAAVVGFLIVFTVVGNVLVVIAVLTSRAL (1) APQNLFLVSLASADILVATLVM PFSLANELM ( 6) QVWCGVYLALDVLFCTSSIVHLCAISLDRYWSVTQ (11) RVKATIVAVWLISAVISFPPLV SLYRQ
 88. ( 12) AAIAAAITFLILFTIFGNALVILAVLTSRSL (1) APQNLFLVSLAAADILVATLII PFSLANELL ( 6) RTWCEVYLALDVLFCTSSIVHLCAISLDRYWAVSR (11) RIKCIILTVWLIAAVISLPPLI YKGDQ
 89. (109) LLTALVLSVIIVLTIIGNILVILSVFTYKPL (1) IVQNFFIVSLAVADLTVALLVL PFNVAYSIL ( 6) IHLCKLWLTCDVLCCTSSILNLCAIALDRYWAITD (11) RVLLLISGVWLLSLLISSPPLI GWNDW
 90. ( 36) VITSLLLGTLIFCAVLGNACVVAAIALERSL (1) NVANYLIGSLAVTDLMVSVLVL PMAALYQVL ( 6) QVTCDLFIALDVLCCTSSILHLCAIALDRYWAITD (10) RPRALISLTWLIGFLISIPPIL GWRTP
 91. ( 49) VLLVMLLALITLATTLSNAFVIATVYRTRKL (1) TPANYLIASLAVTDLLVSILVM PISTMYTVT ( 6) QVVCDFWLSSDITCCTASILHLCVIALDRYWAITD (11) RAAVMIALVWVFSISISLPPFF  WRQA
 92. ( 38) ISLAVVLSVITLATVLSNAFVLTTILLTRKL (1) TPANYLIGSLATTDLLVSILVM PISIAYTIT ( 6) QILCDIWLSSDITCCTASILHLCVIALDRYWAITD (11) HAATMIAIVWAISICISIPPLF  WRQA
 93. ( 22) MLICMTLVVITTLTTLLNLAVIMAIGTTKKL (1) QPANYLICSLAVTDLLVAVLVM PLSIIYIVM ( 6) YFLCEVWLSVDMTCCTCSILHLCVIALDRYWAITN (11) RAALMILTVWTISIFISMPPLF  WRSH
 94. (162) IFVSIVLLIVILGTVVGNVLVCIAVCMVRKL (1) RPCNYLLVSLALSDLCVALLVM PMALLYEVL ( 6) PLLCDIWVSFDVLCCTASILNLCAISVDRYLAITK (11) RMMLCVGIVWLAAACISLPPLLILGNEH
 95. (226) AVTSVLLGLMILVTIIGNVFVIAAIILERNL (1) NVANYLVASLAVADLFVACLVM PLGAVYEIS ( 6) PELCDIWTSCDVLCCTASILHLVAIAVDRYWAVTN (10) RVFMMIFCVWTAAVIVSLAPQF GWKDP
 96. (118) IAMAVVLGLMILVTIIGNVFVIAAIILERNL (1) NVANYLVASLAVADLFVACLVM PLGAVYEIS ( 6) PELCDIWTSCDVLCCTASILHLVAIAADRYWTVTN (10) RVFLMIFCVWFAALIVSLAPQF GWKDP
 97. ( 53) NWPALSIVIIIIMTIGGNILVIMAVSMEKKL (1) NATNYFLMSLAIADMLVGLLVM PLSLLAILY ( 7) RYLCPVWISLDVLFSTASIMHLCAISLDRYVAIRN (11) KAIMKIAIVWAISIGVSVPIPVIGLRDE
 98. ( 74) NWSALLTAVVIILTIAGNILVIMAVSLEKKL (1) NATNYFLMSLAIADMLLGFLVM PVSMLTILY ( 7) SKLCAVWIYLDVLFSTASIMHLCAISLDRYVAIQN (11) KAFLKIIAVWTISVGISMPIPVFGLQDD
 99. ( 53) HWAALLIFAVIIPTIGGNILVILAVSLEKRL (1) YATNYFLMSLAVADLLVGLFVM PIALLTIMF ( 7) LALCPAWLFLDVLFSTASIMHLCAISLDRYIAIKK (11) TAFVKITVVWLISIGIAIPVPIKGIEAD
100. ( 23) ILTACFLSLLILSTLLGNTLVCAAVIRFRHL (2) KVTNFFVISLAVSDLLVAVLVM PWKAVAEIA ( 6) S FCNIWVAFDIMCSTASILNLCVISVDRYWAISS (11) AAFILISVAWTLSVLISFIPVQLSWHKA
101. ( 40) VVTACLLTLLIIWTLLGNVLVCAAIVRSRHL (2) NMTNVFIVSLAVSDLFVALLVM PWKAVAEVA ( 6) A FCDVWVAFDIMCSTASILNLCVISVDRYWAISR (11) MALVMVGLAWTLSILISFIPVQLNWHRD
102. ( 34) NYYATLLTLLIAVIVFGNVLVCMAVSREKAL (1) TTTNYLIVSLAVADLLVATLVM PWVVYLEVV ( 6) RIHCDIFVTLDVMMCTASILNLCAISIDRYTAVAM (12) RVTVMISIVWVLSFTISCPLLF GLNNA
103. ( 28) NYYAMLLTLLVFVIVFGNVLVCIAVSREKAL (1) TTTNYLIVSLAVADLLVATLVM PWAVYMEVV ( 6) RIHCDIFVTLDVMMCTASILNLCAISIDRYTAVAM (12) RVTVMISVVWVLSFAISCPLLF GLNNT
104. ( 29) AYYALSYCALILAIIFGNGLVCAAVLRERAL (1) TTTNYLVVSLAVADLLVATLVM PWVVYLEVT ( 7) RICCDVFVTLDVMMCTASILNLCAISIDRYTAVVM (14) RVALMITAVWVLAFAVSCPLLF GFNTT
105. ( 34) AAALVGGVLLIGAVLAGNSLVCVSVATERAL (1) TPTNSFIVSLAAADLLLALLVL PLFVYSEVQ ( 7) PRLCDALMAMDVMLCTASIFNLCAISVDRFVAVAV (11) RQLLLIGATWLLSAAVAAPVLC GLNDV
106. ( 32) PIVHWVIMSISPVGFVENGILLWFLCF RMR (1) NPFTVYITHLSIADISLLFCIF ILSIDYALD ( 7) YYTIVTLSVTFLFGYNTGLYLLTAISVERCLSVLY (11) QSALVCALLWALSCLVTTMEYV MCIDR
107. ( 75) NIIAPKAVLVSLCGVLLNGTVFWLLCC GAT (0) NPYMVYILHLVAADVIYLCCSA VGFLQVTLL ( 6) FFIPDFLAILSPFSFQVCLCLLVAISTERCVCVLF (11) TSNVVCTLIWGLPFCINIVKSL FLTYT
108. ( 44) AVTNYIFLLLCLCGLVGNGLVLWFFGF SIK (1) TPFSIYFLHLASADGIYLFSKA VIALLNMGT ( 6) DYVRRVSRIVGLCTFFAGVSLLPAISIERCVSVIF (11) LSAGVCALLWLLSFLVTSIHNY FCMFL
109. ( 24) IASPWFAASFCVVGLASNLLALSVLAGARQG (5) SSFLTFLCGLVLTDFLGLLVTG TIVVSQHAA (10) CRLCRFMGVVMIFFGLSPLLLGAAMASERYLGITR (11) RAWATVGLVWAAALALGLLPLL GVGRY
110. ( 25) SVSVAFPITMMVTGFVGNALAMLLVSRSYRR (6) KSFLLCIGWLALTDLVGQLLTS PVVILVYLS (10) GRLCTFFGLTMTVFGLSSLLVASAMAVERALAIRA (10) RATPVLLGVWLSVLAFALLPVL GVGRY
111. ( 26) ILPLVVLGVTFVLGVLGNGLVIWVAGF RMT (1) TVTTICYLNLALADFSFTATLP FLIVSMAMG ( 6) WFLCKLIHIVVDINLFGSVFLIGFIALDRCICVLH (11) LAMKVIVGPWILALVLTLPVFL FLTTV
112. ( 26) IITYLVFAVTFVLGVLGNGLVIWVAGF RMT (1) TVTTISYLNLAVADFCFTSTLP FFMVRKAMG ( 6) WFLCKFVFTIVDINLFGSVFLIALIALDRCVCVLH (11) LAKKVIIGPWVMALLLTLPVII RVTTV
113. ( 26) IFSLLVHGVTFVFGVLGNGLVIWVAGF RMT (1) TVNTICYLNLALADFSFSAILP FRMVSVAMR ( 6) SFLCKLVHVMIDINLFVSVYLITIIALDRCICVLH (11) LAKRVMTGLWIFTIVLTLPNFI FWTTI
114. ( 37) ILALVIFAVVFLVGVLGNALVVWVTAF EAK (1) TINAIWFLNLAVADFLSCLALP ILFTSIVQH ( 6) GAACSILPSLILLNMYASILLLATISADRFLLVFK (11) LAWIACAVAWGLALLLTIPSFLYRVVRE
115. ( 44) YTLSFIYIFIFVIGMIANSVVVWVNIQAKTT (1) YDTHCYILNLAIADLWVVLTIP VWVVSLVQH ( 6) ELTCKVTHLIFSINLFSGIFFLTCMSVDRYLSITY (11) VRRVVCILVWLLAFCVSLPDTY YLKTV
116. ( 28) VMIPTLYSIIFVVGIFGNSLVVIVIYFYMKL (1) TVASVFLLNLALADLCFLLTLP LWAVYTAME ( 6) NYLCKIASASVSFNLYASVFLLTCLSIDRYLAIVH (11) VAKVTCIIIWLLAGLASLPAII HRNVF
117. ( 39) YVVIIAYALVFLLSLLGNSLVMLVILYSRVG (1) SVTDVYLLNLALADLLFALTLP IWAASKVNG ( 4) TFLCKVVSLLKEVNFYSGILLLACISVDRYLAIVH (10) LVKFVCLGCWGLSMNLSLPFFL FRQAY
118. ( 43) YFVVIIYALVFLLSLLGNSLVMLVILYSRVG (1) SVTDVYLLNLALADLLFALTLP IWAASKVNG ( 4) TFLCKVVSLLKEVNFYSGILLLACISVDRYLAIVH (10) LVKFICLSIWGLSLLLALPVLL FRRTV
119. ( 38) IFLPTIYSIIFLTGIVGNGLVILVMGYQKKL (1) SMTDKYRLHLSVADLLFVITLP FWAVDAVAN ( 4) NFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVH (11) AEKVVYVGVWIPALLLTIPDFI FANVS
120. ( 11) TLFPIVYSIIFVLGVIANGYVLWVFARLYPC (3) NEIKIFMVNLTMADMLFLITLP LWIVYYQNQ ( 6) KFLCNVAGCLFFINTYCSVAFLGVITYNRFQAVTR (11) RGISLSLVIWVAIVGAASYFLI LDSTN
121. (102) LFVPSVYTGVFVVSLPLNIMAIVVFILKMKV (1) KPAVVYMLHLATADVLFVSVLP FKISYYFSG ( 6) SELCRFVTAAFYCNMYASILLMTVISIDRFLAVVY (11) RASFTCLAIWALAIAGVVPLVL KEQTI
122. ( 57) TIQPPFLWVLFVLATLENIFVLSVFCLHKSS (1) TVAEIYLGNLAAADLILACGLP FWAITISNN ( 6) ETLCRVVNAIISMNLYSSICFLMLVSIDRYLALVK (11) WAKLYSLVIWGCTLLLSSPMLV FRTMK
123. ( 58) ILISFIYSVVCLVGLCGNSMVIYVILR YAK (2) TATNIYILNLAIADELLMLSVP FLVTSTLLR ( 5) ALLCRLVLSVDAVNMFTSIYCLTVLSVDRYVAVVH (11) VAKVVNLGVWVLSLLVILPIVV FSRTA
124. ( 43) AVLTFIYFVVCIIGLCGNTLVIYVILR YAK (2) TITNIYILNLAIADELFMLGLP FLAMQVALV ( 5) KAICRVVMTVDGINQFTSIFCLTVMSIDRYLAVVH (11) TAKMITMAVWGVSLLVILPIMI YAGLR
consensus                    N                                                                        R
```

Appendix A (Continued)

rhodopsin family receptor (3)

```
 1. (20) NNESFVIYMFVVHFTIPMIIIFFCYGQLVFTV ( 15) AEKEVTRMVIIMVIAFLICWVPYASVAFYIFT ( 9) FMTIPAFFAKSAAIYNPVIYIMMNK ( 37)
 2. (20) NNESFVIYMFVVHFMIPLAVIFFCYGNLVCTV ( 15) AEKEVTRMVIIMVIAFLICWVPYASVAFYIFT ( 9) FMTIPAFFAKSSAIYNPVIYIVMNK ( 40)
 3. (20) NNESYVVYMFVVHFLVPFVIIFFCYGRLLCTV ( 15) AEKEVTRMVVLMVIGFLVCWVPYASVAFYIFT ( 9) FMTLPAFFAKSSALYNPVIYILMNK ( 42)
 4. (20) HNESYVLYMFVIHFIIPVVVIFFSYGRLICKV ( 15) AEKEVTRMVILMVLGFMLAWTPYAVVAFWIFT ( 9) LMAVPAFFSKSSSLYNPIIYVLMNK ( 44)
 5. (20) HNESYVIYMFIVHFTVPMVVIFFSYGRLVCKV ( 15) AEKEVTRMVILMVLGFLLAWTPYAATAIWIFT ( 9) FMTIPAFFSKSSSIYNPIIYVLLNK ( 44)
 6. (20) HNESYVLFLFTFCFGVPLAIIVFSYGRLLITL ( 15) ADREVTKMVVVMVLGFLVCWAPYTAFALWVVT ( 9) LASIPSVFSKSSTVYNPVIYVLMNK ( 43)
 7. (20) RSESYTWFLFIFCFIVPLSLICFSYTOLLRAL ( 15) AEREVSRMVVVMVGSFCVCYVPYAAFAMYMVN ( 9) LVTIPSFFSKSACIYNPIIYCFMNK ( 40)
 8. (20) RSEYYTWFLFIFCFIVPLSLIIFSYSOLLSAL ( 15) AEREVSRMVVVMVGSFCLCYVPYAALAMYMVN ( 9) LVTIPAFFSKSACVYNPIIYCFMNK ( 41)
 9. (20) GVOSYMIVLMVTCCIIPLAIIMLCYLOVWLAI ( 15) AEKEVTRMVVVMIFAYCVCWGPYTFFACFAAA ( 9) MAALPAYFAKSATIYNPVIYVFMNR ( 37)
10. (20) GVOSYMIVLMVTCCITPLSIIVLCYLOVWLAI ( 15) AEKEVTRMVVVMVLAFCFCWGPYAFFACFAAA ( 9) MAALPAFFAKSATIYNPVIYVFMNR ( 37)
11. (20) GVOSYMVVLMVTCCFFPLAIIILCYLOVSLAI ( 15) AEKEVSRMVVVMIVAYCFCWGPYTFFACFAAA ( 9) AAALPAYFAKSATIYNPIIYVFMNR ( 38)
12. (20) GVOSYMIVLMITCCFIPLGIIILCYIAVWWAI ( 15) AEKEVSRMVVVMIMAYCFCWGPYTFFACFAAA ( 9) AAAMPAYFAKSATIYNPVIYVFMNR ( 17)
13. (20) GVASYMITLMLTCCILPLSIIIICYIFVWSAI ( 15) AEKEVSRMVVVMILAFIVCWGPYASFATFSAV ( 9) AAAMPAYFAKSATIYNPIIYVFMNR ( 31)
14. (20) GVASYMVTLLLTCCILPLSVIIICYIFVWNAI ( 15) AEKEVSRMVVVMILAFILCWGPYASFATFSAL ( 9) AAALPAYFAKSATIYNPIIYVFMNR ( 31)
15. (20) GCOSFMLTLMITCCFLPLFIIIVCYLOVWMAI ( 15) AEREVSRMVVVMIVAFCICWGPYASFVSFAAA ( 9) AAALPAYFAKSATIYNPVIYVFMNR ( 39)
16. (18) STRSFILCMYFCGFMLPIIIIAFCYFNIVMSV ( 27) AEMKLAKISMVIITQFMLSWSPYAIIALLAQF ( 9) AAELPVLFAKASAIHNPIVYSVSHP (135)
17. (18) TTRSNILCMYIFAFMCPIVVIFFCYFNIVMSV ( 27) AEMKLAKISIVIVTQFLLSWSPYAVVALLAQF ( 9) AAQLPVMFAKASAIHNPMIYSVSHP (132)
18. (18) NPRSYLIFYSIFVYYIPLFLICYSYWFIIAAV ( 28) AEGKLAKVALVTITLWFMAWTPYLVINCMGLF ( 8) NTIWGACFAKSAACYNPIVYGISHP ( 39)
19. (18) NPRSYLIFYSIFVYYLPLFLICYSYWFIIAAV ( 28) AEGKLAKVALVTISLWFMAWTPYTIINTLGLF ( 8) NTIWGACFAKSAACYNPIVYGISHP ( 39)
20. (18) NPRSYLITYSLFVYYTPLFLICYSYWFIIAAV ( 28) AEGKLAKVALTTISLWFMAWTPYLVICYFGLF ( 8) TTIWGATFAKTSAVYNPIVYGISHP ( 40)
21. (18) DTRLFVACIFFFSFVCPTTMITYYYSQIVGHV ( 29) AEIRIAKAAITICFLFFCSWTPYGVMSLIGAF ( 9) ATMIPACACKMVACIDPFVYAISHP ( 40)
22. (18) DTRLFVGTIFFFSFVCPTLMILYYYSQIVGHV ( 29) AEIRIAKAAITICFLFFVSWTPYGVMSLIGAF ( 9) ATMIPACTCKLVACIDPFVYAISHP ( 39)
23. (27) INELMIFIMSTLLIIIPFFLIVMSYARIISSI ( 6) QGICKVFSTCGSHLSVVSLFYGTIIGLYLCPA ( 7) EMVMAMMYTVVTPMLNPFIYSLRNR ( 17)
24. (27) VNELMIHIMGVIIIVIPFVLIVISYAKIISSI ( 6) QSIHKVFSTCGSHLSVVSLFYGTIIGLYLCPS ( 7) GSAMAMMYTVVTPMLNPFIYSLRNR ( 19)
25. (27) VNELMIYILGGLIIIIPFLLIVMSYVRIFFSI ( 6) QDIYKVFSTCGSHLSVVTLFYGTIFGIYLCPS ( 7) EIAMAMMYTVVTPMLNPFIYSLRNR ( 17)
26. (27) DNELAIFILGGPIVVLPFLLIIVSYARIVSSI ( 6) QSIHKAFSTCGSHLSVVSLFYGTVIGLYLCPS ( 7) ETVMSLMYTMVTPMLNPFIYSLRNR ( 19)
27. (27) VNELVIFVMGGLVIVIPFVLIIVSYARVVASI ( 6) RGIHKIFSTCGSHLSVVSLFYGTIIGLYLCPS ( 7) ETVMAMMYTVVTPMLNPFIYSLRNR ( 19)
28. (27) LNELMILTEGAVVMVTPFVCILISYIHITCAV ( 6) RGGWKSFSTCGSHLAVVCLFYGTVIAVYFNPS ( 7) DMAAAAVMYAVVTPMLNPFIYSLRNS ( 18)
29. (27) PSHLIMNLVPVMLAAISFSGILYSYFKIVSSI ( 6) QGKYKAFSTCASHLSIVSLFYSTGLGVYVSSA ( 7) AASASVMYTVVTPMLNPFIYSLRNK ( 21)
30. (27) LNDLVIYFTLVLLATVPLAGIFYSYFKIVSSI ( 6) HGKYKAFSTCASHLSVVSLFYCTGLGVYLSSA ( 7) SATASVMYTVVTPMVNPFIYSLRNK ( 38)
31. (27) TAELTDFVLAIFILLGPLSVTGASYMAITGAV ( 6) AGRHKAFSTCASHLTVVIIFYAASIFIYARPK ( 7) NKLVSVLYAVIVPLFNPIIYCLRNQ ( 27)
32. (27) VVELVSFGIAFCVILGSCGITLVSYAYIITTI ( 6) RGRHRAFSTCSSHLTVVLIWYGSTIFLHVRTS ( 7) TKAITVLNTIVTPVLNPFIYTLRNK ( 13)
33. (27) VNEWVIFIMGGLILVIPFLLILGSYARIVSSI ( 6) KGICKAFSTCGSHLSVVSLFYGTVIGLYLCSS ( 7) DTVMAMMYTVVTPMLNPFIYSLRNR ( 19)
34. (26) VNEILTLIISVLVLVVPMGLVFISYVLIISTI ( 6) EGRKKAFATCASHLTVVIVHYSCASIAYLKPK ( 7) DQLISVTYTVITPLLNPVVYTLRNK ( 15)
35. (26) VNELVIFIMGGLILVIPFLLIITSYARIVSSI ( 6) IGICKVFSTCGSHLSVVSLFYGTVIGLYLCPS ( 7) ETIMAMMYTVVTPMLNPFIYSLRNK ( 20)
36. (13) TLSOVYILTILILNVVAFFIICACYIKIYFAV ( 9) KDTKIAKKMAILIFTDFTCMAPISFFAISAAF ( 9) SKVLLVLFYPINSCANPFLYAIFTK ( 71)
37. (13) PLSQLYVMSLLVLNVLAFVVICGCYIHIYLTV ( 9) SDTRIAKRMAMLIFTDFLCMAPISFFAISASL ( 9) AKILLVLFHPINSCANPFLYAIFTK ( 64)
38. (13) PLALAYIVFVLTLNIVAFVIVCCCYVKIYITV ( 9) KDTKIAKRMAVLIFTDFICMAPISFYALSAIL ( 9) SKILLVLFYPLNSCANPFLYAIFTK ( 81)
39. (19) YEKVYHICVTVLIYFLPLLVIGYAYTVVGITL ( 18) AKRKVVKMMIVVVCTFAICWLPFHIFFLLPYI (12) VYLAIMWLAMSSTMYNPIIYCCLND ( 97)
40. (19) TLLLYHLVVIALIYFLPLAVMFVAYSVIGLTL ( 19) AKKKFVKTMVLVVLTFAICWLPYHLYFILGSF (12) VYLALFWLAMSSTMYNPIIYCCLNH ( 86)
41. (17) QHFTYHIIVIILVYCFPLLIMGITYTIVGITL ( 18) AKRKVVKMMIIVVMTFAICWLPYHIYFILTAI (12) VYLASFWLAMSSTMYNPIIYCCLNK (104)
42. (17) QHFTYHMIVIVLVYCFPLLIMGITYTIVGITL ( 18) AKRKVVKMMIIVVVTFAICWLPYHIYFILTAI (12) VYLASFWLAMSSTMYNPIIYCCLNK (104)
43. (25) MESLYNILIIILTYFLPIVSMTVTYSRVGIEL ( 18) SKRRVVKMMIVVVLIFAICWLPFHSYFIITSC (12) LYLAIYWLAMSNSMYNPIIYCWMNS (133)
44. (26) ADYAYNLIILVLTTGIPMIVMLICYSLMGRVP ( 18) SKRKVVRMFIAIVSIFAICWLPYHLFFIYAYH (12) MYLGFYWLAMSNAMVNPLIYYWMNK (136)
45. (23) FWKYLDLATFILLYLLPLFIISVAYARVAKKL ( 19) KKKTTVKMLVLVVVLFALVWCPLNCYVLLLSS ( 8) LYFAFHWFAMSSTCYNPFIYCWLNE ( 73)
46. (23) QEYYYTLSLFALQFVVPLGVLIFTYARITIRV ( 19) SKRKMVKMMLTVVIVFTCCWLPFNILQLLLND (11) VWFAFHWLAMSHCCYNPIIYCYMNA ( 76)
47. (26) HRLSYTTLLLVLQYFGPLCFIFICYFKIYIRL ( 18) ETKRINIMLLSIVVAFAVCWLPLTIFNTVFDW (12) LFLLCHLTAMISTCVNPIFYGFLNK ( 59)
48. (18) MQQSWQTFLLLILFLLPGIVMVVAYGLISLEL ( 71) AKKRVIRMLIVIVVLFFLCWMPIFSANAWRAY (12) PISFILLLSYTSSCVNPIIYCFMNK ( 53)
49. (17) VQQTWSVLLLLLLLFFIPGVVIAVAYGLISREL ( 86) AKKRVVRMLLVIVLLFFLCWLPVYSVNTWRAF (12) PISFIHLLSYVSACVNPLVYCFMHR ( 52)
50. (18) VRQTWSVLLLLLLFFVPGVVMAVAYGLISREL ( 86) AKKRVVRMLLVIVVLFFLCWLPLYSANTWRAF (12) PISFIHLLSYASACVNPLVYCFMHR ( 53)
51. (23) YQDVKDWWLFGFYFCMPLVCTAIFYTLMTCEM ( 17) QRREVAKTVFCLVVIFALCWFPLHLSRILKKT (18) MDYIGINLATMNSCINPIALYFVSK ( 53)
52. (25) YKTAKDWWLFSFYFCLPLAITAFFYTLMTCEM ( 16) QRREVAKTVFCLVLVFALCWLPLHLSRILKLT (18) LDYIGINMASLNSCINPIALYLVSK ( 51)
53. (23) HPKIHSWASFLVFYVIPLSIISVYYYFIAKNL ( 20) SRKRLAKTVLVFVGLFAFCWLPNHVIYLYRSY (13) TSICARLLAFTNSCVNPFALYLLSK ( 57)
54. (22) HPKIHSVLIFLVYFLIPLAIISIYYYHIAKTL ( 20) TRKRLAKIVLVFVGCFIFCWFPNHILYMYRSF (13) VTLVARVLSFGNSCVNPFALYLLSE ( 61)
55. (21) TVKVVIQVNTRMSFLFPWLVISILNTVIANKL ( 37) ALRHGVLVLRAVVIAFVVCWLPYHVFRVLMFCY (16) FYMLTNALFYVSSAINPILYNLVSA ( 50)
56. (19) GRRTYVTWIALMVFVAPTLGIAACQVLIFREI ( 32) AVAKTVRMTLVIVVVYVLCWAPFFLVQLWAAW ( 9) PFVLLMLLASLNSCTNPWIYASFSS ( 41)
57. (21) GTRAYVTWMTSGVFVAPVVVLGTCYGFICYHI ( 45) AKIRTVKMTFVIVSAYILCWAPFFIVQMWSVW (12) SITITALLASLNSCCNPWIYMFFSG ( 36)
58. (13) GPKAYITWITLAVYIVPVIVLATCYGLISFKI ( 40) AKIRTVKMTFIIVLAFIVCWTPFFFVQMWSVW ( 9) AFIIVMLLASLNSCCNPWIYMLFTG ( 55)
59. (24) HQAFYNFFTFGCLFIIPLLIMLICNAKIIFAL ( 22) ARLRTLKMTVAFATSFVVCWTPYYVLGIWYWF (11) VNHFFFLFAFLNPCFDPLIYGYFSL ( 0)
60. (19) YYSPIYLMDFGVFYVVPMILATVLYGFIARIL ( 41) SRKQVTKMLAVVVILFALLWMPYRTLVVVNSF ( 9) FLLFCRICIYLNSAINPVIYNLMSQ ( 68)
61. (11) HIDETYLMFWIGVTSVLLLFIVYAYMYILWKA ( 35) MDIRLAKTLVLILVVLIICWGPLLAIMVYDVF ( 9) VFAFCSMLCLLNSTVNPIIYALRSK ( 70)
62. (11) LYHKHYILFCTTVFTLLLLSIVILYCRIYSLV ( 20) ENVALLLKTVIIVLSVFIACWAPLFILLLLDVG (10) RAEYFLVLAVLNSGTNPIIYTLTNK ( 66)
```

Appendix A (Continued)

rhodopsin family receptor (4)

```
63. (13) NNAAILSISFLFMFALMLQLYIQICKIVMRHA ( 16) TTRKGISTLALILGTFAACWMPFTLYSLIADY ( 6)  TYATLLPATYNSIINPVIYAFRNQ ( 15)
64. (26) SMEYMVVYFNFFVWVLPPLLLMVLIYLEVFYLI ( 20) KELKIAKSLALILFLFALSWLPLHILNCITLF ( 9) LIYIAIFLTHGNSAMNPIVYAFRIH ( 33)
65. (26) PMNYMVVYFNFFACVLVPLLLMLGVYLRIFLAA ( 22) KEVHAAKSLAIIVGLFALCWLPLHIINCFTFF (10) LMYLAIVLSHTNSVVNPFIYAYRIR (119)
66. (30) PMSYMVVYFNFFGCVLPPLLIMLVIYIKIFLVA ( 18) REIHAAKSLAMIVGIFALCWLPVHAVNCVTLF (11) AMNMAILLSHANSVVNPIVYAYRNR ( 37)
67. (20) GLDYMVFFSFITWILIPLVVMCIIYLDIFYII ( 19) REFKTAKSLFLVLFLFALCWLPLSIINFVSYF ( 7) AMCLGILLSHANSMMNPIVYACKNK ( 30)
68. (16) SQPIITFGTAMAAFYLPVTVMCTLYWRIYRET (143) KEKKAARTLSAILLAFILTWTPYNIMVLVSTF ( 8) LWELGYWLCYVNSTINPMCYALCNK ( 37)
69. (16) SEPTITFGTAIAAFYMPVTIMTILYWRIYKET (226) KEKKAAQTLSAILLAFIITWTPYNIMVLVNTF ( 8) FWNLGYWLCYINSTVNPVCYALCNK ( 41)
70. (16) SEPTITFGTAIAAFYIPVSVMTILYCRIYRET (215) KERKAAQTLSAILLAFIITWTPYNIMVLVSTF ( 8) LWHLGYWLCYVNSTVNPICYALCNR ( 32)
71. (16) SNAAVTFGTAIAAFYLPVIIMTVLYWHISRAS (167) REKKVTRTILAILLAFIITWAPYNVMVLINTF ( 8) VWTIGYWLCYINSTINPACYALCNA ( 21)
72. (16) SNPAVTFGTAIAAFYLPVIIMTVLYWQISRAS (164) REKKVTRTILAILLAFIITWTPYNVMVLINSF ( 8) VWTIGYWLCYINSTINPACYALCNA ( 21)
73. (16) SNPAVTFGTAIAAFYLPVVIMTVLYIHISLAS (171) RERKVTRTIFAILLAFILTWTPYNVMVLVNTF ( 8) VWSIGYWLCYVNSTINPACYALCNA ( 21)
74. (16) SNPAVTFGTAIAAFYLPVVIMTVLYIHISLAS (172) REKKVTRTIFAILLAFILTWTPYNVMVLINTF ( 8) VWSIGYWLCYVNSTINPACYALCNA ( 21)
75. (16) SNPAVTFGTAIAAFYLPVVIMTILYIHISLAS (175) REKKVTRTIFAILLAFIITWTPYNVMVLINTF ( 8) IWYIGYWLCYVNSTINPACYALCNA ( 21)
76. (17) TNQYITFGTALAAFYFPVTIMCFLYWRIWRET (398) QESKAAKTLSAILLSFIITWTPYNILVLIKPL (10) LWDFFYALCYINSTINPMCYALCNA ( 28)
77. (17) NVTWFKVWTAIINFYLPTLLMLWFYAKIYKAV (194) RERKAAKQLGFIMAAFIICWIPYFIFFMVIAF ( 8) VHWFTIWLGYINSTLNPLIYPLCNE ( 14)
78. (16) VNEVYGLVDGLVTFYLPLLIMCITYYRIFKVA ( 18) REHKATVTLAAVMGAFIICWFPYFTAFVYRGL ( 9) LEAIVLWLGYANSALNPILYAALNR ( 66)
79. (18) TNRAYAIASSVVSFYVPLCIMAFVYLRVFREA ( 66) REQKALKTLGIIMGVFTLCWLPFFLANVVKAF ( 8) LFVFFNWLGYANSAFNPIIY CRSP ( 96)
80. (18) TNRAYAIASSIISFYIPLLIMIFVYLRVYREA ( 49) REHKALKTLGIIMGVFTLCWLPFFLVNIVNVF ( 8) LFVFFNWLGYANSAFNPIIY CRSP (137)
81. (18) TNQAYAIASSIVSFYVPLVIMVFVYSRVFQEA ( 40) KEHKALKTLGIIMGTFTLCWLPFFIVNIVHVI ( 8) VYILLNWIGYVNSGFNPLIY CRSP ( 83)
82. (19) SNWPYVLLSSSVSFYLPLLVMLFVYARVFVVA ( 54) REHRALCTLGLIMGTFTLCWLPFFLANVLRAL ( 9) AFLALNWLGYANSAFNPLIY CRSP ( 52)
83. (11) EEAGYAVFSSVCSFYLPMAVIVVMVCRVYVVA ( 60) REKKAAKTLAIVVGVFVLCWFPFFFVLPLGSL ( 9) VFKVIFWLGYFNSCVNPLIYPCSSR (136)
84. (11) EEPFYALFSSLGSFYIPLAVILVMYCRVYIVA ( 57) REKKAAKTLGIVVGMFILCWLPFFIALPLGSL ( 9) VFKVVFWLGYFNSCLNPIIYPCSSK (162)
85. (11) EEPGYVLFSALGSFYVPLTIILVMYCRVYVVA ( 54) REKKAAKTLGIVVGCFVLCWLPFFLVMPIGSF ( 9) VFKIAFWLGYLNSCINPIIYPCSSQ (135)
86. (16) DQKWYVISSCIGSFFAPCLIMILVYVRIYQIA (144) LEKRFTFVLAVVIGVFVVCWFPFFFTYTLTAV ( 7) LFKFFFWFGYCNSSLNPVIYTIFNH ( 19)
87. (12) DETWYILSSCIGSFFAPCLIMGLVYARIYRVA (137) REKRFTFVLAVVMGVFVLCWFPFFFIYSLYGI (10) LFKFFFWIGYCNSSLNPVIYTVFNQ ( 20)
88. (13) QEAWYILASSIGSFFAPCLIMILVYLRIYLIA (165) REKRFTFVLAVVIGVFVLCWFPFFFSYSLGAI (10) LFQFFFWIGYCNSSLNPVIYTIFNQ ( 19)
89. (13) SQRGYVIYSSLGSFFIPLAIMTIVYIEIFVAT (226) KERRAARTLGIIMGVFVICWLPFFLMYVILPF ( 9) FKNFITWLGYINSGLNPVIYTIFNL ( 13)
90. (12) KDHGYTIYSTFGAFYIPLLLMLVLYGRIFRAA (116) RERKTVKTLGIIMGTFILCWLPFFIVALVLPF (10) LGAIINWLGYSNSLLNPVIYAYFNK ( 17)
91. (13) DHILYTVYSTVGAFYFPTLLLIALYGRIYVEA ( 72) RERKATKTLGIILGAFIVCWLPFFIISLVMPI (10) IFDFFTWLGYLNSLINPIIYTMSNE ( 16)
92. (13) SQISYTIYSTCGAFYIPSVLLIILYGRIYRAA ( 70) RERKATKILGIILGAFIICWLPFFVVSLVLPI (10) LMDFFTWLGYLNSLINPIIYTVFNE ( 16)
93. (14) DHVIYTIYSTLGAFYIPLTLILILYYRIYHAA ( 75) RERKAARILGLILGAFILSWLPFFIKELIVGL ( 8) VADFLTWLGYVNSLINPLLYTSFNE ( 16)
94. (12) QNFAYQIYATLGSFYIPLSVMLFVYYQIFRAA ( 98) KEKKASTTLGIIMSAFTVCWLPFFILALIRPF ( 8) LSSLFLWLGYANSLLNPIIYATLNR ( 52)
95. (14) QDVSYQVFATCCTFYVPLMVILALYWKIYQTA (331) RERKAAKTLAIITGAFVVCWLPFFVMALTMPL ( 9) VASLFLWLGYFNSTLNPVIYTIFSP ( 24)
96. (14) QDVGYQIFATCCTFYVPLLVILFLYWKIYIIA (250) RERKAAQTLAIITGAFVICWLPFFVMALTMSL ( 9) VASLFLWLGYFNSTLNPVIYTIFNP ( 24)
97. (12) NDPNFVLIGSFVAFFIPLTIMVITYCLTIYVL ( 63) NERKASKVLGIVFFVFLIMWCPFFITNILSVL (12) LLNVFVWIGYVCSGINPLVYTLFNK ( 85)
98. (11) ADDNFVLIGSFVSFFIPLTIMVITYFLTIKSL ( 55) NEQKACKVLGIVFFLFVVMWCPFFITNIMAVI (12) LLNVFVWIGYLSSAVNPLVYTLFNK ( 86)
99. (14) RFGSFMLFGSLAAFFAPLTIMIVTYFLTIHAL ( 73) NEQRASKVLGIVFLFFLLMWCPFFITNVTLAL (11) LLQIFVWVGYVSSGVNPLIYTLFNK ( 95)
100. (23) LSRTYAISSSVISFYIPVAIMIVTYTRIYRIA ( 44) RETKVLKTLSVIMGVFVCCWLPFFILNCILPF (14) TFDVFVWFGWANSSLNPIIYA FNA (111)
101. (37) LNRTYAISSSLISFYIPVAIMIVTYTRIYRIA ( 37) KETKVLKTLSVIMGVFVCCWLPFFILNCMVPF (18) TFDVFVWFGWANSSLNPVIYA FNA (114)
102. ( 7) ANPAFVVYSSIVSFYVPFIVTLLVYIKIYIVL (150) KEKKATQMLAIVLGVFIICWLPFFITHILNIH ( 8) LYSAFTWLGYVNSAVNPIIYTTFNI ( 12)
103. ( 7) DNPAFVIYSSIVSFYVPFIVTLLVYVQIYIVL (155) KEKKATQMLAIVLGVFIICWLPFFIIHILNMH ( 8) LYSAFTWLGYVNSAVNPIIYTTFNV ( 12)
104. ( 8) SNPDFVIYSSVVSFYVPFGVTVLVYARIYIVL (153) REKKATQMVVIVLGAFIVCWLPFFLTHVLNTH ( 9) LYRATTWLGYVNSALNPVIYTTFNV ( 12)
105. (10) EDRDYVVYSSVCSFFLPCPLMLLLYWATFRGL ( 88) RERKAMRVLPVVVGAFLLCWTPFFVVHITQAL ( 9) LVSAVTWLGYVNSALNPVIYTVFNA ( 14)
106. (10) RAVIIFIAILSFLVFTPLMLVSSTILVVKIRK ( 5) HSSKLYIVIMVTIIIFLIFAMPMRLLYLLYYE ( 7) LHHISLLFSTINSSANPFIYFFVGS ( 40)
107. (10) KLSGLFHAILSLVMCVSSLTLLIRFLCCSQQ ( 0) QKATRVYAVVQISAPMFLLWALPLSVAPLITD ( 6) TSYLISLFLIINSSANPIIYFFVGS ( 58)
108. ( 9) LNMDISLGILLFFLFCPLMVLPCLALILHVEC ( 6) RSAKLNHVVLAIVSVFLVSSIYLGIDWFLFWV ( 7) PEYVTDLCICINSSAKPIVYFLAGR ( 47)
109. (14) AESGDVAFGLLFSMLGGLSVGLSFLLNTVSVA ( 18) SEVEMMAQLLGIMVVASVCWLPLLVFIAQTVL (18) ELLIYLRVATWNQILDPWVYILFRR ( 30)
110. (29) AFASAFACLGLLALVVTFACNLATIKALVSRC ( 15) ITTETAIQLMGIMCVLSVCWSPLLIMMLKMIF (19) SFLIAVRLASLNQILDPWVYLLLRK ( 39)
111. (29) MLTARGIIRFVIGFSLPMSIVAICYGLIAAKI ( 6) KSSRPLRVLTAVVASFFICWFPFQLVALLGTV (16) LVNPTSSLAFFNSCLNPMLYVFVGQ ( 44)
112. (29) MLTVRGIIRFIIGFSAPMSIVAVSYGLIATKI ( 6) KSSRPLRVLSFVAAAFFLCWSPYQVVALIATV (15) AVDVTSALAFFNSCLNPMLYVFMGQ ( 44)
113. (29) MAKVFLILHFIIGFTVPMSIITVCYGIIAAKI ( 6) KSSRPLRVFAAVVASFFICWFPYELIGILMAV (16) LINPTSSLAFFNSCLNPILYVFMGR ( 46)
114. (18) RERAVAIVRLVLGFLWPLLTLTICYTFILLRT ( 6) RSTKTLKVVVAVVASFFIFWLPYQVTGIMMSF (13) LDSLCVSFAYINCCINPIIYVVAGQ ( 45)
115. (21) WLIGMELVSVVLGFAVPFSIIAVFYFLLARAI ( 6) EKHSSRKIIFSYVVVFLVCWLPYHVAVLLDIF (18) ALHVTQCLSLVHCCVNPVLYSFINR ( 42)
116. (20) LPIGLGLTKNILGFLFPFLIILTSYTLIWKAL ( 11) RNDDIFKIIMAIVLFFFFSWIPHQIFTFLDVL (17) AMPITICIAYFNNCLNPLFYGFLGK ( 52)
117. (19) WRMVLRILPHTFGFIVPLFVMLFCYGFTLRTL ( 6) QKHRAMRVIFAVVLIFLLCWLPYNLVLLADTL (18) ALDATEILGFLHSCLNPLIYAFIGQ ( 40)
118. (19) WRMLLRILPQSFGFIVPLLIMLFCYGFTLRTL ( 6) QKHRAMRVIFAVVLIFLLCWLPYNLVLLADTL (18) ALDATEILGILHSCLNPLIYAFIGQ ( 41)
119. (16) WVVVFQFQHIMVGLILPGIVILSCYCIIISKL ( 6) QKRKALKTTVILILAFFACWLPYYIGISIDSF (18) WISITEALAFFHCCLNPILYAFLGA ( 45)
120. (22) VPVLIIHIFIVFSFFLVFLIILFCNLVIIRTL ( 12) VKRRALWMVCTVLAVFIICFVPHHVVQLPWTL (16) AHQVTLCLLSTNCVLDPVIYCFLTK ( 44)
121. (21) YYAYYFSAFSAVFFFVPLIISTVCYVSIIRCL ( 9) KKSRALFLSAAVFCIFIICFGPTNVLLIAHYS (13) AYLLCVCVSSISSCIDPLIYYYASS ( 49)
122. (20) WEVFTNMLLNVVGFLLPLSVITFCTMQIMQVL ( 12) TERRATVLVLVVLLLFIICWLPFOISTFLDTL (17) ITQIASFMAYSNSCLNPLVYVIVGK ( 54)
123. (19) WLVGFVLYTFLMGFLLPVGAICLCYVLIIAKM ( 14) SERKITLMVMAVVVAVFVICWMPFYVVQLVNVF ( 7) VSQLSVILGYANSCANPILYGFLSD ( 63)
124. (19) WYTGFIIYTFILGFLVPLTIICLCYLFIIIKV ( 14) SEKKVTRMVSIVVAVFIFCWLPFYIFNVSSVS (11) MFDFVVVLTYANSCANPILYAFLSD ( 52)
consensus                                                                                                    P
```

Appendix A (Continued)

G protein α subunit (1)

```
1.  ( 38) ATHRLLLLGAGESGKSTIVKQMRILHV NGFNGE (15) EKATKVQDIKNNLKEAIETIVAAMSNL ( 32) FYEHAKALWE DEGVRACYERSN EY QLID  CAQYFLDKIDVIKQADYVPSDQDLLRCRVLTSGIFETKFQV DKVN  FHMF
2.  ( 38) ATHRLLLLGAGESGKSSIVKQMRILHV NGFNAE ( 0) EKKTKVQDIKNNIKEAIETIVTAMGNL ( 32) FYEHTKTLWQ DEGVRACYERSN EY QLID  CAQYFLDKIDIVKQNDYTPSDQDLLRCRVLTSGIFETKFQV DKVN  FHMF
3.  ( 40) ATHRLLLLGAGESGKSTIVKQMRILHV NGFNPE ( 0) EKKQKILDIRKNVKDAIVTIISAMSTI ( 32) FFDHVKKLWD DEGVKACFERSN EY QLID  CAQYFLERIDSVSLVDYTPTDQDLLRCRVLTSGIFETRFQV DKVN  FHMF
4.  ( 41) ATHRLLLLGAGESGKSTIVKQMRILHV DGFSDS ( 0) EKKQKIDDIKKNIRDAILTITGAMSTL ( 32) FYEHTEELWK DKGVLQTYERSN EY QLID  CAKYFLDRVSTIKNPNYTPNEQDILRCRVLTSGIFETRFQV DKVN  FHMF
5.  ( 35) GTHRLLLLGAGESGKSTIVKQMRILHV NGFSPE ( 0) ERKQKIEDIKRNVRDAILTITGAMSTL ( 32) FYEYTEILWK DKGVQAAFERSN EY QLID  CAQYFLDRVHIIRQAEYTPSEQDILRCRVLTSGIFETKFSV DKVN  FHMF
6.  ( 37) STHRLLLLGAGESGKSTIVKQMRILHI DGFSER ( 0) EKKEKIDAIRKNLRDAICSIAGAMGSL ( 32) FFTYCAKLWK DGGIQETFERSN EY QLID  CAKYFLDKALEVGAPNYIPSEQDILRCRVLTSGIFETKFSV DKVN  FHMF
7.  ( 26) REVKLLLLGAGESGKSTIVKQMKIIHE AGYSEE ( 0) ECKQYKAVVYSNTIQSIIAIIRAMGRL ( 31) LAGVIKRLWK DSGVQACFNRSR EY QLND  SAAYYLNDLDRIAQPNYIPTQQDVLRTRVKTTGIVETHFTF KDLH  FKMF
8.  ( 31) REVKLLLLGAGESGKSTIVKQMKIIHE AGYSEE ( 0) ECKQYKAVVYSNTIQSIIAIIRAMGRL ( 31) LAGVIKRLWK DGGVQACFNRSR EY QLND  SAAYYLNDLDRIAQNSYIPTQQDVLRTRVKTTGIVETHFTF KDLH  FKMF
9.  ( 31) KEVKLLLLGAGESGKSTIVKQMKIIHE DGYSED ( 0) ECKQYKVVVYSNTIQSIIAIIRAMGRL ( 31) LAGVIKRLWR DGGVQACFSRSR EY QLND  SASYYLNDLDRISQSNYIPTQQDVLRTRVKTTGIVETHFTF KDLY  FKMF
10. ( 22) KEVKLLLLGAGESGKSTIVKQMKIIHE DGYSEE ( 0) ECRQYKVVVYSNTIQSIIAIIRAMGRL ( 31) LAGVIQRLWE DSGVQACFSRSR EY QLND  SASYYLSDIERIAQGSYIPTQQDVLRTRVKTTGIVETHFTF KDLY  FKMF
11. ( 31) REVKLLLLGAGESGKSTIVKQMKIIHE DGYSEE ( 0) ECRQYRAVVYSNTIQSIMAIVKAMGNL ( 32) LSGVIRRLWA DHGVQACFGRSR EY QLND  SAAYYLNDLERIAQSDYIPTQQDVLRTRVKTTGIVETHFTF KDLH  FKMF
12. ( 31) REVKLLLLGAGESGKSTIVKQMKIIHE EGYSEE ( 0) DCKQYKPVVYSNTIQSMIAIIRAMGSL ( 31) LAAVMKRLWA DGGVQGCFSRSR EY QLND  SASYYLNALDRLAAPGYIPTQQDVLRTRVKTTGIVETHFTF KDLH  FKMF
13. ( 31) REVKLLLLGAGESGKSTIVKQMKIIHE KGYSQE ( 0) ECLQYKPVVYSNAIQSMIAIIKAMGQL ( 31) LSGIMKRLWK DVGVQECFSRSR EY QLND  SAEYYLNALDRISAPGYIPTEQDVLRTRVKTTGIVETHFTF KDLH  FKMF
14. ( 32) SEVKLLLLGAGESGKSTIVKQMKIIHD TGYSQE ( 0) ECEEYRRVVFSNTVQSLMVIIRAMGRL ( 31) IVLLMKKLWA DGGVQQTFARSR EY QLND  SAGYYLNSLDRIAQPNYIPTQQDVLRTRVKTTGIIETHFSC KQLH  FKLF
15. ( 31) KDVKLLLLGAGESGKSTIVKQMKIIHE DGFSGE ( 0) DVKQYKPVVYSNTIQSLAAIVRAMDTL ( 32) LLSAMMRLWG DSGIQECFNRSR EY QLND  SAKYYLDSLDRIGAADYQPTEQDILRTRVKTTGIVETHFTF KNLH  FRLF
16. ( 31) KDVKLLLLGAGESGKSTIVKQMKIIHE DGFSGE ( 0) DVKQYKPVVYSNTIQSLAAIVRAMDTL ( 32) LLSAMVRLWA DSGIQECFNRSR EY QLND  SAKYYLDSLDRIGAPDYQPTEQDILRTRVKTTGIVETHFTF KNLH  FRLF
17. ( 31) KDIKLLLLGAGESGKSTIVKQMKIIHE SGFTAE ( 0) DFKQYRPVVYSNTIQSLVAILRAMPTL ( 32) LLAAMKRLWQ DAGVQECFSRSN EY QLND  SAKYFLDDLDRLGAKDYQPTEQDILRTRVKTTGIVEVHFSF KNLN  FKLF
18. ( 31) KDIKLLLLGAGESGKSTIVKQMKIIHE GGFTSE ( 0) DNKQYKPVVYSNTIQSLVAIIRAMGTL ( 32) LLAAMKRLWV DSGVQECLGRAN EY QLND  SAKYFLDDLDRLGAKDYMPTEQDILRTRVKTTGIVEVHFSF KNLN  FKLF
19. ( 31) KDIKLLLLGAGESGKSTIVKQMKIIHE SGFTAE ( 0) DYKQYKPVVYSNTVQSLVAILRAMSNL ( 32) LLSSMKRLWG DAGVQDCFSRSN EY QLND  SAKYFLDDLERLGEAIYQPTEQHILRTRVKTTGIVEVHFTF KNLN  FKLF
20. ( 31) REIKLLLLGTSNSGKSTIVKQMKIIHS GGFNLE ( 0) ACKEYKPLIIYNAIDSLTRIIRALAAL ( 32) LLGVMRRLWA DPGAQACFSRSS EY HLED  NAAYYLNDLERIAAADYIPTVEDILRSRDMTTGIVENKFTF KELT  FKMV
21. ( 27) RTVKLLLLGAGESGKSTIVKQMKIIHQ DGYSLE ( 0) ECLEFIAIIYGNTLQSILAIVRAMTTL ( 31) MSDIIQRLWK DSGIQACFERAS EY QLND  SAGYYLSDLERLVTPGYVPTEQDVLRSRVKTTGIIETQFSF KDLN  FRMF
22. ( 31) KTVKLLLLGAGESGKSTIVKQMKIIHQ DGYSPE ( 0) ECLEFKAIIYGNVLQSILAIIRAMTTL ( 31) LVEVIRRLWK DGGVQACFERAA EY QLND  SASYYLNQLERITDPEYLPSEQDVLRSRVKTTGIIETKFSV KDLN  FRMF
23. ( 31) RTVKLLLLGAGESGKSTIVKQMKIIHK NGYSKQ ( 0) ECMEFKAVVYSNTLQSILAIVKAMTTL ( 31) LAEIIKRLWG DPGIQACFERAS EY QLND  SAAYYLNDLDRLTAPGYVPNEQDVLHSRVKTTGIIETQFSF KDLN  FRMF
24. ( 37) RELKLLLLGTGESGKSTFIKQMRIIHG SGYSDE ( 0) DKRGFTKLVYQNIFTAMQAMIRAMDTL ( 30) YVDAIKSLWN DPGIQECYDRRR EY QLSD  STKYYLNDLDRVADPSYLPTQQDVLRVRVPTTGIIEYPFDL QSVI  FRMV
25. ( 37) RELKLLLLGTGESGKSTFIKQMRIIHG AGYSEE ( 0) DKRGFTKLVYQNIFTAMQAMIRAMETL ( 30) YVSAIKTLWE DPGIQECYDRRR EY QLSD  SAKYYLTDVDRIATLGYLPTQQDVLRVRVPTTGIIEYPFDL ENII  FRMV
26. ( 33) RELKLLLLGTGESGKSTFIKQMRIIHG SGYSDE ( 0) DRKGFTKLVYQNIFTAMQAMIRAMDTL ( 30) QVAAIKQLWL DPGIQECYDRRR EY QLSD  SAKYYLTDIERIAMPSFVPTQQDVLRVRVPTTGIIEYPFDL ENII  FRMV
27. ( 31) RELKLLLLGTGESGKSTFIKQMRIIHG SGYSDE ( 0) DKRGYIKLVFQNIFMAMQSMIKAMQML ( 30) YLNAIKTLWD DAGIQECYDRRR EY QLTD  SAKYYLSDLARIEQADYLPTEQDILRARVPTTGILEYPFDL DGIV  FRMV
28. ( 40) GELKLLLLGPGESGKSTFIKQMRIIHG AGYSEE ( 0) ERKGFRPLVYQNIFVSMRAMIEAMERL ( 30) YAAAMQWLWR DAGIRACYERRR EF HLLD  SAVYYLSHLERITEEGYVPTAQDVLRSRMPTTGINEYCFSV QKTN  LRIV
29. ( 40) EELKLLLLGPGESGKSTFIKQMRIIHG VGYSEE ( 0) DRRAFRLLIYQNIFVSMQAMIDAMDRL ( 30) YAVAMQYLWR DAGIRACYERRR EF HLLD  SAVYYLSHLERISEDSYIPTAQDVLRSRMPTTGINEYCFSV KKTK  LRIV
30. ( 53) RLVKILLLGAGESGKSTFLKQMRIIHG REFDQK ( 0) ALLEFRDTIFDNILKGSRVLVDARDKL ( 34) YVPALSALWR DSGIREAFSRRS EF QLGE  SVKYFLDNLDRIGQLNYFPSKQDILLARKATKGIVEHDFVI KKIP  FKMV
31. ( 46) RLVKILLLGAGESGKSTFLKQMRIIHG QDFDQR ( 0) AREEFRPTIYSNVIKGMRVLVDAREKL ( 38) YLPAIRALWE DSGIQNAYDRRR EF QLGE  SVKYFLDNLDKLGVPDYIPSQQDILLARRPTKGIHEYDFEI KNVP  FKMV
32. (130) RQVKLLLLGAGESGKSTFLKQMRIIHG VNFDYE ( 0) LLLEYQSVIYQNVIRGMQVLLDAREKL ( 31) YAPPISRLWQ DRGIRRAFERRR EF QISD  SVSYFLDEIQRLATPDYVPTHKDILHCRKATKGVYEFCVKV QNIP  FVFV
33. ( 29) KDVKLLLLGPGESGKSTIFKQMKIIQEDGGYSVE ( 0) ELLEYRAFVYSNCISQMEALLTASAKL ( 27) LAADIKHLWE DKGIKETYAQKD KHF QLND  SAAYFFDNIDRYMREDFVPNEQDVLRCRVRTTGIQESEFTF DKIR  LKIV
34. ( 34) GEIKLLLLGAGESGKSTIAKQMKIIHL NGFNDE ( 0) EKSSYKTIIYNNTVGSMRVLVNAAEEL ( 30) LAQDIKALWA DPGIQNTFQRSS EF QLND  SAAYYFDSIDRISQPLYLPSENDVLRSRTKTTGIIETVFEI QNST  FRMV
35. ( 29) NEVKLLLLGAGESGKSTISKQMKIIHQ SGYSNE ( 0) ERKEFKPIITRNILDNMRVLLDGMGRL ( 36) QGKKIKALWT DPGVKQAMRRAN EFSTLPD  SAPYFFDSIDRMTSPVYIPTDQDILHTRVMTRGVHETNFEI GKIK  FRLV
36. ( 31) NIIKLLLLGAGESGKSTVLKQMKIIHN SGFSQE ( 0) EISNKRNVVCANTVQAMGALLDGMKQL ( 32) MFNALTELWA DKGVQCAYDKRE FF YLHD  SAKYFLDRIARVHTPNYVPTENDILHTRVPTMGVIEVNFTI KGKF  FRVF
37. ( 31) RTVKLLLLGAGECGKSTVLKQMRLLTS KQYTDE ( 0) ELLTQAKLVYTNIVIEMDHLVKAMPAA ( 31) AADHVEKLWK DPVVKRLYAERK EL NIRDIGDNTEYFFENLPRISKEDYHPNATDTLLLRTKTTGIVEVGFEI KKVK  FRVF
38. ( 31) KVVKLLLLGAGECGKSTVLKQMRILHD HGFTAE ( 0) EAEQQKSVVFNNTLQAMTAILKGMEAL ( 32) LANAIQALWN DKAVQQVIAKGN EF QMPE  SAPHFLSSLDRIKLPDYNPTEQDILLSRIKTTGIVEVKFQM KSVD  FRVL
39. ( 72) NDIKVLLLGAGDSGKTTIMKQMRLLYS PGFSQV ( 0) VRKQYRVMIFENIISSLCLLLEAMDNS ( 30) IYEAVHALTL DTKLRTVQSCGT NL SLLD  NFYYYQDHIDRIFDPQYIPSDQDILHCRIKTTGISEETFLL NRHH  YRFF
40. ( 32) NEIKMLLLGAGESGKSTVLKQMKLIHH GGYSDQ ( 0) EKDSYKEIIFSNTVQSMRAILDALPAL ( 30) IADAIRQLWA DPGLKEAVRRSR EF QLND  SAVYYFNSIDRMSAPGYLPTDQDILRSRVKTTGITETTFKV GELT  YKLF
41. (121) KELKVLLLGAGESGKSTVLQQLKILHQ NGFSEQ ( 0) EIKEYIPLIYQNLLEIGRNLIQARTRF ( 34) IAGVISTLWA LPSTQDLVNGPNASKF YLMD  STPYFMENFTRITSPNYRPTQQDILRSRQMTSGIFDTVIDMGSDIK MHIY
42. ( 39) NEIKLLLLGAGESGKSTVLKQLKLLHQ GGFSHQ ( 0) ERLQYAQVIWADAIQSMKILIIQARKL (141) IAKAIKQLWNNDKGIKQCFARSN EF QLEG  SAAYYFDNIEKFASPNYVCTDEDILKGRIKTTGITETEFNI GSSK  FKVL
43. ( 39) KGVKLLLLGAGESGKSTVLKQLKLLHK GGFTQQ ( 0) ERRQYSHVIWCDVIQSMKVLIIQARKL ( 98) IAEAIHKLWKLDSGIKKCFDRSN EF QLEG  SADYYFDNVVNFADTNYLSTDLDILKGRIKTTGITETDFLI KSFQ  FKVL
44. ( 36) HIRKLLLLGAGESGKSTIFKQIKLLFQ TGFDEG ( 0) ELKSYVPVIHANVYQTIKLLHDGTKEF ( 38) IAEGIETLWK DPAIQETCARGN EL QVPD  CTKYLMENLKRLSDINYIPTKEDVLYARVRTTGVVEIQFSPVGENKKSGEVYRLF
45. ( 37) HIQKLLLLGAGDSGKSTIFKQIKLLFQ TGFDEE ( 0) ELKNYIPVIHANVYQTTKILHDGSKEL ( 38) LVQDIEALWK DPAIQETLLRGN EL QVPD  CAHYFMENLERFSDVHYIPTKEDVLFARIRTTGVVEIQFSPVGENKKSGEVYRLF
consensus           LLLG   GK    0                    L                    L R T G
```

Appendix B

G protein α subunit (2)

```
1.  DVGGQRDERRKWIQCFN DVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKQDLLAEKVLAGKSKIE  DYFPEFARYTTP (18) FIRDEFLR ( 0) ISTASGDGRHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHL (6)
2.  DVGGQRDERRKWIQCFN DVTAIIFVVASSSYNMVIREDNHTNRLQEALNLFKSIWNNRWLRTISVILFLNKQDLLAEKVNAGKSKIE  DYFPEFARYTTP (18) FIRDEFLR ( 0) ISTASGDGRHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHL (6)
3.  DVGGQRDERRKWIQCFN DVTAIIYVAACSSYNMVIREDNNTNRLRESLDLFESIWNNRWLRTISIILFLNKQDMLAEKVLAGKSKIE  DYFPEYANYTVP (18) FIRDLFLR ( 0) ISTATGDGKHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHL (6)
4.  DVGGQRDERRKWIQCFN DVTAIIFVTACSSYNMVLREDPTONRLRESLDLFKSIWNNRWLRTISIILFLNKQDLLAEKIKAGKSKLS  EYFSEFNKYQTP (21) FIRDEFLR ( 0) ISTASGDGKHYCYPHFTCAVDTENIKRVFNDCRDIIQRMHL (6)
5.  DVGGQRDERRKWIQCFN DVTAIIFVTACSGYNMVLREDATONRLKESLDLFKSIWNNRWLRTISVILFLNKQDLLAEKVKAGKSKIE  DYFPEYARYQVP (18) FIRDEFLR ( 0) ISTASGDGRHYCYPHFTCAVDTENIRRVFDDCRDIIQRMHL (6)
6.  DVGGQREERRKWIQCFN DVTAIIFVAACSSYNMVLREDPSONRVKESLELLASIWNNRWLRNISVILFLNKQDLLTEKVLAGKSKIE  VYFPHYATYQAP (19) FFRDEFLK ( 0) VTSNNNGGRHYCYPHLTCAVDTENIRRVFNDCRDIIQRMHL (6)
7.  DVGGQRSERKKWIHCFE GVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  KKSPLT  ICYPEYAGSNTY ( 5) YIQCQFED ( 0) LNKRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKNNL (6)
8.  DVGGQRSERKKWIHCFE GVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  KRSPLT  ICYPEYPGSNTY ( 5) YIQCQFED ( 0) LNKRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKNNL (6)
9.  DVGGQRSERKKWIHCFE GVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNKKDLFEEKI  KRSPLT  ICYPEYTGSNTY ( 5) YIQCQFED ( 0) LNRRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKNNL (6)
10. DVGGQRSERKKWIHCFE GVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFIDTSIILFLNKKDLFEEKI  SRSPLT  ICYPEYSGSNTY ( 5) YIQCQFED ( 0) LNRRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKSNL (6)
11. DVGGQRSERKKWIHCFE GVTAIIFCVALSAYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  THSPLT  ICFPEYTGANKY ( 5) YIQSKFED ( 0) LNKRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKNNL (6)
12. DVGGQRSERKKWIHCFE GVTAIIFCVALSAYDLVLAEDEEMNRMHESMKLFDSICNNKWFTETSIILFLNKKDLFEEKI  TKSPLT  ICFPEYTGSNTY ( 5) YIQMQFED ( 0) LNKRKDQKEIYT  HFTCATDTNNIQFVFDAVTDVIIKNNL (6)
13. DVGGQRSERKKWIHCFE GVTAIIFIVAMSEYDLTLAEDQEMNRQMESMKLFDSICNNKWFTETSIILFLNKKDLFEEKI  KKSPLT  ICFPEYTGANTY ( 5) YIQLQFEN ( 0) LNKKKDTKEIYS  HFTCATDTNNVQFVFDAVTDVIIKNNL (6)
14. DVGGQRSERKKWIHCFE GVTAIIFCVALSGYDLVLAEDEEMNRMIESLKLFDSICNSKWFVETSIILFLNKKDLFEEKI  KRSPLT  ICFPEYTGTNTF ( 5) YIRMKFEN ( 0) LNKRKDQKEIYT  HLTCATDTNNVKFVFDAVTDVIIKNNL (6)
15. DVGGQRSERKKWIHCFE DVTAIIFCVALSGYDOVLHEDETTNRMHESLKLFDSICNNKWFTDTSIILFLNKKDIFEEKI  KKSPLT  ICFPEYTGPSAF ( 5) YIQAQYES ( 0) KNKSAH KEIYS  HVTCATDTNNIQFVFDAVTDVIIAKNL (6)
16. DVGGQRSERKKWMHCFE DVTAIIFCVALTGYDOVLHEDETTNRMHESLKLFDSICNNKWFTDTSIILFLNKKDIFQEKI  KSSPLT  ICFPEYTGPNSF ( 5) HTQHQYES ( 0) RNKSEN KEIYT  HITCATDTQNIQFVFDAVTDVIIAYNL (6)
17. DVGGQRSERKKWIHCFE DVTAIIFCVAMSEYDOVLHEDETTNRMQESLKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  RKSPLT  ICFPEYTGGQEY ( 5) YIQAQFEA ( 0) KNKSTS KEIYC  HMTCATDTNNIQFVFDAVTDVIIANNL (6)
18. DVGGQRSERKKWIHCFE DVTAIIFCVAMSEYDOVLHEDETTNRMQESLKLFDSICNNKWFTETSIILFLNKKDLFEEKI  KKSPLT  ICFPEYTGKQMY ( 5) YIQAQFEA ( 0) KNKSSA KEIYC  HQTCATDTNNIQFVFDAVTDVIIANNL (6)
19. DVGGQRSERKKWIHCFE DVTAIIFCVAMSEYDQLLHEDETTNRMHESLKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  KKSPLT  ICFPEYSGRQDY ( 5) YIQAQFEA ( 0) KNKSAN KEIYC  HMTCATDTTNIQFVFDAVTDVIIANNL (6)
20. DVGGQRSERKKWIHCFE GVTAIIFCVELSGYDLKLYEDNQTSRMAESLRLFDSICNNNWFINTSLILFLNKKDLLAEKI  RRIPLT  ICFPEYKGQNTY ( 5) YIQRQFED ( 0) LNRNKETKEIYS  HFTCATDTSNIQFVFDAVTDVIIQNNL (6)
21. DVGGQRSEPKKWIHCFE GVTCIIFIAALTAYDMVLVEDDEVNRMHESLHLFNSICNHRYFATTSIVLFLNKKDVFFEKV  KKAHLS  ICFPDYDGPNTY ( 5) YIKVQFLE ( 0) LNMRRDVKEIYS  HMTCATDTQNVKFCFDAVTDIIIKENL (6)
22. DVGGQRSERKKWIHCFE GVTCIIFCAALSAYDMVLVEDDEVNRMHESLHLFNSICNHKFFAATSIVLFLNKKDLFEEKI  KKVHLS  ICFPEYDGNNSY ( 5) YIKSQFLD ( 0) LNMRKDVKEIYS  HMTCATDTQNVKFVFDAVTDIIIKENL (6)
23. DVGGQRSERKKWIHCFE GVTCIIFCAALSAYDMVLVEDEEVNRMHESLHLFNSICNHKYFATTSIVLFLNKKDLFQEKV  TKVHLS  ICFPEYTGPNTF ( 5) YIKNQFLD ( 0) LNLKKEDKEIYS  HMTCATDTQNVKFVFDAVTDIIIKENL (6)
24. DVGGQRSERRKWIHCFE NVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKI  MYSHLV  DYFPEYDGPQRD ( 6) FILKMFVD ( 0) LNPDSD KIIYS  HFTCATDTENIRFVFAAVKDTILQLNL (6)
25. DVGGQRSERRKWIHCFE NVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEDKI  LYSHLV  DYFPEFDGPQRE ( 6) FILKMFVD ( 0) LNPDSD KIIYS  HFTCATDTENIRFVFAAVKDTILQLNL (6)
26. DVGGQRSERRKWIHCFE SVTSIIFLVALSEYDQVLAECDNENRMEESKALFRTIITYPWFLNSSVILFLNKKDLLEEKI  MYSHLI  SYFPEYTGPKQD ( 6) FILKLYQD ( 0) QNPDKE KVIYS  HFTCATDTENIRFVFAAVKDTILQLNL (6)
27. DVGGQRSERRKWIHCFE NVTSIIFLVALSEYDQILFESDNENRMEESKALFRTIITYPWFQNSSVILFLNKKDLLEEKI  MYSHLV  DYFPEYDGPKQD ( 6) FVLKKYLA ( 0) CNPDPE RQCYS  HFTTATDTENIKLVFCAVKDTIMQNAL (6)
28. DVGGQKSERKKWIHCFE NVIALIYLASLSEYDQCLEENNQENRMKESLALFGTILELPWFKSTSVILFLNKTDILEEKI  PTSHLA  TYFPSFQGPKQD ( 6) FILDMYTR (12) KKGARS RRLFS  HYTCATDTQNIRKVIKDVRDSVLARYL (6)
29. DVGGQRSERRKWIHCFE NVIALIYLASLSEYDQCLEENDQENRMEESLALFSTILELPWFKSTSVILFLNKTDILEDKI  HTSHLA  TYFPSFQGPRRD ( 6) FILDMYAR (12) RKGSRA RRFFA  HFTCATDTQSVRSVFKDVRDSVLARYL (6)
30. DVGGQRSQRQKWFQCFD GITSILFMVSSSEYDQVLMEDRRTNRLVESMNIFETIVNNKLFFNVSIILFLNKMDLLVEKV  KSVSIK  KHFPDFKGDPHR ( 6) YLVQCFDR ( 0) KRRNRS KPLFH  HFTTAIDTENIRFVFHAVKDTILQENL (6)
31. DVGGQRSERKRWFECFD SVTSILFLVSSSEFDQVLMEDRQTNRLTESLNIFETIVNNRVFSNVSIILFLNKTDLLEEKV  QVVSIK  DYFLEFEGDPHC ( 6) FLVECFRG ( 0) KRRDQQQRPLYH  HFTTAINTENIRLVFRDVKDTILHDNL (6)
32. DVGGQRTQRQKWTRCFDSSVTSIIFLVSSSEFDQVLAEDRKTNRLEESKNIFDTIVNNATFKGISIILFLNKTDLLEQKVCNPETDIR  WYYPHFNGNPHS ( 6) FILQMFMS ( 0) VRRSSSISRIYH  HFTTAIDTRNINVVFNSVKDTILQRNL (6)
33. DVGGQRSQRRKWIHCFD CVTAVIFVAAMSDYDQVLREDESVNRTRESLALFKEIVNCDYFKETPIVLFLNKKDLFKEKL  KRVPLQ  SCFSDYTGPNKY ( 5) FIQSQYLA ( 0) QGPSP  RTIYT  HATCAVDTENIKFVFRAVRQTILSQAL (3)
34. DVGGQRSERKKWMHCFQ EVTAVIFCVALSEYDLKLYEDDTTNRMQESLKLFKEICNTKWFANTAMILFLNKRDIFSEKI  TKTPIT  VCFKEYDGPQTY ( 5) FIKQQFIN ( 0) QNENPK KSIYP  HLTCATDTNNILVVFNAVKDIVLNLTL (7)
35. DVGGQRSERKKWLSCFD DVTAVVFCVALSEYDLLLYEDNSTNRMLESLRVFSDVCNS WFVNTPIILFLNKSDLFREKI  KHVDLS  ETFPEYKGGRDY ( 5) YIKERFWQ ( 0) INKTEQ KAIYS  HITCATDTNNIRVVFEAVKDIIFTQCV (7)
36. DVGGQRSQRKKWIHCFD DAKAMIYVASLSEYDQVLLEDNTTNRMHESIQLFKQVINNKYFVNTSVILFLNKIDLFEEKIVTKKRSLG  IAFESFSGPSQD ( 6) FVEKKYRS ( 0) MAENKE KNIYC  HHTCATDTQQVQYVLDAVLDTILSTKL (6)
37. DVGGQRSERKKWIHCFE DVNAIIFIAALSEYNEVLFEDETTNRMIESMRLFESICNSRWFHNTNIILFLNKKDLFEEKI  KKENIH  KAFPEYRGEQNY ( 5) FIKTKFEA ( 0) LSNNPK KTFYV  HETCATDTNQVQKILDSVISMIIQSNL (6)
38. DVGGQRSERKKWIHCFE DVNAIIFIAAISEYDQVLFEDETTNRMIESMRLFESICNSRWFINTSMILFLNKKDLFAEKI  KRTSIK  SAFPDYKGAQTY ( 5) YIEEKFDG ( 0) LNANPE KTIYM  HQTCATDTDQVQMILDSVIDMIIQANL (6)
39. DVGGQRSERRKWIHCFE NVTALLFLVSLAGYDQCLVEDNSGNQMQEALLLWDSICNSSWFSESAMILFLNKLDLFKRKG  SHFPIQ  KHFPDYQEVGST (20) YFYLKFES ( 0) LNRIAS RSCYC  HFTTATDTSLLQRVMVSVQDTIMSNNL (5)
40. DVGGQRSERKKWIHCFE NVTALVFLVSLSEYDQMLYEDESVNRMQEALTLFDSICNSRWFVKTSIILFLNKIDLFAEKL  PARRS  TYFPDFTGGDNY ( 5) YLLHRFVS ( 0) LNQSAATKQIYA  HYTCATDTQQIKFVLSAIQDILLQLHL (6)
41. DVGGQRSERKKWIHCFD NVTLVIFCVSLSEYDQTLMEDKNQNRFQESLVLFDNIVNSRWFARTSVVLFLNKIDLFAEKL  RKVPME  NYFPDYTGGSDI ( 5) YILWRFVQ ( 0) LNRAN  LSIYP  HVTQATDTSNIRLVFAAIKETILENTL (7)
42. DAGGQRSERKKWIHCFE GITAVLFVLAMSEYDQMLFEDERVNRMHESIMLFDTLLNSKWFKDTPFILFLNKIDLFEEKV  KSMPIR  KYFPDYQGRVGD ( 6) YFEKIFLS ( 0) LNKTN  KPIYV  KRTCATDTQTQMKFVLSAVTDLIIQQNL (6)
43. DAGGQRSVRKKWIHCFE DITAVLFVLAISEYDQNLFEDERVNRMHESIVLFDSLCNSKWFANTPFILFLNKIDIFENKI  KKNPLK  NYFPDYDGKPDD ( 6) FFETNFLK ( 0) INQTN  KPIYV  HRTCATDSKSMKFVLSAVTDMIVQQNL (6)
44. DVGGQRNERRKWIHLFE GVTAVIFCAAISEYDQTLFEDEQKNRMMETKELFDWVLKQPCFEKTSFMLFLNKFDIFEKKV  LDVPLNVCEWFRDYQPVSSG (13) KFEELYYQ ( 0) NTAPDRVDRVFKI YRTTALDQKLVKKTFKLVDETLRRRNL (6)
45. DVGGQRNERRKWIHLFE GVTAVIFCAAISEYDQTLFEDERKNRMMETKELFEWVLKQPCFEKTSFMLFLNKFDIFEQKV  PKVPLNACEWFKDYQSVSTG (13) KFEESYFQ ( 0) CTAPDRVDRVFKI YRTTALDQKLVKKTFKLVDETLRRRNL (6)
consensus D GGQ     W  F                  E      E                  LFLNK D   K                                                  T A
```

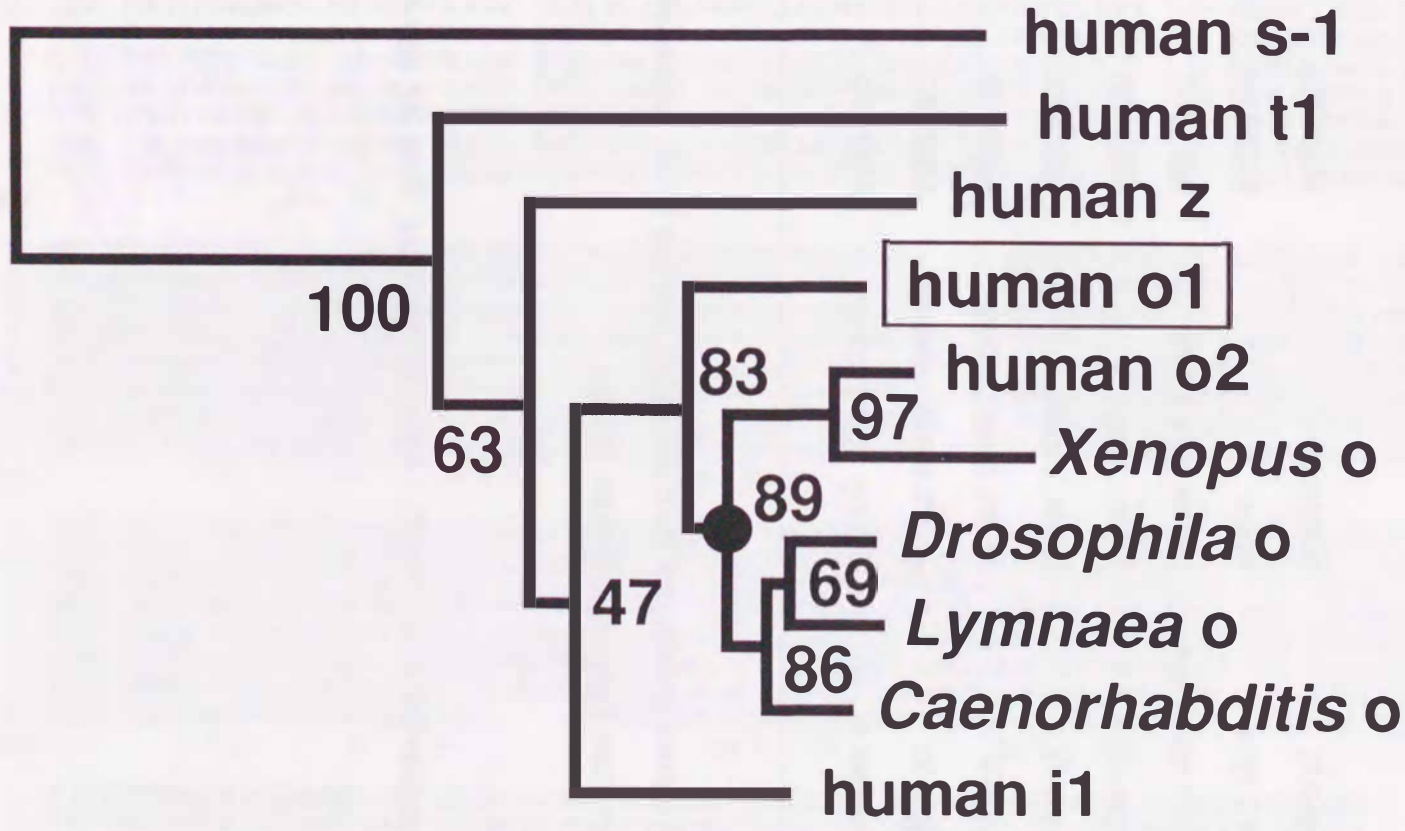
Appendix B (Continued)

G protein α subunits (αo subunit exon 7 and 8 region)

```
          1. (240) TNRMHESLMLFDSICNNKFFIDTSIILFLNKKDLFGEKI  KKSPLTICFPEYTGPNTYED
          2. (240) TNRMHESLKLFDSICNNKWFTDTSIILFLNKKDIFEEKI  KKSPLTICFPEYTGPSAFTE
          3. (240) TNRMHESLKLFDSICNNKWFTDTSIILFLNKKDIFQEKI  KSSPLTICFPEYTGPNSFTE
          4. (240) TNRMQESLKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  RKSPLTICFPEYTGGQEYGE
          5. (240) TNRMQESLKLFDSICNNKWFTETSIILFLNKKDLFEEKI  KKSPLTICFPEYTGKQMYQE
          6. (240) TNRMHESLKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  KKSPLTICFPEYSGRQDYHE
          7. (234) MNRMHESMKLFDSICNNKWFTDTSIILFLNKKDLFEEKI  KKSPLTICYPEYAGSNTYEE
          8. (240) TSRMAESLRLFDSICNNNWFINTSLILFLNKKDLLAEKI  RRIPLTICFPEYKGQNTYEE
          9. (235) VNRMHESLHLFNSICNHRYFATTSIVLFLNKKDVFFEKV  KKAHLSICFPDYDGPNTYED
         10. (262) TNRLQEALNLFKSIWNNRWLRTISVILFLNKQDLLAEKVLAGKSKIEDYFPEFARYTTPED
consensus            R  E   LF SI N       S  LFLNK D   EK             P

          1. AAA ( 0) YIQAQFESKNRSPN KEIYC  HMTCATDTNNIQVVFDAVTDIIIANNLRGCGLY (0)
          2. AVA ( 0) YIQAQYESKNKSAH KEIYS  HVTCATDTNNIQFVFDAVTDVIIAKNLRGCGLY (0)
          3. AVA ( 0) HTQHQYESRNKSEN KEIYT  HITCATDTQNIQFVFDAVTDVIIAYNLRGCGLY (0)
          4. AAA ( 0) YIQAQFEAKNKSTS KEIYC  HMTCATDTNNIQFVFDAVTDVIIANNLRGCGLY (0)
          5. ASA ( 0) YIQAQFEAKNKSSA KEIYC  HQTCATDTNNIQFVFDAVTDVIIANNLRGCGLY (0)
          6. ASA ( 0) YIQAQFEAKNKSAN KEIYC  HMTCATDTTNIQFVFDAVTDVIIANNLRGCGLY (0)
          7. AAA ( 0) YIQCQFEDLNKRKDTKEIYT  HFTCATDTKNVQFVFDAVTDVIIKNNLKDCGLF (0)
          8. AAV ( 0) YIQRQFEDLNRNKETKEIYS  HFTCATDTSNIQFVFDAVTDVIIQNNLKYIGLC (0)
          9. AGN ( 0) YIKVQFLELNMRRDVKEIYS  HMTCATDTQNVKFCFDAVTDIIIKENLKDCGLF (0)
         10. ATP (13) FIRDEFLRISTASGDGRHYCYPHFTCAVDTENIRRVFNDCRDIIQRMHLRQYELL (0)
consensus A                              Y   H TCA DT N    F    D I    L    L
```

Appendix B'(a)



Appendix B'(b)

Appendix C1

Expressions and Accession Numbers of Adenylyl Cyclases

No.	Gene	Expression	Accession No.
1.	rat AC type-II	olfactory	M55075
2.	bovine AC type-I	brain	M25579
3.	*Drosophila* rutabaga	mushroom body	M81887
4.	rat AC type-II	brain, lung	M80550
5.	rat AC type-IV	brain, others	M80633
6.	rat AC type-V	heart, brain, others	M96159
7.	rat AC type-VI	heart, brain, others	M96160
8.	*Dictyostelium* AC-A	during aggregation	M87279

Note. - Sequence data are from GenBank release 77.0. Expression data were taken from Tang and Gilman (1992).

References

Tang, W.-J. and Gilman, A. G. (1992). Adenylyl cyclases. Cell 70, 869-872.

adenylyl cyclase

```
1. (306) FNTMYMYRHENVSILFADIVGFTQLSSAC SAQELVKLLNELFARFDKLAAKYHQLRIKILGDCYYCICGLPDYR ( 0) EDHAVCSILMGLAMVEAISY
2. (292) FHKIYIQRHDNVSILFADIVGFTGLASQC TAQELVKLLNELFGKFDELATENHCRRIKILGDCYYCVSGLTQPK ( 0) TDHAHCCVEMGLDMIDTITS
3. (262) FHRIYIQKHENVSILFADIVGFTVLSSQC SAQELVRLLNELFGRFDQLAHDNHCLRIKILGDCYYCVSGLPEPR ( 0) KDHAKCAVEMGLDMIDAIAT
4. (276) FHNLYVKRHTNVSILYADIVGFTRLASDC SPGELVHMLNELFGKFDQIAKENECMRIKILGDCYYCVSGLPISL ( 0) PNHAKNCVKMGLDMCEAIKK
5. (260) FHSLYVKRHQGVSVLYADIVGFTRLASEC SPKELVLMLNELFGKFDQIAKEHECMRIKILGDCYYCVSGLPLSL ( 0) PDHAINCVRMGLDMCRAIRK
6. (293) FHKIYIQKHDNVSILFADIEGFTSLASQC TAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEAR ( 0) ADHAHCCVEMGMDMIEAISS
7. (378) FHKIYIQKHDNVSILFADIEGFTSLASQC TAQELVMTLNELFARFDKLAAENHCLRIKILGDCYYCVSGLPEAR ( 0) ADHAHCCVEMGVDMIEAISL
8. (425) IVIPEPEEYKSCSILCFDIVQFTNMSAKLDSPSRLVDLLTQVFREFDTVVLRNGCQKIKTDGDAYICACGLKSKK (95) HFEKLIDVAIEIMNLDVLKE
consensus                 S L  DI  FT          LV  L   F  FD          IK  GD Y C  GL

1. VREKTKTGVDMRVGVHTGTVLGGVLGQKRWQYDVWSTDVTVANKMEAGGIPGRVHISQSTMDCLK ( 0) GEFDVEPGDGGSRCDYLDEKGIETYLIIASKPEVKK
2. VAEATEVDLNMRVGLHTGRVLCGVLGLRKWQYDVWSNDVTLANVMEAAGLPGKVHITKTTLACLN ( 0) GDYEVEPGHGHERNSFLKTHNIETFFIVPS HRRKI
3. VVEATDVILNMRVGIHTGRVLCGVLGLRKWQFDVWSNDVTLANHMESGGEPGRVHVTRATLDSLS ( 0) GEYEVEAGHGDERSSYLRDHGVDTFFIVPPPHRRKP
4. VRDATGVDINMRVGVHSGNVLCGVIGLQKWQYDVWSHDVTLANHMEAGGVPGRVHISSVTLEHLN ( 0) GAYKVEEGDGEIRDPYLKQHLVKTYFVINPKGERRS
5. LRVATGVDINMRVGVHSGSVLCGVIGLQKWQYDVWSHDVTLANHMEAGGVPGRVHITGATLALLA ( 0) GAYAVERADMEHRDPYLRELGEPTYLVIDPWAEEED
6. VREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGGKAGRIHITKATLNYLN ( 0) GDYEVEPGCGGERNAYLKEHSIETFLILRCTQKRKE
7. VREVTGVNVNMRVGIHSGRVHCGVLGLRKWQFDVWSNDVTLANHMEAGGRAGRIHITRATLQYLN ( 0) GDYEVEPGRGGERNAYLKEQCIETFLILGASQKRKE
8. TGNTEGIQVQFRCGIAAGSVYGGVIGSQKYQFDIWGDTIARSHTLEQLGQPGKVHVGETIMTHKN (21) HDYEFHKAHGECITSYFVDWKDDYREKKKKDLSCDF
consensus           R G   G V  GV G    Q D W          E  G  G  H

1. (407) KRDEELYSQSYDEIGVMFASLPNFADFYTEESINNGGIECLRFLNEIISDFDSLLDNPKFRVITKIKTIGSTYMAASGVTP (22) HLADLADFALAMKD
2. (370) PRNMDLYYQSYSQVGVMFASIPNFNDFYIELDGNNMGVECLRLLNEIIADFDELMDKDFYKDLEKIKTIGSTYMAAVGLAP (12) HLSTLADFAIEMFD
3. (491) RNNMELYHQSYAKVGVIFASVPNFNEFYTEMDGSDQGLECLRLLNEIIADFDELLKEDRFRGIDKIKTVGSTYMAVVGLIP (14) HMTALIEYVKAMRH
4. (400) LKNEELYHQSYDCVCVMFASIPDFKEFYTESDVNKEGLECLRLLNEIIADFDDLLSKPKFSGVEKIKTIGSTYMAATGLSA (15) HIGTMVEFAYALVG
5. (392) RRNEDLYHQSYECVCVLFASIPDFKEFYSESNINHEGLECLRLLNEIIADFDELLSKPKFSGVEKIKTIGSTYMAATGLNA (15) HLGTMVEFAVALGS
6. (405) RRNDELYYQSCECVAVMFASIANFSEFYVELEANNEGVECLRLLNEIIADFDEIISEDRFRQLEKIKTIGSTYMAASGLND ( 9) HIKALADFAMKLMD
7. (403) RRNDELYYQSCECVAVMFASIANFSEFYVELEANNEGVECLRLLNEIIADFDEIISEERFRQLEKIKTIGSTYMAASGLNA ( 9) HITALADYAMRLME
8. (437) ETKGIVYVQPHQDVSIMFIQIAGFQEY       DEPKDLIKKLNDIFSFFDGLLNQKYGGTVEKIKTIGNTYMAVSGLDG ( 0) SPSFLEKMSDFALD
consensus             Y Q       F    F                  LN I   FD          KIKT G TYMA  G

1. TLTNINNQSFNNFMLRIGMNKGGVLAGVIGARKPHYDIWGNTVNVASRMESTGVMGNIQVVEETQVILREYGFRFVRRGPIFVKGKGELLTFFLKGR (  22)
2. VLDEINYQSYNDFVLRVGINVGPVVAGVIGARRPQYDIWGNTVNVASRMDSTGVQGRIQVTEEVHRLLRRGSYRFVCRGKVSVKGKGEMLTYFLEGR (  74)
3. SLQEINSHSYNNFMLRVGINIGPVVAGVIGARKPQYDIWGNTVNVASRMDSTGVPGYSQVTQEVVDSLVGSHFEFRCRGTIKVKGKGDMVTYFLCDS (1094)
4. KLDAINKHSFNDFKLRVGINHGPVIAGVIGAQKPQYDIWGNTVNVASRMDSTGVLDKIQVTEETSLILQTLGYTCTCRGIINVKGKGDLKTYFVNTE (  12)
5. KLGVINKHSFNNFRLRVGLNHGPVVAGVIGAQKPQYDIWGNTVNVASRMESTGVLGKIQVTEETARALQSLGYTCYSRGVIKVKGKGQLCTYFLNTD (  10)
6. QMKYINEHSFNNFQMKIGLNIGPVVAGVIGARKPQYDIWGNTVNVASRMDSTGVPDRIQVTTDMYQVLAANTYQLECRGVVKVKGKGEMMTYFLNGG (   4)
7. QMKHINEHSFNNFQMKIGLNMGPVVAGVIGARKPQYDIWGNTVNVSSRMDSTGVPDRIQVTTDLYQVLAAKGYQLECRGVVKVKGKGEMTTYFLNGG (   3)
8. VKAYTNSVAI SRVVRIGISHGPLVAGCIGISRAKFDVWGDTANTASRMQSNAQDNEIMVTHSVYERLNKLFYFDDEKEIL VKGKGKMVTHVLKGK (  50)
consensus     N          G   G   AG IG      D WG T N  SRM S       V      L             VKGKG   T
```

Appendix C2

Appendix D1

Expressions and Accession Numbers of Phosphodiesterases

No.	Gene	Expression	Accession No.
1.	human cAMP PDE (PDE2)	brain, heart, kidney	M37744
		testis[a]	
2.	rat cAMP PDE (PDE4)	brain, heart, liver,	J04563
		kidney, testis	
3.	rat cAMP PDE (PDE3)	brain, heart, liver,	M25349
		kidney, testis	
4.	rat cAMP PDE (PDE1)	testis, kidney	M25347
5.	*Drosophila* cAMP PDE (*dunce*)	NI	X55167
6.	rat calmodulin-dependent PDE (CaM-PDE)	NI	M94537
7.	bovine calmodulin-dependent PDE 61-kDa (CaM-PDE) (partial)	NI	A40282[b]
8.	human cGMP-inhibited cAMP PDE (cGI PDE)	heart	M91667
9.	mouse cGMP PDE α	rod cell	X60664
10.	mouse cGMP PDE β	rod cell	X55968
11.	bovine cGMP PDE	cone cell	M37838
12.	bovine cGMP-stimulated cNMP PDE (cGS PDE)	ubiquitious	M73512
13.	*Saccharomyces* cAMP PDE (PDE2)	NI	M14563

Note. - Sequence data are from GenBank release 77.0 and NBRF release 36.0.
NI = not identified.

a Rat expression data from Swinnen et. al. (1989).

b Sequence data from NBRF release 36.0.

References

Swinnen, J. V., Joseph, D. R., and Conti, M. (1989). Molecular cloning of rat homologues of the *Drosophila melanogaster* dunce cAMP phosphodiesterase: evidence for a family of genes. Proc. Natl. Acad. Sci. USA *86*, 5325-5329.


```
PDE

 1. (228) DVAYHNSLHAADVLQSTHVLLATPALDAVFTDLEILAALFAAAIHDVDHPGVSNQFLINTNSELALMY NDESVLEN HHLAVG FKLLQEYNC
 2. (227) DVAYHNSLHAADVAQSTHVLLSTPALDAVFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELALMY NDESVLEN HHLAVG FKLLQEEHC
 3. (233) DVAYHNNIHAADVVQSTHVLLSTPALEAVFTDLEILAAIFASAIHDVDHPGVSNQFLINTNSELALMY NDSSVLEN HHLAVG FKLLQEENC
 4. ( 97) NVAYHNSIHAADVVQSAHVLLGTPALEAVFTDLEVLAAIFACAIHDVDHPGVSNQFLINTNSELALMY NDSSVLEN HHLAVG FKLLQGENC
 5. (242) DNPFHNSLHAADVTQSTNVLLNTPALEGVFTPLEVGGALFAACIHDVDHPGLTNQFLVNSSSELALMY NDESVLEN HHLAVA FKLLQNQGC
 6. (217) KNPYHNQIHAADVTQTVHCFLLRTGMVHCLSEIEVLAIIFAAAIHDYEHTGTTNSFHIQTKSECAILY NDRSVLEN HHISSV FRMMQDDEM
 7. (213) KNPYHNLIHAADVTQTVHYIMLHTGIMHWLTELEILAMVFAAAIHDYEHTGTTNNFHIQTRSDVAILY NDRSVLEN HHVSAA YRLMQEEEM
 8. (791) SGFTHGHMGYVFSKTYNVTDDKYGCLSGNIPALELMALYVAAAMHDYDHPGRTNAFLVATSAPQAVLY NDRSVLEN HHAAAAWNLFMSRPEY
 9. (554) RITYHNWRHGFNVGQTMFSLLVTGKLKRYFTDLEALAMVTAAFCHDIDHRGTNNLYQMKSQNPLAKL  HGSSILER HHLEFG KTLLRDESL
10. (552) RITYHNWRHGFNVAQTMFTLLMTGKLKSYYTDLEAFAMVTAGLCHDIDHRGTNNLYQMKSQNPLAKL  HGSSILER HHLEFG KFLLAEESL
11. (552) AVTYHNWRHGFNVGQTMFTLLMTGRLKKYYTDLEAFAMLAAAFCHDIDHRGTNNLYQMKSTSPLARL  HGSSILER HHLEYS KTLLQDESL
12. (631) DPPYHNWMHAFSVSHFCYLLYKNLELTNYLEDMEIFALFISCMCHDLDHRGTNNSFQVASKSVLAALYSSEGSVMER HHFAQA IAILNTHGC
13. (260) VNKFHNFRHAIDVMQATWRLCTYLLKDN   PVQTLLLCMAAIGHDVGHPGTNNQLLCNCESEVAQNF KNVSILENFHRELFQ  QLLSEH W
consensus          H                                          HD  H G  N         A       S  E  H

 1. DIFQNLSKRQRQSLRKMVIDMVLATDMSKHMTLLADL (25) VLRNMVHCADLSNPTKPLELYRQWTDRIMAEFFQQGDRERE RGMEISPMC (255)
 2. DIFQNLTKKQRQTLRKMVIDMVLATDMSKHMSLLADL (25) VLRNMVHCADLSNPTKSLELYRQWTDRIMEEFFQQGDKERE RGMEISPMC (132)
 3. DIFQNLTKKQRQSLRKMAIDIVLATDMSKHMNLLADL (25) VLQNMVHCADLSNPTKPLQLYRQWTDRIMEEFFRQGDRERE RGMEISPMC (148)
 4. DIFQNLSTKQKLSLRRMVIDMVLATDMSKHMSLLADL (25) VLQSLVHCADLSNPAKPLPLYRQWTERIMAEFFQQGDRERE SGLDISPMC ( 59)
 5. DIFCNMQKKQRQTLRKMVIDIVLSTDMSKHMSLLADL (25) VLENLVHCADLSNPTKPLPLYKRWVALLMEEFFLQGDKERE SGMDISPMC (139)
 6. NIFINLTKDEFVELRALVIEMVLATDMSCHFQQVKTM (14) ALSLLLHAADISHPTKQWSVHSRWTKALMEEFFRQGDKEAE LGLPFSPLC (126)
 7. NVLINLSKDDWRDLRNLVIEMVLSTDMSGHFQQIKNI (14) TMSLILHAADISHPAKSWKLHHRWTMALMEEFFLQGDKEAE LGLPFSPLC (124)
 8. NFLINLDHVEFKHFRFLVIEAILATDLKKHFDFVAKF (20) VCQMCIKLADINGPAKCKELHLQWTDGIVNEFYEQGDEEAS LGLPISPFM (151)
 9. NIFQNLNRRQHEHAIHMMDIAIIATDLALYFKKRTMF (29) VMAMMMTACDLSAITKPWEVQSKVALLVAAEFWEQGDLERTVLQQNPIPMM ( 98)
10. NIYQNLNRRQHEHVIHLMDIAIIATDLALYFKKRTMF (29) VMAMMMTACDLSAITKPWEVQSKVALLVAAEFWEQGDLERTVLDQQPIPMM ( 97)
11. NIFQNLNKRQYETVIHLFEVAIIATDLALYFKKRTMF (29) IMAMMMTACDLSAITKPWEVQSQVALLVANEFWEQGDLERTVLQQQPIPMM ( 96)
12. NIFDHFSRKDYQRMLDLMRDIILATDLAHHLRIFKDL (18) LLCLLMTSCDLSDQTKGWKTTRKIAELIYKEFFSQGDLEKA MGNRPMEMM ( 93)
13. PLKLSISKKKFD    FISEAILATDMALHSQYEDRL (10) LISLIIKAADISNVTRTLSISARWAYLITLEFNDCALLETFHKAHRPEQDC ( 85)
consensus                    TD                    D                    EF      E
```

Appendix D2

Appendix E1

Expressions and Accession Numbers of Phospholipase Cs

No.	Gene	Expression	Accession No.
(A) Phosopholipase C γ (Activated by Protein Tyrosine Kinases)			
1.	human PLC-γ1	NI	M34667
2.	human PLC-γ2	NI	X14034;M37238
(B) Phosopholipase C δ (Activated by Unknown Factors)			
3.	rat PLC-δ1	adrenal > brain	M20637
4.	bovine PLC-δ2	brain	S14113[a]
(C) Phosopholipase C β (Activated by G-proteins)			
5.	rat PLC β3	NI	M99567
6.	rat PLC-β1	NI	M20636
7.	human PLC-β2	NI	M95678
8.	*Drosophila plc-21*	adult and larval brain	M60452
9.	bovine retina PLC-β	retina	L13935
10.	*Drosophila norpA*	eye	J03138
(D) *Dictyostelium* and *Saccharomyces* Phosopholipase C			
11.	*Saccharomyces PLC1*	NI	D12738
12.	*Dictyostelium* DdPLC	all stages of development	M95783

Note. - Sequence data are from GenBank release 77.0 and NBRF release 36.0. > indicates that tissue expressions of the right are weaker than those of the left. NI = not identified.

[a] Sequence data from NBRF release 36.0.

PLC

```
 1. (322) NNPLSHYWISSSHNTYLTGDQFSSESSLEAYARCLRMGCRCIELDCWDG (4) PVIYHGHTLTTKIKFSDVLHTIKEHAFVASEYPVILSIEDHCSIA QQRN
 2. (314) NNPLSHYWISSSHNTYLTGDQLRSESSPEAYIRCLRMGCRCIELDCWDG (4) PVIYHGWTRTTKIKFDDVVQAIKDHAFVTSSFPVILSIEEHCSVE QQRH
 3. (298) DQPLSHYLVSSSHNTYLLEDQLTGPSSTEAYIRALCKGCRCLELDCWDG (4) PIIYHGYTFTSKILFCDVLRAIRDYAFKASPYPVILSLENHCSLE QQRV
 4. (292) TQPLNHYYINSSHNTYLVGDQLCGQSSVEGYIRALKRGCRCVEVDIWDG (4) PIVYHGHTLTSRIPFKDVVAAIGQYAFQTSDYPVILSLENHCSWE QQEI
 5. (259) TQPLSAYFINSSHNTYLTAGQLAGTSSVEMYRQALLWGCRCVELDVWKG (6) PFITHGFTMTTEVPLRDVLEAIAETAFKTSPYPVILSFENHVDSAKQQAK
 6. (314) TQPLNHYFINSSHNTYLTAGQFSGLSSAEMYRQVLLSGCRCVELDCWKG (6) PIITHGFTMTTDIFFKEAIEAIAESAFKTSPYPIILSFENHVDSPRQQAK
 7. (318) SQPLSHYFINSSHNTYLTAGQLAGNSSVEMYRQVLLSGCRCVELDCWKG (6) PVITHGFTMTTEISFKEVIEAIAECAFKTSPFPILLSFENHVDSPKQQAK
 8. (320) DQPMSHYFINSSHNTYLTGHQLTGKSSVEIYRQCLLAGCRCVELDFWNG (4) PVIVHGYTFVPEIFAKDVLEAIAESAFKTSEYPVILSFENHCNPR QQAK
 9. (151) DHPLAHYFISSSHNTYLTGRQFGGKSSVEMYRQVLLAGCRCVELDCWDG (6) PIITHGKAMCTDILFKDVIQAIKETAFVTSEYPVILSFENHCSKY QQYK
10. (321) DQPLAHYYINSSHNTYLSGRQIGGKSSVEMYRQTLLAGCRCVELDCWNG (6) PIVTHGHAYCTEILFKDCIQAIADCAFVSSEYPVILSFENHCNRA QQYK
11. (324) SKPLSYYFINSSHNTYLSGHQLKGLSTSEMYTNTLRQGCKCVELDVWDG (4) PIIFHGNTLTSQIKFSHVCETIKARGFETSPYPVILSLEVHCSVP QQIM
12. (382) SKPLNHYFIASSHNTYLLGKQIAETPSVEGYIQVLQQGCRCVEIDIWDG (3) PVVCHG FLTSAIPLKTVIRVIKKYAFITSPYPLIISLEINCNKD NQKL
consensus        P   Y   SSHNTYL   Q      E Y   L  GC C E D W G    P   HG              I    F  S  P   S E       Q
```

```
 1. MAQYFKKVLGDTL ( 8) ADGLPSPNQLKRKILIKH (544) LSRIYPKGQRLDSSNYDPLPMWICGSQLVALNFQTPDKPMQMNQALF ( 5) CGYVLQPSTM (22)
 2. MAKAFKEVFGDLL ( 8) ADQLPSPSQLREKIIIKH (526) LTRVYPKGQRVDSSNYDPFRLWLCGSQMVALNFQTADKYMQMNHALF ( 5) TGYVLQPESM (21)
 3. MARHLRAILGPIL ( 8) TTSLPSPEQLKGKILLKG (107) LSRIYPAGWRTDSSNYSPVEMWNGGCQIVALNFQTPGPEMDVYLGCF ( 5) CGYVLKPAFL (24)
 4. IVRHLTEILGDQL ( 9) PTQLPSPEDLRGKILVKG (109) LSRVYPSGLRTDSSNYNPQEFWNAGCQMVAMNMQTAGLEMDLCDGLF ( 5) CGYVLKPDFL (27)
 5. MAEYCRSIFGEAL (12) GTPLPSPQDLMGRILVKN (177) LSRIYPKGTRVDSSNYMPQLFWNVGCQLVALNFQTLDLPMQLNAGVF ( 5) SGYLLKPEFM (24)
 6. MAEYCRTIFGDML (12) GVPLPSPEDLRGKILIKN (133) MSRIYPKGTRMDSSNYMPQMFWNAGCQMVALNFQTMDLPMQQNMAVF ( 5) SGYLLKHEFM (24)
 7. MAEYCRLIFGDAL (12) GVPLPSPMDLMYKILVKN (127) LSRIYPKGTRVDSSNYMPQLFWNAGCQMVALNFQTVDLAMQINMGMY ( 5) SGYRLKPEFM (24)
 8. IANYCREIFGDML (12) NMDLPPPAMLRRKIIIKN (187) LSRVYPAGTRFDSSNFMPQLFWNAGCQLVALNFQTLDLAMQLNLGIF ( 5) SGYLLKPEFM (24)
 9. MSKYCEDLFGDLL (12) GRPLPSPNDLKRKILIKK (156) MSRIYPKGGRVDSSNYMPQIFWNSGCQMVSLNYQTPDLAMQLNQGKF ( 5) CGYLLKPDFM (24)
10. LAKYCDDFFGDLL (12) GLPLPPPCKLKRKILIKN (135) MSRIYPKGTRADSSNYMPQVFWNAGCQMVSLNFQSSDLPMQLNQGKF ( 5) CGYLLKPDFM (24)
11. MANHMKEIFGEML ( 7) TKELPTLDSLKYKILLKG (126) LLRVYPRGTRFDSSNFDPMPGWSIGCQLAALNQQTSSEPMWINDGMF ( 5) CGYVLKPPCL (25)
12. ASLIMREVLAEQL ( 6) TDKLPSPRELKHKILLKS (121) LMRVYPHVLRYKSSNFNPIPFWKAGVQMVATNWQTNDIGQQLNLAMF (12) SGYVLKPKKL (27)
consensus             L         LP    L   I  K          R YP   R  SSN  P   W  G Q    N Q                    GY L
```

```
 1. IEVLGARHLP ( 7) CPFVEIEV AGAEYDSTKQ  KTEFV VDNGLNPVWPAK (5) ISNPEFAFLRFVV (11) LAQATFPVKGLKTGYRAVPL ( 98)
 2. VKVLGARHLP ( 7) CPFVEVEI CGAEYGNNKF  KTTVV NDNGLSPIWAPT (7) IYDPNLAFLRFVV (11) LAHATYPIKAVKSGFRSVPL ( 85)
 3. VRIISGQQLP (10) DPKVIVEI HGVGRDTGSR  QTAVI TNNGFNPRWDME (4) VTVPDLALVRFMV (11) IGQSTIPWNSLKQGYRHVHL ( 21)
 4. LQVISGQQLP (11) DPLVRVEI FGVRPDTTRQ  ETSYV ENNGFNPYWGQT (4) ILVPELALLRFVV (11) IGQYTLPWSCMQQGYRHIHL ( 28)
 5. VKVISGQFLS ( 4) GIYVEVDM FGLPVDTRRK YRTRTS QGNSFNPVWDEE (6) VVLPTLASLRIAA ( 7) VGHRILPVSAIRSGYHYVCL (410)
 6. ITVISGQFLS ( 4) RTYVEVEL FGLPGDPKRR YRTKLSPSTNSINPVWKEE (6) ILMPELASLRVAV ( 7) LGHRIIPINALNSGYHHLCL (403)
 7. VKIISGQFLS ( 4) GTYVEVDM FGLPVDTRRKAFKTKTS QGNAVNPVWEEE (6) VVLPSLACLRIAA ( 7) IGHRILPVQAIRPGYHYICL (440)
 8. ITVLSGQFLT ( 4) NTFVEVDM YGLPADTVRKKFRTKTV RDNGMNPLYDEE (6) VVLPELASIRIAA ( 7) IGHRVLPVIGLCPGYRHVNL (470)
 9. VQVISGQFLS ( 4) GTYVEVDM YGLPTDTIRKEFRTRMV MNNGLNPVYNEE (6) VILPDLAVLRIAV ( 7) IGQRILPLDGLQAGYRHISL (374)
10. VKVIAGQFLS ( 4) GTYVEVDM FGLPSDTVKKEFRTRLV ANNGLNPVYNED (6) VVLPDLAVLRFGV ( 7) LGQRILPLDGLQAGYRHVSL (309)
11. VNVISARQLP (11) DPYVTLSI VGTHFDQKVE  KTKVI DNNGFNPHWGEE (4) LYNSQLSMLLIRV (10) IGHHCIRVENIRPGYRILKL ( 21)
12. IRILSTQLLP (24) EPTMPISIDKGTRISATEA  STKSS QGNGFNPIWDAE (4) LKDTDLTFIKFMV ( 6) IASVCLKLNYLRMGYRHIPL ( 20)
consensus          L                G          T     N   P                                                G     L
```

Appendix E2