G Protein Diversity

Thesis by

Michael Paul Strathmann

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1991

(Defended March 12, 1991)

The flow of information from hormone receptors through heterotrimeric G proteins to intracellular effectors constitutes a basic form of signal transduction that has been found in every eukaryotic cell examined. This mode of signaling is the basis for intercellular communication in one-celled organisms, such as *S. cerevisiae*, and simple eukaryotes that undergo limited development, e.g., *Dictyostelium*. We examined G protein diversity in the mouse to explore how G protein-mediated signal transduction has adapted to the complex signaling processes that define a multicellular organism. We found that the diversity of G protein alpha subunits is generated by a large number of distinct genes and by products of alternative splicing. Five new genes were found to encode alpha subunits that fall into two new classes. These classes, which are defined by amino acid sequence identity, have been conserved in distantly related animals; the four classes known to exist in mammals are also present in *Drosophila* , and three of these classes have been found in nematodes. The gene encoding $G\alpha_o$ in mammals, an alpha subunit earlier characterized by biochemical means, was found to undergo alternative splicing to produce transcripts encoding two forms of the protein.

In addition to the diversity among alpha subunits, we found a novel beta subunit. This was particularly surprising because biochemical evidence suggested that beta-gamma dimers are interchangeable. The possibility of combinatorial associations of alpha, beta, and gamma subunits to produce functionally distinct heterotrimers must be considered.

The considerable task of analyzing gene families by cloning and sequencing necessitated the development of new techniques for these purposes. We used the polymerase chain reaction to amplify cDNA with degenerate oligonucleotide primers. The primers were designed to hybridize to regions of DNA that are conserved in all members of the gene family. In addition, we developed a technique for randomly inserting sequencing primer sites throughout a cloned region of DNA. For this

purpose, we used the transposon γδ coupled with an efficient scheme for isolating and characterizing the insertions.

G PROTEIN DIVERSITY

Introduction: Section I

G protein-mediated signal transduction: A history of the hormone-stimulated adenylate cyclase system

The field of signal transduction has been shaped by studies of hormone-stimulated glycogen breakdown in liver cells. Sutherland and colleagues discovered a heat-stable, soluble factor which could stimulate phosphorylase activity in cell extracts. This factor was produced upon addition of epinephrine or glucagon to the particulate fraction of liver homogenates (1). An identical substance, later shown to be cyclic AMP (2), was produced by particulate preparations from a variety of tissues (3).

These studies led to the concept of signal transduction by second messengers. Certain hormones, the first messengers, stimulate the cell-membrane associated protein, adenylate cyclase (4), which produces the intracellular second messenger cAMP. The change in second-messenger levels then influences one or more cellular processes (5).

Studies on the stimulation of cAMP synthesis in adipose cells by a variety of hormones suggested that different receptors are free to interact with a common pool of adenylate cyclase enzymes (6, 7). Maximal stimulation by two or more hormones (epinephrine, ACTH, and glucagon) were never additive, yet the specific antagonistic effect of pronethalol on epinephrine suggested that the receptors were separate (6). The cell fusion studies of Orly and Schramm demonstrated that receptors from one source can effectively couple to adenylate cyclase from another source (8).

The study of glucagon effects in liver extracts again resulted in a critical observation concerning the mode of signal transduction in eukaryotic cells. Rodbell and coworkers found that both GTP and GDP reduced the affinity of membrane receptors for glucagon (9). The nucleotides did not act competitively with the hormone suggesting that they had separate binding sites. In addition, guanine nucleotides were shown to be essential for hormone-stimulated adenylate cyclase activity (10). These studies also indicated that guanine nucleotide and fluoride ion might affect adenylate cyclase by a related mechanism (10). Fluoride ion had been shown many years earlier

to stimulate adenylate cyclase (11). Interestingly, this ion was originally used by Sutherland to inhibit a phosphatase in assays for hormone stimulation of phosphorylase. The realization that the assays worked only in the absence of fluoride ultimately led to the discovery of the ion's effect on adenylate cyclase (12).

After the initial observation by Rodbell, the stimulation of adenylate cyclase by several different hormones including prostaglandins (13), TSH (14), oxytocin (15), and epinephrine (16) was shown to be influenced by guanine nucleotides. Maguire et al. found that the decrease in the affinity of receptors for hormones in the presence of guanine nucleotides was specific for agonists; the binding of antagonists was not affected (17). Foreshadowing the discovery of other guanine-nucleotide modulated second-messenger systems, the affinity of angiotensin II for its receptor, which does not couple to adenylate cyclase, was shown to be affected by GTP (18). Several years after the discovery that guanine nucleotides modulate receptor affinity for hormone came the realization that, in fact, hormones modulate the effects of GTP (19).

The observation that poorly hydrolyzable analogs of GTP (Gpp(NH)p, GTPγS, Gpp(CH_2)p) activate adenylate cyclase, often in the absence of hormone, yet compete with GTP in their actions, led several groups to postulate that a GTPase is involved in the adenylate cyclase system (20, 21, 22). It was the work of Cassel and Selinger, however, that established the central role of guanine nucleotide exchange and hydrolysis in the modulation of adenylate cyclase activity. Working with turkey erythrocyte membranes, they found a catecholamine-stimulated GTPase (23). The actions of agonists on the GTPase could be inhibited by the beta-adrenergic antagonist propranolol. The addition of cAMP or the inactivation of the endogenous adenylate cyclase with N-ethylmaleimide did not affect the GTPase. Also, the affinity and specificity for GTP exhibited by the GTPase were similar to those values reported for catecholamine-stimulated adenylate cyclase.

The stimulatory effects of Gpp(NH)p on adenylate cyclase were known to be stable to removal of hormone and excess nucleotide in avian erythrocyte membranes (21, 24). Cassel and Selinger showed that, in the presence of isoproterenol, intact Gpp(NH)p was released from membranes that were washed free of excess nucleotide (25). The release of nucleotide was proportional to the decline in adenylate cyclase activity. They proposed a model in which hormone produces a change in the guanine-nucleotide binding site, leading to an equilibrium between guanine nucleotide at the binding site and in the medium. They concluded that Gpp(NH)p persistently activates adenylate cyclase because it cannot be hydrolyzed; hence under physiological conditions, the GTPase activity serves to hydrolyze GTP and return the system to its basal state. This model was supported by further studies that showed the displacement of GDP by GTP in the presence of hormone and its subsequent hydrolysis to GDP (26). The significance of the GTPase activity as a "shut off" mechanism is exemplified by the effect of cholera toxin on adenylate cyclase activity. Cholera toxin was known to produce its pathological effects through the stimulation of adenylate cyclase (27, 28). Cassel and Selinger showed that the toxin reduces the catecholamine-stimulated GTPase activity (29).

In turkey erythrocyte membranes, the number of guanine-nucleotide binding sites that were coupled to the adenylate cyclase system was estimated to be equal to the number of beta-adrenergic receptors and to the number of adenylate cyclase catalytic components (25). Tokovsky and Levitzki, however, demonstrated that receptors act catalytically rather than stoichiometrically (30). They showed that up to a 10-fold loss of beta-adrenergic receptors, due to alkylation, slowed proportionally the rate of activation of adenylate cyclase by Gpp(NH)p, but the maximum rate of cAMP synthesis did not change.

Since the initial discovery of guanine-nucleotide effects on adenylate cyclase activity, it had been hypothesized that the protein responsible for these regulatory effects could be distinct from the receptor and catalytic components. The existence of a separate, regulatory protein was demonstrated by two groups.

Pfeuffer was able to label pigeon erythrocyte membranes with a photoactivatable GTP analog, which was a potent activator of adenylate cyclase (31). In solubilized membranes, only two proteins (42kD and 23kD) were labeled. 95% of the nucleotide binding sites could be separated form adenylate cyclase activity by centrifugation of the solubilized membranes through a sucrose density gradient. The majority of the 42kD protein remained associated with adenylate cyclase activity. Pfeuffer made use of a GTP-Sepharose affinity matrix to separate solubilized membranes into two fractions. The bound fraction, which was released from the matrix in the presence of GTP or Gpp(NH)p, retained only 3% Gpp(NH)p stimulated adenylate cyclase activity, whereas the unadsorbed fraction retained 15% activity. Recombination of the two fractions restored 60% of the control adenylate cyclase activity. Pfeuffer's work, however, did not show that the guanine-nucleotide binding sites were necessarily separable from the catalytic activity.

Ross and Gilman (32) made use of somatic cell variants to demonstrate that Gpp(NH)p-stimulated adenylate cyclase activity requires at least two components with different thermal stabilities. They found that a detergent extract of wild-type membranes, which lost hormone-stimulated adenylate cyclase activity upon solubilization, completely lost Gpp(NH)p- and fluoride-stimulated activity upon heating at 37°C. This extract, however, could restore activity to solubilized membranes from a phenotypically adenylate cyclase deficient (cyc-) cell line. The cyc- cells had been shown to possess beta-adrenergic receptors (33). The factor(s) contributed by the cyc- cells was labile to heating at 30°C whereas the factor(s) from

the heated wild-type extract showed complicated inactivation kinetics at 50°C. Ross et al. (34) showed that cyc- cells possess Mn^{2+}-dependent adenylate cyclase activity. The protein responsible for this activity had hydrodynamic properties identical to the Gpp(NH)p-stimulated adenylate cyclase. These two proteins shared other properties including thermolability. Ross et al. concluded that the component supplied by the cyc- cells in the reconstitution experiment was the catalytic protein, whereas the heat-stable factor(s) was the regulatory element.

Gilman's group subsequently purified the regulatory (G/F) proteins from rabbit liver (35, 36) and turkey erythrocytes (37). The ability to reconstitute fluoride-stimulated adenylate cyclase in the cyc- membranes was used to measure activity. The protein from rabbit liver consisted of three polypeptides with molecular weights of 52, 45, and 35kD at a stoichiometric ration of 1:1 (52kD+45kD:35kD). The form from turkey erythrocytes contained only the 45 and 35kD polypeptides. Hydrodynamic measurements gave a molecular weight of 70kD, but activating agents (fluoride, GTPγS, or Gpp(NH)p) produced an apparent dissociation of subunits to produce a 50kD protein that could activate adenylate cyclase (35-38). In addition, cholera toxin (35) and photoactivatable GTP (39) labeled the 52 and 45kD polypeptides.

Separation of the G/F component into alpha (52 and 45kD polypeptides) and beta (35kD) subunits was achieved by Northup et al. The alpha subunit was shown to be sufficient to activate adenylate cyclase in a preparation of the catalyst that was resolved from other components of the enzyme system (40). Beta subunits were found to increase the initial rate of deactivation of fluoride-activated G/F alpha subunits by 160-fold (41). In addition, beta subunits suppressed the initial rate of activation by GTPγS (39). These properties led Northup et al. (39, 40) to conclude that G/F is an alpha-beta dimer that dissociates upon activation; the alpha subunits stimulate adenylate cyclase, whereas the beta subunits have a repressive role. Eventually, the beta subunit

was recognized to be tightly associated with a small polypeptide (6-10kD), the gamma subunit (42, 43).

Purification of the regulatory component, beta-adrenergic receptors (44, 45) and the catalytic component of adenylate cyclase (46-49), coupled with efforts to reconstitute hormone-stimulated activity in phospholipid vesicles (50-52) enabled investigators to verify in a defined system the GTP-regulatory cycle that was established by the key experiments cited above. This cycle, which is driven by hormone-receptor catalyzed, guanine-nucleotide exchange and subsequent GTP hydrolysis, is central to a wide variety of signal tranducing pathways. In fact, the observation that phototransduction is analogous in many ways to hormone-stimulated adenylate cyclase (53, 54), hinted at the enormous diversity of "G protein-mediated" signaling systems which we are beginning to uncover.

References

1. Rall, T. W., Sutherland, E. W. and Berthet, J. (1957) *J. Biol. Chem.* **224**, 463-475.

2. Lipkin, D., Cool, W. H. and Markham, R. (1959) *J. Amer. Chem.* Soc. **81**, 6198-6203.

3. Rall, T. W. and Sutherland, E. W. (1958) *J. Biol. Chem.* **232**, 1065-1076.

4. Davoren, P. R. and Sutherland, E. W. (1963) *J. Biol. Chem.* **238**, 3016-3023.

5. Sutherland, E. W. and Robison, G. A. (1966) *Pharmacol. Rev.* **18**, 145-161.

6. Butcher, R. W., Baird, C. E. and Sutherland, E. W. (1968) *J. Biol. Chem.* **243**, 1705-1712.

7. Birnbaumer, L. and Rodbell, M. (1969) *J. Biol. Chem.* **244**, 3477-3482.

8. Orly, J. and Schramm, M. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4410-4414.

9. Rodbell, M., Krans, M. J., Pohl, S. L. and Birnbaumer, L. (1971) *J. Biol. Chem.* **246**, 1872-1876.

10. Rodbell, M., Birnbaumer, L., Pohl, S. L. and Krans, M. J. (1971) *J. Biol. Chem.* **246**, 1877-1882.

11. Sutherland, E. W., Rall, T. W. and Menon, T. (1962) *J. Biol. Chem.* **237**, 1220-1227.

12. Sutherland, E. W. (1972) *Science* **177**, 401-408.

13. Krishna, G., Harwood, J. P., Barber, A. J. and Jamieson, G.A. (1972) *J. Biol. Chem.* **247**, 2253-2254.

14. Wolff, J. and Cook, G. H. (1973) *J. Biol. Chem.* **248**, 350-355.

15. Bockaert, J., Roy, C. and Jard, S. (1972) *J. Biol. Chem.* **247**, 7073-7081.

16. Leray, F., Chambaut, A. M. and Hanoune, J. (1972) *Biochem. Biophys. Res. Commun.* **48**, 1385-1391.

17. Maguire, M. E., Van Arsdale, P. M. and Gilman, A. G. (1976) *Mol. Pharmacol.* **12**, 335-339.

18. Glossmann, H., Baukal, A. J. and Catt, K. J. (1974) *J. Biol. Chem.* **249**, 825-834.

19. Rodbell, M., Lin, M. C. and Salomon, Y. (1974) *J. Biol. Chem.* **249**, 59-65.

20. Rodbell, M., Lin, M. C., Salomon, Y., Londos, C., Harwood, J. P., Martin, B. R., Rendell, M. and Barman, M. (1975) in *Advances in Cyclic Nucleotide Research,* eds. Drummond, G. I., Greengard, P., and Robison, G. A. (Elsevier, Amsterdam) Vol. 5, pp. 3-29.

21. Pfeuffer, T. and Helmreich, E. J. M. (1975) *J. Biol. Chem.* **250**, 867-876.

22. Blume, A. J. and Foster, C. J. (1976) *J. Biol. Chem.* **251**, 3399-3404.

23. Cassel, D. and Selinger, Z. (1976) *Biochim. Biophys. Acta* **452**, 538-551.

24. Schramm, M. and Rodbell, M. (1975) *J. Biol. Chem.* **250**, 2232-2237.

25. Cassel, D. and Selinger, Z. (1977) *J. Cyclic Nucleotide* Res. **3**, 11-22.

26. Cassel, D. and Selinger, Z. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4155-4159.

27. Kimberg, D. V., Field, M., Johnson, J., Henderson, A. and Gershon, E. (1971) *J. Clin. Invest.* **50**, 1218-1230.

28. Sharp, G. W. G. and Hynie, S. (1971) *Nature* **229**, 266-269.

29. Cassel, D. and Selinger, Z. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3307-3311.

30. Tolkovsky, A. M. and Levitski, A. (1978) *Biochemistry* **17**, 3795-3810.

31. Pfeuffer, T. (1977) *J. Biol. Chem.* **252**, 7224-7234.

32. Ross, E. M. and Gilman, A. G. (1977) *J. Biol. Chem.* **252**, 6966-6969.

33. Insel, P. A., Maguire, M. E., Gilman, A. G., Bourne, H. R., Coffino, P. and Melmon, K. L. (1976) *Mol. Pharmacol.* **12**, 1062-1069.

34. Ross, E. M., Howlett, A. C., Ferguson, K. M. and Gilman, A. G. (1978) *J. Biol. Chem.* **253**, 6401-6412.

35. Northup, J. K., Sternweis, P. C., Smigel, M. D., Schleifer, L. S., Ross, E. M. and Gilman, A. G. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6516-6520.

36. Sternweis, P. C., Northup, J. K., Smigel, M. D. and Gilman, A. G. (1981) *J. Biol. Chem.* **256**, 11517-11526.

37. Hanski, E., Sternweis, P. C., Northup, J. K., Dromerick, A. W. and Gilman, A. G. (1981) *J. Biol. Chem.* **256**, 12911-12919.

38. Howlett, A. C. and Gilman, A. G. (1980) *J. Biol. Chem.* **255**, 2861-2866.

39. Northup, J. K., Smigel, M. D. and Gilman, A. G. (1982) *J. Biol. Chem.* **257**, 11416-11423.

40. Northup, J. K., Smigel, M. D., Sternweis, P. C. and Gilman, A. G. (1983) *J. Biol. Chem.* **258**, 11369-11376.

41. Northup, J. K., Sternweis, P. C. and Gilman, A. G. (1983) *J. Biol. Chem.* **258**, 11361-11368.

42. Hildebrandt, J. D., Codina, J., Risinger, R. and Birnbaumer, L. (1984) *J. Biol. Chem.* **259**, 2039-2042.

43. Bokoch, G. M., Katada, T., Northup, J. K., Ui, M. and Gilman, A. G. (1984) *J. Biol. Chem.* **259**, 3560-3567.

44. Shorr, R. G. L., Strohsacker, M. W., Lavin, T. N., Lefkowitz, R. J. and Caron, M. G. (1982) *J. Biol. Chem.* **257**, 12341-12350.

45. Homcy, C. J., Rockson, S. G., Countaway, J. and Egan, D.A. (1983) *Biochemistry* **22**, 660-668.

46. Pfeuffer, E., Drehev, R.-M., Metzger, H. and Pfeuffer, T. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3086-3090.

47. Smigel, M. D. (1986) *J. Biol. Chem.* **261**, 1976-1982.

48. Coussen, F., Haiech, J., D'Alayer, J. and Monneron, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6736-6740.

49. Yeager, R. E., Heideman. W., Rosenberg, G. B. and Storm, D. R. (1985) *Biochemistry* **24**, 3776-3783.

50. Pedersen, S. E. and Ross, E. M. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7228-7232.

51. Brandt, B. R., Asana, T., Pedersen, S. E. and Ross, E. M. (1983) *Biochemistry* **22**, 4357-4362.

52. Cerione, R. A., Codina, J., Benovic, J. L., Lefkowitz, R. J., Birnbaumer, L. and Caron, M. G. (1984) *Biochemistry* **23**, 4519-4523.

53. Shinozawa, T., Sen, I., Wheeler, G. and Bitensky, M. (1979) *Prog. Clin. Biol. Res.* **31**, 581-734.

54. Fung, B. K.-K., Hurley, J. B. and Stryer, L. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 152-156.

Introduction: Section II

Diversity of G proteins in signal transduction

Melvin I. Simon, Michael P. Strathmann

and Narasimhan Gautam[1]

M. I. Simon is the Anne P. and Benjamin F. Biaggini Professor of Biological Sciences, in the Division of Biology, California Institute of Technology, Pasadena, California 91125.
M. P. Strathmann is a graduate student in the Division of Biology, California Institute of Technology, Pasadena, California 91125.
[1]N. Gautam is Assistant Professor in the Department of Anesthesiology, Washington University School of Medicine, St. Louis, Missouri 63110.

Science, in press

Abstract.

The heterotrimeric guanine nucleotide-binding proteins (G proteins) act as switches that regulate intracellular information processing circuits connecting cell surface receptors to a variety of effectors. This system is present in all eukaryotic cells and it mediates metabolic, humoral, neural and developmental functions. More than a hundred different kinds of receptors and many different effectors have been described. The G proteins that coordinate receptor-effector activity are derived from a large gene family. At present, the family is known to contain at least 16 different genes that encode the alpha subunit of the heterotrimer, four that encode beta subunits, and multiple genes encoding gamma subunits. Specific transient interactions between these components generate the pathways that modulate cellular responses to complex chemical signals.

Introduction.

One of the remarkable features of all biological systems is their ability to process and respond to enormous amounts of information. Much of this information is provided to individual cells in the form of changes in concentration of hormones, growth factors, neuromodulators, or other molecules. These ligands interact with transmembrane receptors and the binding event is transduced into an intracellular signal. The characterization of different families of cell surface receptors and of a variety of different mechanisms of signal transduction has been the focus of intensive recent research. In this article we examine the signal processing mechanisms associated with one of these systems, the G-protein coupled receptors (1).

Molecular cloning, biochemical, and pharmacological studies have revealed a very large family of transmembrane receptor proteins that share a characteristic topological structure. They all have seven membrane spanning domains and show considerable amino acid sequence homology (2). They appear to process information regarding changes in specific ligand binding by a mechanism that involves interaction with guanine nucleotide-binding proteins (G proteins). Signal transducing G proteins occur in two forms: the "small G proteins" that are generally found as single polypeptides composed of about 200 amino acids and the heterotrimeric G proteins that are made up of alpha, beta, and gamma subunits. The small G proteins are involved in regulating cell growth, protein secretion, and intracellular vesicle interaction (3). The heterotrimeric G proteins are associated with signal transduction from cell surface receptors (4) and are thought to act as switches that can exist in either of two states depending on bound nucleotide (Figure 1). Signal transduction is initiated by ligand binding which stabilizes an alternate conformational form of the receptor and thus transmits information across the cell membrane. This leads to a complex series of events that we understand only in broad outline. The ligand bound receptor initiates

two processes; one leads to desensitization, and occurs through receptor modification (5) and the other is a signal generating process that begins with the activation of the heterotrimeric G protein (Figure 1A). Interaction of the G-protein with the activated receptor promotes the exchange of GDP bound to the α subunit for GTP and the subsequent dissociation of the α-GTP complex from the β-γ heterodimer (Figure 1B). A single receptor can activate multiple G protein molecules, thus amplifying the ligand binding event. The α subunit with bound GTP and the free β-γ subunit may interact with effector proteins that further amplify the signal. Such effectors include ion channels and enzymes that generate regulatory molecules or second messengers. Low molecular weight second messengers such as cyclic adenosine-monophosphate (cAMP) or inositol triphosphate, in turn generate dramatic intracellular changes including selective protein phosphorylation, gene transcription, cytoskeleton reorganization, secretion, and membrane depolarization. Termination of the signal occurs when GTP bound by the G protein α subunit is hydrolysed to GDP. The alpha subunit then reassociates with beta-gamma subunits.

More than a hundred different G protein-coupled receptors have been found in mammals, including distinct receptors that bind the same ligand. At least five different muscarinic receptors, more than eight different adrenergic receptors, five different serotonin receptors, and four different opsins (2) have been identified. A growing family of receptors and receptor subtypes that respond to purines, bombesin, bradykinin, thrombin, histamine, dopamine, ecosinoids, vasopressin, growth hormone releasing hormone, and somatostatin are being cloned and characterized. Distinct forms, or subtypes, of the receptors that respond to the same ligand may be differentiated by the intracellular responses that they elicit. Specific receptor subtypes are coupled to different second messenger pathways and to the regulation of different ion channels. Since a single receptor subtype can be coupled to multiple effectors and

multiple receptor subtypes can activate a single effector (2, 4, 6), the G-protein coupled interactions form complicated networks. Furthermore, characterization of effectors has revealed that they too are specified by extensive gene families. Cloning studies have identified different types of adenylyl cyclases (7), five different phospholipase C isotypes (8), and multiple types of phospholipase A2 (9). There is also evidence for multiple calcium, potassium, and possibly sodium channels that are responsive to G proteins (10). This array of receptors and effectors raises questions about the nature of the information processing circuits that are formed. It remains unclear how many different G proteins are required to couple the receptor and effector subtypes, how specific receptors are linked through G-proteins to form autonomous circuits, how circuits interact with each other, and how they are reshaped during cellular differentiation. Nor do we know what controls the specificity, the level of amplification, the timing, and the crosstalk between signals in these circuits.

The coordination of this network of circuits is a formidable task in which the G proteins play a central role. To understand how the complex family of G proteins function, we must identify the components of the G-protein mediated networks and the nature of their specific interactions.

G Protein α Subunit Diversity.

Cloning and sequencing techniques have been most productive in identifying and classifying new α subunits. When the deduced amino acid sequences of all of the α subunits that have been cloned (more than 30 different cDNAs) are aligned, approximately 20% of the amino acids are found to be invariantly conserved (11). Amino acid sequence similarity provides a direct measure of the relatedness of different alpha subunits. Using a variety of probes, multiple G proteins have been identified in all of the eukaryotic organisms that have been examined. A classification of the G

alpha subtypes found in mammals based on amino acid sequence similarity is shown in Figure 2. The family is made up of four classes and each class is composed of specific isotypes. Thus, the G_S class includes both the $G\alpha_S$ and $G\alpha_{olf}$ isotype (12). $G\alpha_{olf}$ shows 88% amino acid sequence identity with $G\alpha_S$ and is able to activate adenylyl cyclase to increase intracellular cAMP levels. It is expressed in specific neural tissues and is highly enriched in neurons in the olfactory epithelium. Individual isotypes are found to be highly conserved between species; for example, there are no amino acid differences between $G\alpha_S$ isolated from humans and from mice. It is therefore easy to identify the α subunit isotypes in a variety of mammals even though some of them such as $G\alpha_{i1}$ and $G\alpha_{i3}$ (13) differ in only 6% of their amino acid sequences. In addition to amino acid similarity, there is conservation at the level of gene structure. The genes encoding $G\alpha_{i1}$, $G\alpha_{i2}$, $G\alpha_{i3}$, $G\alpha_o$, and alpha subunits of the transducin rod and cone photoreceptors all conserve the positions of their introns and exons (14).

The functional role of specific α subunits is not obvious from their structural classification. One of the most effective tools in implicating G proteins in specific functions in intact cells has been the use of pertussis toxin (PTX) (15). The toxin uncouples the receptor from its G protein and thus blocks signal transduction by receptors that cause decreases in cyclic AMP levels, that regulate ion channels, and that activate phospholipases. Since members of the G_i class of alpha subunits contain sites susceptible to modification by PTX, they are expected to mediate activation of the pertussis toxin sensitive processes. Indeed, the $G\alpha_i$ subtypes and $G\alpha_o$ have been shown to have a role in regulating ion channels. In addition, activated $G\alpha_i$ subunits lower intracellular cAMP levels and $G\alpha_o$ has been implicated in increasing phosphoinositide release (16). However, the mechanism by which $G\alpha_i$ lowers cAMP levels is not clear (see below). Only in the case of the $G\alpha_S$ and $G\alpha_{t1}$ (rod transducin) proteins have preparations of the proteins been reconstituted with their purified

effectors and receptors. Transducin couples rhodopsin to the activation of retinal phosphodiesterase (17) while the ubiquitous $G\alpha_s$ and its splice variants activate purified reconstituted adenylyl cyclase (18).

GTP dependent signaling pathways that are resistant to PTX and activate phospholipases have also been described. Recent cloning experiments have identified a variety of new G-proteins that function in these pathways (19). Eight G alpha cDNA clones belonging to three different classes of G alpha subunits have been characterized. They include two novel classes, G12 (19) and G_q (20) (Figure 2); all of these lack the cysteine residue four amino acids from the C terminal end that is the target for PTX mediated ADP ribosylation (Figure 3). The $G\alpha_z$ (21) and $G\alpha_q$ (22) proteins have been isolated and shown to be refractory to PTX modification.

The $G\alpha_z$ subunit bears some resemblance to the G_i class (23); however, it differs markedly in biochemical properties. The $G\alpha_z$ heterotrimer has been purified from bovine brain and the α subunit has been expressed in *E coli* (21). The purified α subunit shows a very slow rate of guanine nucleotide exchange and an unusual Mg^+ ion dependence when compared to the $G\alpha_s$ and $G\alpha_i$ proteins. Furthermore, its intrinsic guanosine triphosphatase (GTPase) activity is extremely slow, approximately 100 times slower than that determined for the other G protein α subunits. While the function of $G\alpha_z$ remains obscure, its kinetic properties and its distribution is interesting. It is found primarily in neurons, particularly cells with long axonal processes (24).

$G\alpha_q$ and $G\alpha 11$ are ubiquitously distributed and they lack a site for PTX modification (20). The amino acid sequences of the $G\alpha 11$ and $G\alpha_q$ isotypes differ from each other by less than 12% and almost all of these changes are confined to the NH_2- terminal region of the molecule. This region may be important in determining the specificity of interaction with the β-γ subunit and the relative rate of nucleotide

exchange and hydrolysis (see below). Thus, while $G\alpha_q$ and $G\alpha 11$ are often found in the same cell, they may be responsible for generating signals with different time constants and they could interact with different subsets of receptors and effectors. Recent results clearly point to the involvement of $G\alpha_q$ and $G\alpha 11$ in PTX resistant coupling to phospholipase C activation. A novel 42 kd protein that activates phospholipase C in a PTX resistant fashion was partially purified (22, 25). The 42 kd G protein had amino acid sequence identity with the $G\alpha_q$ clone (21). In reconstitution experiments the 42 kd protein specifically activates the beta isotype of phospholipase C (26, 27) and not the gamma or delta form (27). Antisera prepared against peptides with the specific sequences found at the C-terminus and in other regions of $G\alpha_q$ react with the 42 kd protein. Finally, in *Drosophila,* a cDNA clone with 76% amino acid identity to mammalian $G\alpha_q$ (20) was found to be expressed in eye tissue (28). It may represent the G-protein that couples *Drosophila* rhodopsin to the activation of the phospholipase C that is involved in the phototransduction cascade. Taken together all of this work clearly indicates that $G\alpha_q$ and $G\alpha 11$ are involved in coupling one type of phospholipase C to a specific set of receptor subtypes.

There are three other isotypes in the G_q class. They all show restricted patterns of tissue specific expression. $G\alpha 14$ is found primarily in stromal and epithelial cells (29), while $G\alpha 15$ and $G\alpha 16$ (30) are found in cells derived from the hematopoietic lineage. $G\alpha 15$ is found in murine B lymphocytes and $G\alpha 16$ in human T lymphocytes and both are found in myeloid cells. Since there are multiple phospholipase C isotypes (8), it is possible that the other members of the G_q class interact with different members of the phospholipase family.

$G\alpha 12$ and $G\alpha 13$ (31, 32) represent yet another class of potential pertussis toxin resistant α subunits (Figure 2 and 3). Both $G\alpha 12$ and $G\alpha 13$ mRNA are expressed ubiquitously. Again, we know very little about the function of these $G\alpha$

proteins. There is some evidence indicating that a member of the Gα12 family has been found in *Drosophila* melanogaster. A cDNA clone corresponding to the gene Concertina (cta) in *Drosophila* has marked amino acid sequence homology with Gα12 (32). It has been suggested that the Concertina gene might play a role in regulating embryonic development in *Drosophila.*

The diversity of α subunit structure is further extended by differential splicing of complex genes. At least two variants can be generated by differential splicing of the Go gene and most preparations of Go contain both polypeptides (33). These variants differ in amino acid sequence at the COOH-terminal half of the protein and may therefore have different receptor and effector specificities (see below). It has been suggested that the Go variants mediate PTX-sensitive activation of phospholipase C (34).

The use of the polymerase chain reaction (PCR) has revealed new G protein alpha subunits (19). However, it is not clear whether all members of the family have been identified. It is possible that some α subunits are expressed only in a small subset of cells, or that the sequences of novel subunits could diverge in conserved regions such that the probing techniques would not detect them. Splice variants of alpha subunit cDNAs may have been overlooked during the search. However, there has not been a great deal of pressure to search for new G proteins, since in terms of assigning functions to known G proteins, we already face an embarrassment of riches.

Structure function relationships among the Gα subunits.

Unfortunately, the crystal structure is not yet known for any alpha subunit of the heterotrimeric G proteins. An approximation of the three dimensional structure of the G alpha subunit has been developed by modeling the amino acid sequences based on the crystal structure of the small G-protein Ras and of elongation factor-TU,

another GTP binding protein (35). Figure 3 shows the disposition of functional domains in a "normalized" G alpha subunit based on site-specific mutagenesis and studies with chimeric genes. This work has been reviewed recently (36) and we will briefly summarize some of the results. Mutations in the Gα subunit that are analogous to those that have been studied in Ras exemplify both the differences and the similarities between the two systems. Ras ordinarily hydrolyzes GTP very slowly (approximately 100 times slower than $G\alpha_s$). The rate of hydrolysis is accelerated by interaction with another protein called the GTPase activating protein (GAP). The substitution of valine for glycine in the A-box (see Figure 3) in ras lowers its GTPase activity and also results in its inability to be activated by the GAP protein. The homologous change introduced in $G\alpha_s$ results in a two- to four-fold change in GTPase activity (37). However, amino acid changes at the Arginine residue that is modified by ADP ribosylation with cholera toxin (38), or in other nearby residues (e.g., gln-227) leads to dramatic (30- to 100-fold) decreases in intrinsic GTPase activity. It has been suggested that this portion of the molecule (from residue approximately 100-230, Figure 3) has a critical role in regulating the GTPase activity. Recently, similar mutations have been found in $G\alpha_{i2}$ and $G\alpha_s$ by screening tumors for amino acid changes in G alpha subunit proteins. These results suggest that mutations that lock the G protein in the GTP bound form are dominant and in some tissues may lead to changes in growth control and oncogenesis (39).

The NH_2-terminal region of the G alpha subunit is thought to be involved in interaction with the beta-gamma subunit. This notion is supported by the finding that proteolysis of the N terminus prevents the G alpha subunit from binding to the beta-gamma subunit (40). The N-terminal region is also the site of myristoylation on some of the G-proteins ($G\alpha_o$ and $G\alpha_{i1}$, $G\alpha_{i2}$ and $G\alpha_{i3}$). Myristoylation increases the affinity of the alpha subunit for beta-gamma and facilitates heterotrimer formation (41).

It has been suggested that the C terminal region of the G protein is involved in receptor interactions. This suggestion is supported by the observation that modification of the G alpha subunit by pertussis toxin blocks its interaction with receptor. Antibodies or peptides that specifically interact with C terminal regions of some of the G alpha proteins also block interaction with receptors (42). G-protein-effector interactions have been examined by constructing chimeric α subunits. The results suggest that sequences in the C-terminal half of the G-alpha subunit can determine effector specificity (43). However, this domain is not well defined. Reconstituted systems containing purified components may eventually allow mapping of the amino acid residues required for specific protein binding and thus pinpoint the nature of the interactions.

Experiments demonstrating a direct role of G alpha subunits in the activation of ion channels have dramatically increased our understanding of G protein function. In addition to activating adenylyl cyclase, $G\alpha_S$ subunits have been shown to regulate calcium channels (44). $G\alpha_O$ regulates a variety of neuronal and atrial potassium channels as well as calcium channels in dorsal root ganglia (45). The $\alpha i3$ and $\alpha i1$ proteins appear to activate potassium channels (10, 16). The use of patch-clamp techniques, highly purified or recombinant G proteins and the observation of rapid channel activation suggest that the effect of the activated alpha subunit is directly on the channel. It has been known for a long time that channels can be gated by the activities of some second messengers. Thus, in addition to a direct effect on ion channel function, channels can be regulated indirectly through G protein interaction with specific effectors which in turn activate second messengers and kinases that can modify channel activity.

G alpha subunit evolution.

There is no evidence for a cell surface receptor coupled G-protein system in bacteria. In fungi, homologs of the subunits of heterotrimeric G proteins do exist. Two genes encoding G protein alpha subunits have been found in yeast. One of the heterotrimeric G proteins in yeast is coupled to the mating type receptor. The mechanism of action of G-proteins is different in yeast (*S. cerevisiae*) than in multicellular organisms; it appears that the beta-gamma heterodimer rather than the alpha subunit interacts with effector (46). Nonetheless, the overall features of the system are conserved, since mammalian G protein alpha subunits can restore partial function to yeast mutants. Mammalian $G\alpha_s$ and $G\alpha_i$ subunits are apparently able to interact well with the yeast beta-gamma subunit and thus inhibit the mating-type pathway, but the mammalian proteins respond poorly to the yeast mating type receptor because they lack the appropriate receptor specificity (47). If, however, the gene for the mating type receptor is replaced by the gene for the appropriate mammalian receptor, e.g., the β-adrenergic receptor, then catecholamines will trigger the yeast mating response (47). This kind of hybrid system promises to be very useful in screening for effective receptor agonists and antagonists.

In *Dictyostelium discoideum*, the slime mold, an extensive family of G protein alpha subunit genes has been described. They play important roles in regulating cellular aggregation and development (48). The amino acid sequences of the Gα subunits maintain the universally conserved GTP binding and hydrolysis motifs (Figure 3) but do not bear a direct relationship to the G protein classes found in mammalian cells. On the other hand, several of the G proteins from simple multicellular organisms show a clear relationship to the classes found in mammals. $G\alpha_s$ and $G\alpha_o$ homologues have been identified in *C. elegans* (49). In *Drosophila* (50) sequences homologous to G_i, G_s, G_q, and G12 alpha subunit classes have all

been found. These sequences generally bear from 60-75% amino acid identity with the homologous mammalian proteins. If the amino acid sequence identity among members of classes of G alpha subunits reflect their interaction with similar subsets of effectors or receptors, we can expect insights into mammalian G-protein function from studies on these simpler systems.

Beta-gamma subunit diversity and function.

Four distinct beta subunit isotypes have been found in mammals (51, 52). They share more than 80% amino acid sequence identity. However, an individual isotype, Gβ1, cloned from different mammalian sources has identical amino acid sequence. The same is true for Gβ2, suggesting that each isotype has a conserved sequence and may also have conserved function. Gβ1, Gβ2, and Gβ3 are ubiquitously expressed while Gβ4 is abundant in brain and lung tissue but is found at low levels in other tissues (52). All of the beta subunits are made up of eight segments of amino acid sequence; each segment shares a repetitive 40 amino acid sequence motif that is characterized by a number of distinct amino acids, including a tryptophan-aspartic acid pair. This motif has been referred to as the WD-40 repeat (51). The repeat structure is common to beta subunits found in all organisms. Recently, this same motif has been found as part of a large number of other proteins (53). The function of the WD-40 repeat is not known and its distribution in other genes has not as yet provided a useful clue. One portion of the beta sequence that does not show the repeat structure is the first 30-40 amino acids on the NH_2-terminal end of beta. Crosslinking studies with transducin suggest that this is the region where beta and gamma subunits may interact (54).

Diversity amongst the gamma subunits has been demonstrated by electrophoretic, immunological, and protein sequencing techniques (55). However,

these data do not allow us to distinguish between heterogeneity due to primary structure or to post-translational modification. cDNA clones corresponding to four distinct isotypes have been found (56). Peptide sequences obtained from purified gamma subunit proteins suggest the existance of at least two more. One gamma isotype, Gγ1, is expressed only in photoreceptors, while another, Gγ2, is expressed at different levels in all tissues that were examined; Gγ3 is expressed primarily in brain and in testis. The proteins are most divergent at their NH_2-terminal sequence and they share considerable sequence homology at their COOH-terminus. The amino acid sequence near the COOH-terminus of the gamma subunits (as deduced from the cDNA), resembles the Ras oncogene sequence with a characteristic cysteine four residues from the end of the protein. The gamma subunits are all modified (57) by the removal of the three carboxyl terminal amino acids adjacent to the cysteine, and by the carboxymethylation, and isoprenylation of the terminal cysteine residue. The modification of Gγ1 most closely resembles the modification of Ras; a 15 carbon farnesyl group is found on the carboxymethylated cysteine (58). Without modification, Gγ1 complexed to beta is inactive (59). Thus, isoprenylation and perhaps reversible carboxymethylation could play roles in the regulation of signal transduction. Gamma subunits extracted from brain were also found to be polyisoprenylated. However, they are modified by the addition of a 20 carbon all trans-geranylgeranyl moiety (60). This addition may be required to anchor the gamma subunit in the membrane.

There are a few examples of direct effects of beta-gamma on purified components of mammalian signaling systems. Beta-gamma addition to rod photoreceptor outer segments apparently activates phospholipase A2 (61). The activation of phospholipase A2 with subsequent release of arachadonic acid could account for some of the observations that addition of βγ to membrane patches activates

potassium channels. However, inhibitors of phospholipase A2 do not block acetylcholine activation of K^+ channels and recent reports suggest that beta-gamma addition can inhibit K^+ channel activation by interaction with alpha subunits (62). Part of the confusion in these and other experiments designed to look for an effect of $\beta\gamma$ may result from the complex nature of the beta-gamma mixtures that are added back to membrane preparations and from the complexity of the membrane patch system.

The suppressive effect of added beta-gamma subunits on the activity of GTP activated alpha subunits has been demonstrated in a number of contexts. In fact, this is the basis for one explanation of the mechanism of G_i mediated inhibition of adenylyl cyclase (63). Indirect evidence has implicated activation of $G\alpha_i$ in lowering intracellular cAMP levels; however, the purified alpha subunit shows only mild activity in reconstituted systems. This led to the hypothesis that activation of $G\alpha_i$ frees beta-gamma subunits to interact with endogenous $G\alpha_S$ and, thus, activation of $G\alpha_i$ operates indirectly by inactivating $G\alpha_S$. The scheme has been criticized on kinetic and other grounds (64, 16). However, it remains a consistent explanation of the data. Another hypothesis is that $\beta\gamma$ subunits act directly. This notion is based on the observation that beta-gamma subunits added to adenylyl cyclase inhibit its activity (65). Again, some of these effects might be indirect; they may result from the interaction of beta-gamma with calmodulin thereby inhibiting the Ca^{2+}-calmodulin-mediated stimulation of some of the adenylyl cyclase activity (66).

Beta-gamma subunits act during the signaling process to coordinate cellular responses. They have several functions: (a) stabilizing the interaction of alpha subunits with receptors and perhaps inducing formation of appropriate receptor complexes exhibiting specific activation kinetics, (b) modulating the effects of activated alpha subunits, and (c) regulating, at least indirectly, channel and phospholipase activity. The extensive modification of the gamma subunits, the diversity of beta and

gamma, as well as the crosstalk between beta-gamma subunits that are associated with different alpha subunits, all point to an important function for beta-gamma in establishing specific receptor/G protein associations and in integrating the effects and timing of the various G-protein mediated circuits.

G-protein Networks.

In exploratory experiments with cloned cell lines, appropriate probes have been used to detect $G\alpha_s$, $G\alpha_{i1}$, $G\alpha_{i3}$, $G\alpha_o$, $G\alpha_z$, $G\alpha_q$, and $G\alpha 11$ mRNA in the same cell. Presumably, there are also detectable amounts of four beta and at least three gamma subunit genes expressed in some cloned cell lines. If all of these subunits associated combinatorially and at random, there would be almost one hundred different kinds of heterotrimers. Different combinations could have different affinities for individual receptors. These combinatorial relations could regulate the association between G protein and receptor and the kinetics of the activation response, since each receptor when transiently activated would interact with the subset of combinations of G-protein subunits for which it had the highest affinity. However, there may be a mechanism that assembles the heterotrimer in a specific manner and transports specific assemblies to intracellular compartments that are enriched for the presence of appropriate receptors or effectors. There is evidence for compartmentalization (67) and for the preferential formation of subsets of beta-gamma heterodimers associated with specific alpha subunits. Reconstitution experiments have shown that $G\alpha_s$, $G\alpha i$, and transducin have different affinities for different $\beta\gamma$ heterodimers (68) and the differential elution of α subunits from $\beta\gamma$ columns further attests to differences in their relative affinities (69). A $\beta\gamma$ complex isolated from placenta has been shown to have distinctive properties (70), and different gamma subunits were found to be associated in heterotrimers with the same alpha subunit when isolated from different tissues (71).

A precise role for $\beta\gamma$ in determining receptor or effector specificity *in vivo* remains obscure.

Network specificity can also be controlled by feedback processes. Activation of a particular G protein-coupled pathway can generate second messengers that regulate protein kinases. The kinases, in turn, can influence the information processing system. There are a number of examples where the addition of ligand leads to rapid phosphorylation of the G alpha subunit. In *Dictyostelium* the $G\alpha 2$ protein plays a critical role in transducing signals from cyclic AMP receptors. The addition of ligand to the cells results in very rapid phosphorylation of $G\alpha 2$ (72). Phosphorylation is transient and it is not clear how this modification affects the activity of $G\alpha 2$. It could lead to inactivation and thus reflect a desensitization or adaptation process. In the yeast mating type system where mutants in the alpha subunit lead to constitutively active beta-gamma subunits, a process (not necessarily protein phosphorylation) that desensitizes and reverses the long term effects of free beta-gamma results in adaptation (73). There is also evidence for the rapid phosphorylation of activated G_k alpha subunits and of G_{i2} alpha subunits (74). However, it is not clear how these modifications affect function. There are a variety of other modifications including myrisoylation, isoprenylation, carboxymethylation and ADP-ribosylation that could also be regulated to modulate the activity of different G proteins.

Some of the circuitry mediated by G proteins is presumably "hardwired" and can serve the function of signal distribution. For example, $G\alpha_s$, when activated, is capable of both opening specific Ca^{++} ion channels and increasing steady state levels of cyclic AMP. Thus, these responses are presumably coordinated by the activation of a single $G\alpha_s$ protein. G protein similarities can generate crosstalk between circuits, resulting in signal integration. If two G proteins such as $G\alpha_{i2}$ and $G\alpha_o$ are activated by different receptors and characteristically deliver signals to distinct effectors but are

capable of interacting at low efficiency with other effectors, we would expect that the activation of either one of these pathways could elicit activation of the other. Thus, crossactivation can be an essential part of the information transducing circuit. On the other hand, parts of the intracellular system may be built to shield against crosstalk. For example, (75) it has been found that in polarized renal epithelium cells $G\alpha_{i2}$ is localized to the basal lateral membrane of the cell while $G\alpha_{i3}$ is found in the golgi and in apical membranes. By confining specific G proteins and perhaps their effectors to local regions of the cell, their ability to crosstalk or interact could be effectively stymied.

There are other proteins that appear to augment the function of G proteins. Nucleoside diphosphate kinase (NDPK) has been reported to form complexes with the various GTP binding proteins (76). The enzyme is a source of GTP and its association with the alpha subunits may reflect functional interactions. There is also indirect evidence suggesting that perhaps the small GTP binding proteins such as Ras bound to GAP (77) can interact with the G alpha subunit.

The properties of G-protein networks are well suited to the needs of the complex processing that takes place in the nervous system. Many neuromodulators and neurotransmitters operate through G-protein coupled pathways and G-proteins can effect specific ion channel function. It is therefore not surprising that long term potentiation (LTP) in the mammalian hippocampus is thought to involve G proteins (78). In aplysia, associative learning is known to be mediated by multiple G-protein pathways (79). We are beginning to get indications of a most exciting aspect of the complexity of G-protein networks--their role in coordination and integration of information during neural function.

Role of G proteins in differentiation.

As cells differentiate, presumably their signaling characteristics change and new pathways are expressed and integrated into the function of the cell. Perhaps the clearest example of expression of new G proteins as a function of differentiation is found in *Dictyostelium*, where the transitions from amoeba to aggregate and from slug to stalk and spore are all accompanied by the appearance of new G alpha protein subunits (80). Presumably, the new transducers can establish and generate circuits that produce a pattern of second messengers that help stabilize the differentiated state of these cells. Studies of the effects of one G-protein coupled pathway on another suggest that there are feedback mechanisms that regulate gene expression of the components of the pathway (81).

The notion that G proteins may initiate or stabilize developmental changes is an interesting one. There are a variety of regulatory peptide factors and hormones that can influence cell growth. Many of these agents, such as thrombin, bombesin, serotonin, and angiotensin, bind to specific receptors that are coupled to G proteins. In some cases, pertussis toxin has been reported to block the proliferative response. The toxin has also been used to show that growth factor receptors may also be coupled to G proteins (82). Marked effects on the signaling mechanisms of insulin-like growth factor (IGF) have been described and peptides that contain a sequence derived from the IGF receptor (83) have been shown to carry the specificity to activate certain G proteins. This implies that apart from the seven-pass membrane receptors, a whole new class of receptors with single transmembrane segments could also be coupled to G-proteins.

G proteins may also play a role in modulating cell movement, cytoskeletal structure, and chemotaxis. G proteins have been reported to interact with tubulin (84), collagen (85), and actin (86). Furthermore, pertussis toxin blocks some of these

interactions. One hope is that genetic systems can be used to analyze complex functions of G-protein pathways. In *Drosophila*, a maternal effects gene called concertina (cta) that is required for appropriate morphological development has been cloned and shown to be a G alpha subunit (32). G proteins may act synergistically with growth factors to direct a cell along a particular developmental pathway. Thus, signals from specific hormones and mediators that induce high, intracellular concentrations of a particular second message may be integrated to facilitate a growth factor initiated response. For example, ligands that activate G protein-coupled receptors that increase cyclic AMP concentration in cells enhance the differentiation of PC12 cells initiated by nerve growth factor. Tyrosine kinase-coupled receptors and G protein-coupled receptors activate different phospholipase C isotypes. These two pathways could act synergistically. Thus, the activation of G protein mediated pathways in the appropriate context might initiate, facilitate, or amplify changes required for cellular differentiation.

Cloning and sequencing have been used to establish the nature of the G-proteins, and experiments utilizing purified protein to reconstitute steps of signalling pathways *in vitro* have helped to define the mechanism of action of G protein-coupled systems. However, these approaches give us little insight into the integrative properties of the system or into its function *in vivo*. New microscopes and specific fluorescent markers may allow us to visualize the effects of different circuits simultaneously. Applications of homologous recombination and antisense oligonucleotides should provide ways to inactivate individual genes and allow assessments of their function. Extensive studies with dominant mutants in mammalian systems and in transgenic mice have proven to be useful and may be extended to define the function of individual components and provide the means to analyze the complex interactions within the signal transduction network.

References

1. The citations are not comprehensive. They are intended only to provide the reader with an entry into some aspects of the recent literature.

2. J. R. Raymond, M. Hnatowich, M. G. Caron and R. Lefkowitz, *ADP-Ribosylating Toxins and G-Proteins*, J. Moss and N. Vaughan, Eds., (American Society of Microbiology, 1990), pp. 163-176; H. Dohlman, J. Thorner, M. G. Caron, R. J. Lefkowitz, *Ann. Rev. Biochem.* **54**, in press (1991).

3. A. Hall, *Science* **249**, 635 (1990).

4. A. Gilman, *Annu. Rev. Biochem.* **56**, 615 (1987); T. Pfeuffer and E. J. M. Helmreich, *Curr. Topics in Cell. Reg.* **29**, 129 (1988).

5. W. P. Hausdorff, M. G. Caron and R. J. Lefkowitz, *Fed. Amer. Soc. Exper. Biol.* **4**, 2881 (1990).

6. E. M. Ross, *Neuron* **3**, 1412 (1989); L. Birnbaumer, *Annu. Rev. Pharm. and Toxicol.* **30**, 675 (1990).

7. J. Krupinski et al., *Science* **244**, 1558 (1990).

8. S. G. Rhee, P.-G. Suh, S.-H. Ryu, S. Y. Lee, *Science* **244**, 546 (1989).

9. E. Diez and S. Mong, *J. Biol. Chem.* **265**, 14654 (1990).

10. A. Brown and L. Birnbaumer, *Annu. Rev. Physiol.* **52**, 197 (1990); G. Schulz, W. Rosenthal, F. Heschiler, W. Trautwein, *Annu. Rev. Physiol.* **52**, 275 (1990).

11. M. Lochrie and M. I. Simon, *Biochem.* **27**, 4957 (1988).

12. D. T. Jones and R. R. Reed, *Science* **244**, 790 (1989).

13. D. T. Jones and R. R. Reed, *J. Biol. Chem.* **262**, 14241 (1987).

14. Y. Kaziro *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **53**, 209 (1988).

15. M. Ui., *ADP-Ribosylating Toxins and G-Proteins*, J. Moss and M. Vaughan, Eds., (American Society of Microbiology, 1990), pp. 45-77.

16. L. Birnbaumer *et al.*, *Recent Prog. in Hormone Research* **45**, 121 (1989).

17. M. Chabre and P. Deterre, *Eur. J. Biochem.* **179**, 255 (1989).

18. M. P. Graziano, P. J. Casey and A. G. Gilman, *J. Biol. Chem.* **262**, 11375 (1987); M. P. Graziano, F. Freissmuth and A. G. Gilman, *J. Biol. Chem.* **264**, 409 (1989).

19. M. Strathmann, T. Wilkie and M. I. Simon, *Proc. Natl. Acad. Sci. USA* **86**, 7407 (1989).

20. M. Strathmann and M. I. Simon, *Proc. Natl. Acad. Sci. USA* **87**, 9113-9117 (1990).

21. P. J. Casey, H. K. W. Fong, M. I. Simon and A. G. Gilman, *J. Biol. Chem.* **265**, 2383 (1990).

22. I.-H. Pang and P. Sternweis, *J. Biol. Chem.* **265**, 18707 (1990).

23. M. Maksuoka, M. Itoh, T. Kozasa and Y. Kaziro, *Proc. Natl. Acad. Sci. USA* **85**, 5384 (1988); H. K. W. Fong, K. K Yoshimoto, P. Eversole-Cire and M. I. Simon, *Proc..Natl. Acad. Sci. USA* **85**, 3066 (1988).

24. D. R. Hinton *et al.*, *J. Neuroscience* **10**, 2763 (1990).

25. S. J. Taylor, J. A. Smith and J. H. Exton, *J. Biol. Chem.* **265**, 17150 (1990).

26. S. J. Taylor, H. Z. Chae, S. G. Rhee and J. H. Exton, *Nature*, in press (1991).

27. A. V. Smrcka, J. R. Hepler, K. O. Brown and P. C. Sternweis, *Science*, in press (1991).

28. Y.-J. Lee, M. B. Dobbs, M. L. Verardi and D. R. Hyde, *Neuron* **5**, 889-898 (1990).

29. T. Wilkie, M. P. Strathmann and M. I. Simon, in preparation.

30. T. Amatruda, D. Steele, V. Slepak and M. I. Simon, *Proc. Natl. Acad. Sci. USA*, in press (1991).

31. M. P. Strathmann and M. I. Simon, *Proc. Natl. Acad. Sci. USA* , in press (1991).

32. S. Parks and E. Weischaus, *Cell* **64**, 447 (1991).

33. W. H. Hsu *et al.*, *J. Biol. Chem* **265**, 11220 (1990); M. Strathmann, T. Wilkie and M. I. Simon, *Proc. Natl. Acad. Sci. USA* **87**, 6477 (1990).

34. A. Kikuchi *et al.*, *J. Biol. Chem.* **261**, 3901 (1986); D. A. Ewald, I.-H. Pang, P. C. Sternweis and R. J. Miller, *Neuron* **2**, 1185 (1989); T. M. Moriarty *et al.*, *Nature* **343**, 79 (1990).

35. K. R. Halliday, *J. Cyclic Nucleotide Protein Phos. Res.* **9**, 435 (1984); S. B. Masters, R. M. Stroud and M. R. Bourne, *Protein Engineer.* **1**, 47 (1988); S. R. Holbrook and S. H. Kim, *Proc. Natl Acad. Sci. USA* **86**, 1751 (1989).

36. H. R. Bourne, C. A. Landis and S. B. Masters, *Proteins: Structure Function and Genetics* **6**, 222 (1989); H. R. Bourne, D. A. Sanders and F. McCormick, *Nature* **348**, 125 (1990); ibid **349**, 117 (1991).

37. M. P. Graziano and A. G. Gilman, *J. Biol Chem* **264**, 15465 (1989); S. B. Masters *et al.*, *J. Biol. Chem.* **264**, 15467 (1989).

38. M. Freissmuth and A. G. Gilman, *J. Biol. Chem* **264**, 21907 (1989).

39. C. Landis *et al.*, *Nature* **340**, 692 (1989); J. Lyons *et al.*, *Science* **249**, 655 (1990).

40. H. Hamm *et al.*, *Science* **241**, 832 (1988); E. Neer, L. Pulsifer and L. G. Wolf, *J. Biol. Chem* **263**, 8996 (1988).

41. T. L. Z. Jones *et al.*, *Proc. Natl Acad. Sci. USA* **87**, 568 (1990); S. M. Mumby, R. O. Heukeroth, J. I. Gordon and A. G. Gilman, Ibid **87**, 728 (1990).

42. D. Deretic and H. Hamm, *J. Biol. Chem.* **262**, 10839 (1987); K. Sullivan *et al.*, *Nature* **330**, 758 (1987).

43. S. B. Masters *et al.*, *Science* **241**, 448 (1988); S. Osawa *et al.*, *Mol. Cell Biol.* **10**, 2931 (1990); S. Gupta *et al.*, *Science* **249**, 662 (1990).

44. A. Yatani *et al.*, *J. Biol Chem.* **263**, 9887 (1988).

45. A. Yatani *et al.*, *Science* **249**, 1163 (1990); J. Hescheler *et al.*, *Nature* **325**, 445 (1987).

46. D. Dietzel and J. Kurjan, *Cell* **50**, 1001 (1987); A. Miyajima *et al. Cell* **50**, 1011 (1987); M. Nakafuku *et al., Proc. Natl. Acad. Sci USA* **85**, 1374 (1988).

47. K. Matsumoto *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **53**, 567 (1988); K. King *et al.*, *Science* **250**, 121 (1990).

48. P. Devreotes. *Science* **245**, 1054 (1989).

49. M. Lochrie, J. Mendel, P. Sternberg and M. I. Simon *Cell Regulation*, in press (1991).

50. C. J. Schmidt, S. G. Fazio, Y.-K. Chow and E. J. Neer, *Cell Regulation* **1**, 125 (1989); A. Yarfitz, N. M. Provost and J. B. Hurley, *Proc. Natl Acad. Sci. USA* **85**, 7134 (1988); N. M. Provost, D. E. Somers and J. B Hurley, *J. Biol. Chem.* **263**, 12070 (1988); F. Quan, W. J. Wolfgang and M. A Forte, *Proc. Natl Acad. Sci. USA* **86**, 4321 (1989); J. Yoon, R. D. Shortridge, B. T. Bloomquist, S. Schneuwly, M. H. Perdew and W. L. Pak, *J. Biol. Chem.* **264**, 18536 (1989); S. M. de Sousa, L. L. Hoveland, S. Yarfitz and J. B. Hurley, *J. Biol. Chem.* **259**, 18544 (1989); N. C. Thambi, F. Quan, W. I. Wolfgang, A. Spiegel and M. Forte, *J. Biol. Chem.* **259**, 18552 (1989).

51. T. Sugimoto *et al.*, *FEBS Letters* **191**, 235 (1985); J. Codina *et al.*, *FEBS Letters* **207**, 187 (1986); H. K. W. Fong *et al.*, *Proc. Natl Acad. Sci. USA* **83**, 2162 (1986); H. K. W. Fong *et al.*, Ibid **84**, 3792 (1987); B. Gao, A. G.

Gilman and J. D. Robishaw, Ibid **84,** 6122 (1987); M. Levine *et al.*, Ibid **87**, 2329 (1990).

52. E. von Weizsäcker, M. P. Strathmann and M. I. Simon, *Mol. Cell. Biol.*, submitted (1991).

53. M. A. Dalrymple *et al.*, *Cell* **58**, 811 (1989); R. Ruggieri *et al.*, *Proc. Natl. Acad. Sci. USA* **86**, 8778 (1989); (b) F. Guillemot *et al.*, Ibid **86**, 4594 (1990); D. A. Hartley, A. Preiss and S. Artavanis-Tsakonas, *Cell* **55**, 785 (1988).

54. J. Bubis and H. G. Khorana, *J. Biol. Chem.* **265**, 12995 (1990).

55. P. Gierschik *et al.*, *Proc. Natl. Acad. Sci USA* **82**, 727 (1985); J. D. Hildebrandt *et al.*, *J. Biol. Chem* **260**, 14867 (1985); D. J. Roof, M. L. Applebury and P. C. Sternweis, Ibid **260**, 16242 (1985).

56. J. B. Hurley *et al.*, *Proc. Natl Acad. Sci. USA* **81**, 6948 (1984); K. Yatsunami, B. V. Pandya, D. D. Oprian and H. G. Khorana, Ibid **82**,1936 (1985); N. Gautam, M. Baetscher, R. Aebersold and M. I. Simon, *Science* **244**, 971 (1989); J. Robishaw, V. K. Kalman, C. R. Moomaw and C. A Slaughter, *J. Biol. Chem.* **264**, 15758 (1989); N. Gautam, J. Northup, H. Tamir and M. I. Simon, *Proc. Natl Acad. Sci. USA* **87**, 7973 (1990); T. Asano, R. Morishita, T. Kobayashi and K. Kato, *FEBS Lett.* **12,** 41 (1990).

57. S. Clark *et al.*, *Proc. Natl. Acad. Sci. USA* **85**, 4643 (1988); L. Gutierrez *et al.*, *EMBO J.* **8**, 1093 (1989).

58. C. C. Farnsworth *et al.*, *J. Biol. Chem* **264**, 20422 (1989); B. K.-K. Fung, H. Yamane, I. M. Ota and S. Clarke, *FEBS Lett* **260**, 313 (1990); P. J. Casey, P. A. Solski, C. J. Der and J. E. Buss, *Proc. Natl Acad. Sci USA* **86**, 8323 (1989).

59. Y. Fukada *et al.*, *Nature* **346**, 658 (1990).

60. H. K. Yamane *et al.*, *Proc. Natl Acad. Sci. USA* **87**, 5868 (1990); S. Mumby *et al.*, Ibid **87**, 5873 (1990).

61. C. L. Jelsema and J. Axelrod, *Proc. Natl. Acad. Sci. USA* **84**, 3623 (1987).

62. K. Okabe *et al.*, *J. Biol. Chem* **265**, 12854 (1990).

63. A. G. Gilman, *Cell* **36**, 577 (1984).

64. A. Levitzki, *FEBS Lett.* **211**, 133 (1987).

65. T. Katada *et al.*, *J. Biol. Chem.* **259**, 3568 (1984); T. Katada, M. Oinuma, M. Ui, *J. Biol. Chem.* **261**, 8182 (1986).

66. T. Katada, K. Kusakabe, M. Oinuma and M. Ui, *J. Biol. Chem.* **262**, 11897 (1987).

67. S. M. Strittmatter *et al.*, *Nature* **344**, 836 (1990).

68. R. A. Cerione *et al.*, *Biochem.* **26**, 1485 (1987); M.-J. Im *et al.*, *J. Receptor Res.* **7**, 17 (1987); R. E. Kohnken and J. D. Hildebrandt, *J. Biol. Chem.* **264**, 20688 (1989); P. J. Casey, M. P. Graziano and A. G. Gilman, *Biochem.* **28**, 611 (1989).

69. I.-H. Pang and P. C. Steinweis, *Proc. Natl Acad. Sci. USA.* **86**, 7814 (1989).

70. T. Evans *et al.*, *J. Biol Chem.* **262**, 176 (1987).

71. N. Gautam, J. Northup, H. Tamir and M. I. Simon, *Proc. Natl Acad. Sci. USA* **87**, 7973 (1990).

72. R. E. Gunderson and P. N. Devreotes, *Science* **248**, 591 (1990).

73. I. Miyajima, K. Arai and K. Matsumoto, *Mol. Cell. Biol.* **9**, 2289 (1989); D. E. Stone and S. I. Reed, *Mol. Cell Biol.* **10**, 4439 (1990).

74. Y. Zick *et al.*, *Proc. Natl Acad. Sci. USA* **83**, 9294 (1986); K. E. Carlson, L. F. Brass and D. R. Manning, *J. Biol. Chem.* **264**, 13298 (1989).

75. L. Ercolani *et al.*, *Proc. Natl Acad. Sci. USA* **87**, 4635 (1990).

76. K. Ohtsuki, T. Ikeuchi and M. Yokoyama, *Biochem. Biophys Acta* **882**, 322 (1986); N. Kimura and N. Shimada, *J. Biol. Chem.* **263**, 4647 (1988).

77. A. Yatani *et al.*, *Cell* **61**, 769 (1990).

78. J. W. Goh and P. S. Pennefather, *Science* **224**, 980 (1989).

79. A. Volterra and S. A. Siegelbaum, *Proc. Natl Acad. Sci. USA* **85**, 7810 (1988).

80. A. Kumagai *et al.*, *Cold Spring Harbor Symp. Quant. Biol* **53**, 675 (1989).

81. J. R. Hadcock, M. Ros, D. L. Watkins and C. C. Malbon, *J. Biol. Chem.* **265**, 14784 (1990).

82. K. Seuwen, I. Magnaldo and J. Pouyssegur, *Nature* **335**, 254 (1988).

83. J. Pouyssegur, C. Kahan and K. Seuwen, *Growth Factors*, V. R. Sara, Ed., (Raven Press, New York, 1990), pp. 85-100; T. Okamoto *et al.*, *Cell* **62**, 709 (1990).

84. N. Wang, K. Yar and M. M. Rasenick, *J. Biol. Chem.* **265**, 1239 (1990).

85. G. Walker and L. Y. Bourguignon, *FASEB J.* **4**, 2930 (1990).

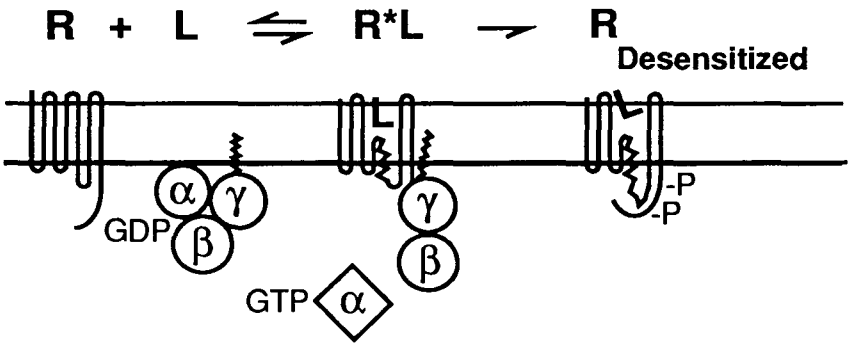
86. T. Bengtsson, E. Sarndahl, O. Stendahl and T. Andersson, *Proc. Natl. Acad. Sci. USA* **87**, 2921 (1990).

Figure 1. Receptor-G protein mediated signal transduction.

(A) Receptor (R) associates with a specific ligand (L) to stabilize an activated form of the receptor (R*), which catalyzes the exchange of GDP for GTP at a site on the alpha subunit of specific G-proteins. The βγ heterodimers may remain associated with the membrane through a 20 carbon isoprenyl modification ( ) of the gamma subunit. The receptor is desensitized by specific phosphorylation (-P).

(B) The G-protein cycle. Pertussis toxin (PTX) blocks the catalysis of GTP exchange by the receptor. Different activated alpha subunits (αGTP) and or the βγ heterodimers can interact with different effectors (E). Cholera toxin (CTX) blocks the GTPase activity of some alpha subunits, fixing them in an activated form.

A.


R + L ⇌ R*L → R
Desensitized
α
γ
GDP
β
GTP
α
γ
β
-P
-P

B.

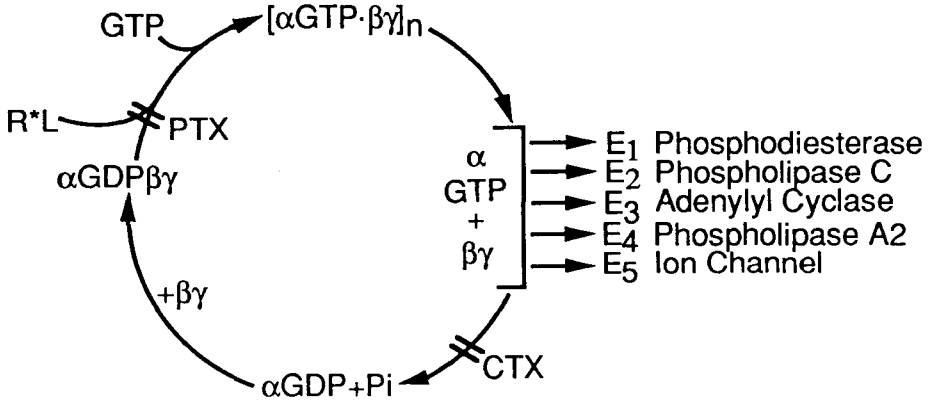

GTP
[αGTP·βγ]n
R*L
PTX
αGDPβγ
α
GTP
+
βγ
E1 Phosphodiesterase
E2 Phospholipase C
E3 Adenylyl Cyclase
E4 Phospholipase A2
E5 Ion Channel
+βγ
αGDP+Pi
CTX

Figure 2. Relationships among mammalian G protein alpha subunits. Alpha subunits are grouped by amino acid sequence identity. Branch junctions approximate values calculated for each pair of sequences. This figure extends the relationships shown by Kaziro et al. (ref. 14) and clearly defines four classes of alpha subunits. The splice variants of $G\alpha_s$ are not shown.

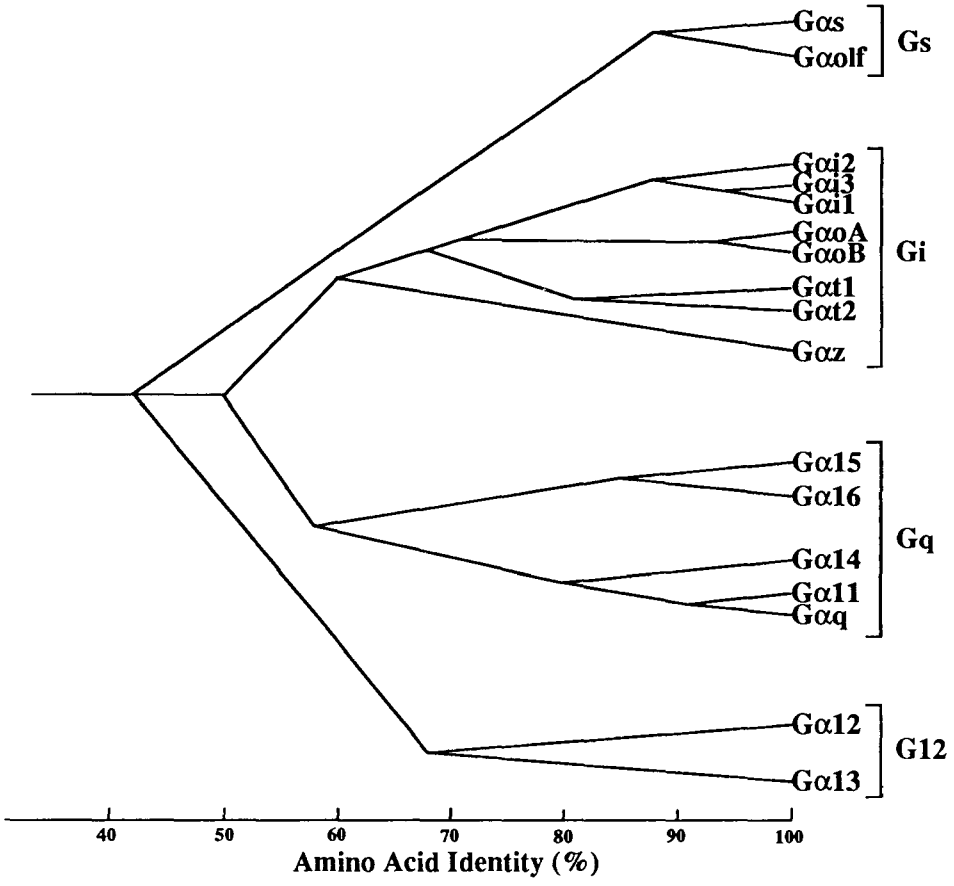
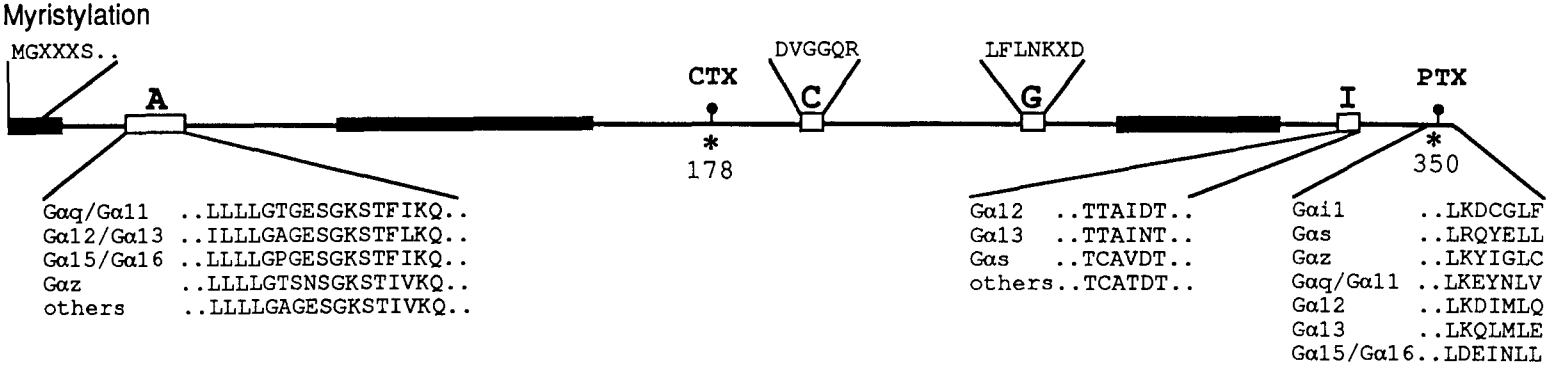

Gαs
Gαolf
Gs
Gαi2
Gαi3
Gαi1
GαoA
GαoB
Gαt1
Gαt2
Gαz
Gi
Gα15
Gα16
Gα14
Gα11
Gαq
Gq
Gα12
Gα13
G12
40
50
60
70
80
90
100
Amino Acid Identity (%)

Figure 3. The distribution of sequence specificity among the alpha subunits. The empty boxes represent the highly conserved domains found in all G-proteins. These sequences are thought to be directly involved in interaction with the guanine nucleotide (see bold letters, A, C, G, and I). The single letter amino acid code is used to show the distinctive sequences for some of the alpha subunits. The full boxes indicate the regions that show the highest levels of amino acid sequence diversity. They are at the N-terminus, between residues corresponding to amino acid 90-160 and amino acid 280-320 in the $G\alpha_{i1}$ sequence. The site for cholera toxin modification (Arginine 178), pertussis toxin modification (Cysteine 350), and myristoylation are shown.


Myristylation
MGXXXS..
A
CTX
DVGGQR
C
LFLNKXD
G
I
PTX
*
178
*
350
Gαq/Gα11 ..LLLLGTGESGKSTFIKQ..
Gα12/Gα13 ..ILLLGAGESGKSTFLKQ..
Gα15/Gα16 ..LLLLGPGESGKSTFIKQ..
Gαz ..LLLLGTSNSGKSTIVKQ..
others ..LLLLGAGESGKSTIVKQ..
Gα12 ..TTAIDT..
Gα13 ..TTAINT..
Gαs ..TCAVDT..
others..TCATDT..
Gαi1 ..LKDCGLF
Gαs ..LRQYELL
Gαz ..LKYIGLC
Gαq/Gα11 ..LKEYNLV
Gα12 ..LKDIMLQ
Gα13 ..LKQLMLE
Gα15/Gα16..LDEINLL

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 7407–7409, October 1989
Biochemistry

Diversity of the G-protein family: Sequences from five additional α subunits in the mouse

(GTP-binding protein/signal transduction/polymerase chain reaction)

MICHAEL STRATHMANN, THOMAS M. WILKIE, AND MELVIN I. SIMON

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125

Contributed by Melvin I. Simon, July 11, 1989

ABSTRACT **Biochemical analysis has revealed a number of guanine nucleotide-binding regulatory proteins (G proteins) that mediate signal transduction in mammalian systems. Characterization of their cDNAs uncovered a family of proteins with regions of highly conserved amino acid sequence. To examine the extent of diversity of the G protein family, we used the polymerase chain reaction to detect additional gene products in mouse brain and spermatid RNA that share these conserved regions. Sequences corresponding to six of the eight known G protein α subunits were obtained. In addition, we found sequences corresponding to five newly discovered α subunits. Our results suggest that the complexity of the G protein family is much greater than previously suspected.**

G proteins are components of eukaryotic signal transducing systems. They relay information from activated membrane receptors to intracellular effectors (1). Many different kinds of transmembrane receptors, such as hormone, neurotransmitter, and sensory receptors, are coupled via G proteins to a variety of effectors, including adenylyl cyclase, phospholipases, phosphodiesterases, and potassium and calcium channels (2, 3). The G proteins share a common structure; they are heterotrimers composed of α, β, and γ subunits. The G-protein α subunit binds GDP and, upon interaction with the appropriate activated receptor, exchanges the GDP for GTP. This exchange is accompanied by dissociation of the α and βγ subunits, allowing α (and perhaps βγ) to interact with effectors. An intrinsic GTPase activity restores the α subunit to its initial GDP bound state.

Multiple genes have been found to encode the G-protein subunits, with the greatest diversity thus far found among the α subunits. Molecular cloning of cDNAs resulted in the identification of eight genes encoding α subunits in mammals (4). Some of these gene products appear to be restricted to one cell type and to be involved in a single function. For example, two different genes encoding phototransducers were discovered; one is expressed in rod photoreceptors and the other is expressed in cone cells (5). Other α subunits are ubiquitous; G_s (stimulatory G protein) is found in most cells that have been examined and appears to subserve multiple functions (6).

It is not clear how many different G proteins are required to mediate the varied signaling systems in mammals. Perhaps a small number of genes encode a few G proteins, which can be adapted to multiple tasks. Alternatively, a large number of different G proteins may be utilized, with unique proteins dedicated to specific functions. We have examined G-protein diversity in the mouse to explore further how G-protein-mediated signal transduction has adapted to the complex signaling processes that define a multicellular organism.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Materials. *Thermus aquaticus* DNA polymerase (*Taq* polymerase) and polymerase chain reaction (PCR) buffer were obtained from Perkin–Elmer/Cetus (Norwalk, CT). Other enzymes were obtained from New England Biolabs (Beverly, MA). Oligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) DNA synthesizer.

PCR. cDNA was prepared from total RNA from adult mouse brain and mouse round spermatids. Spermatids were prepared by the staput gradient technique (7) and provided to us by A. Bellvé (Columbia University, New York). Reverse transcriptase reaction conditions were identical to those described by the manufacturer (cDNA synthesis system, BRL) except the oligonucleotide primer oMP20 or oMP21 was used at 5 μM in place of oligo(dT) (see Fig. 1 legend for a description of the primers). An aliquot of the reverse transcriptase reaction (cDNA synthesized from 100 ng of RNA) was used for PCR amplification (94°C, 1 min; 37°C, 1.5 min; 72°C, 2 min; 35 rounds followed by a 10-min incubation at 72°C). PCR was performed on a Cetus DNA thermal cycler. Conditions for PCR were identical to those described by the manufacturer (Gene Amp Kit, Cetus) except each primer was present at 5 μM. Reaction mixtures were electrophoresed on NuSieve agarose gels (3%), and products of the appropriate size were isolated. The fragments were treated with Klenow fragment and digested with *Eco*RI and *Bam*HI or phosphorylated with T4 polynucleotide kinase and cloned into Bluescript vector (Stratagene). Clones were sorted by single lane sequencing (see text). The complete nucleotide sequence of the cloned PCR products was obtained by sequencing double-stranded Bluescript plasmid using the M13 universal and reverse primers.

Northern Analysis. ^{32}P-labeled hybridization probes were generated by random priming of gel-purified PCR fragments (Multiprime DNA labeling system, Amersham). The fragments were amplified from plasmid DNA with the appropriate PCR primers (see Fig. 1). Blots (see Fig. 2 legend) were hybridized at 42°C in 50% formamide and washed at 55°C–60°C in 0.5× SET (0.5% sodium lauryl sulfate in 10 mM Tris·HCl/0.25 mM EDTA, pH 7.5).

RESULTS AND DISCUSSION

The amino acids thought to allow the G-protein α subunit to interact with guanine nucleotides lie in several domains (8). These domains are highly conserved among all known G proteins from yeast, *Drosophila*, and mammals (4, 9). To probe α-subunit diversity, we designed four sets of mixed oligonucleotides corresponding to two of these conserved regions for use in the PCR (10, 11) (Fig. 1). cDNA from total mouse brain and from purified mouse spermatids was amplified, and DNA fragments of the expected size were cloned.

Abbreviations: PCR, polymerase chain reaction; G protein, guanine nucleotide-binding regulatory protein.

```
          oMP41     oMP19                                                    oMP20,21
          DVGGQR    KWIHCF                                                   FLNKKD
Gs:   DVGGQRDERRKWIQCFNDVTAIIFVVASSSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTISVILFLNKQD (295)
Gi1:  DVGGQRSERKKWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKD (272)
Gi2:  DVGGQRSERKKWIHCFEGVTAIIFCVALSAYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKD (273)
Gi3:  DVGGQRSERKKWIHCFEGVTAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICNNKWFTDTSIILFLNKKD (272)
Go:   DVGGQRSERKKWIHCFEDVTAIIFCVALSGYDQVLHEDETTNRMHESLMLFDSICNNKFFIDTSIILFLNKKD (273)
Gx:   DVGGQRSERKKWIHCFEGVTAIIFCVELSGYDLKLYEDNQTSRMAESLRLFDSICNNNWFINTSLILFLNKKD (273)
Tr:   DVGGQRSERKKWIHCFEGVTCIIFIAALSAYDMVLVEDDEVNRMHESLHLFNSICNHRYFATTSIVLFLNKKD (268)
Tc:   DVGGQRSERKKWIHCFEGVTCIIFCAALSAYDMVLVEDDEVNRMHESLHLFNSICNHKFFAATSIVLFLNKKD (272)

Gα10:       DERRKWIQCFNDVTAIIYVAACSSYNMVIREDNNTNRLRESLDLFESIWNNRWLRTISIIL
Gα11:       SERRKWIHCFENVTSIMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPWFQNSSVIL
Gα12:       SQRQKWFQCFDGITSILFMVSSSEYDQVLMEDRRTNRLVESMNIFETIVNNKLFFNVSIIL
Gα13:       SERKRWFECFDSVTSILFLVSSSEFDQVLMEDRQTNRLTESLNIFETIVNNRVFSNVSIIL
Gα14:             ESVTSIIFLVALSEYDQVLAECDNENRMEESKALFRTIITYPWFLNSSVIL
CONSENSUS:  SER KWI CF  VT I F   S YD VL ED  NR  ES  LF  I N   F   S IL
```

FIG. 1. Predicted amino acid sequence from five G-protein α subunits. The amino acid sequence from the relevant region of the eight previously cloned α-subunit cDNAs are designated by standard one-letter nomenclature. The gene name is on the left and the position of the last amino acid in the sequence is numbered in parentheses on the right. Three blocks of six conserved amino acids, indicated at the top, were chosen from these sequences to synthesize completely degenerate oligonucleotide primers: **DVGGQR** [oMP41, GTCTAGAGA(CT)GT(ACGT)-GG(ACGT)GG(ACGT)CA(AG)(AC)G], **KWIHCF** [oMP19, CGGATCCAA(AG)TGGAT(ACT)CA(CT)TG(CT)TT], **FLNKKD** [oMP20, GGAATTC(AG)TC(CT)TT(CT)TT(AG)TT(ACGT)AG(AG)AA; oMP21, GGAATTC(AG)TC(CT)TT(CT)TT(AG)TT(CT)AA(AG)AA]. The oligonucleotides made to FLNKKD are in the antisense orientation. The underlined sequence at the 5′ end of each oligonucleotide contains a restriction endonuclease site to facilitate cloning the PCR products. The arrowheads above the amino acid sequence of the PCR primers indicate the direction of polymerization. Gα10, -11, -12, and -13 were cloned from PCR-amplified mouse brain cDNA by using oligonucleotide primer oMP41 paired with either oMP20 or -21 in separate reactions. Gα11 and -14 were cloned from PCR-amplified mouse spermatid cDNA by using oligonucleotide primers oMP19, -20, and -21 in the same reaction. The DNA sequence of G_s, G_o, G_z, and either G_i1 or G_i3 was obtained from at least two clones derived independently from brain and spermatid cDNA; the remaining two G_i clones were obtained from spermatid cDNA. The DNA sequence of G_s and G_i2 was identical to the mouse lymphoma cDNA sequence (12). The DNA sequence of G_o, G_z, G_i1 or G_i3, and Gα11 agreed between the corresponding clones obtained from brain and spermatid cDNA. Translation of each DNA sequence in the appropriate reading frame was identical to the published amino acid sequences for the corresponding rodent cDNA clones. The consensus amino acids were compiled from the sequence of all 13 α-subunit sequences between the DVGGQR and FLNKKD PCR primers. A consensus position was defined by amino acid identity in 11 of the 13 sequences shown. Boldface type in the α-subunit sequences indicates agreement with the consensus.

Individual clones were grouped according to the pattern obtained from a single lane of DNA sequence, and the full DNA sequence was determined for representatives of each pattern.* Newly discovered α-subunit clones were easily identified by comparison with the known α-subunit sequences since greater than half of the amino acids in the cloned region are conserved.

Fig. 1 presents the amino acid sequence of the known mammalian G-protein α subunits along with the predicted amino acid sequence of five PCR products. A consensus sequence was compiled from the most conserved amino acids in the region targeted by PCR. Each sequence contains >80% of the consensus amino acids. These α-subunit sequences are clearly distinct from the other classes of GTP binding proteins, in particular ARF (ADP ribosylation factor) and members of the *ras*-like family (13), which only share homology in some of the GTP-binding domains. While we cannot rigorously classify the newly discovered α-subunit sequences, certain features are noteworthy. Gα11 and Gα14 appear to be closely related to each other, with 92% amino acid identity in the cloned region. Also, Gα12 and Gα13 are similar to each other, with 77% amino acid identity. Gα10 is most closely related to G_s, with 85% amino acid identity. An amino acid sequence identical to that identified here as Gα10 was described independently by Jones and Reed (14). They showed that this sequence is part of an α subunit, G_{olf}, that they found activates adenylate cyclase and is primarily expressed in rat olfactory neurons.

Northern blots were done to examine the expression of mRNA corresponding to these α-subunit sequences in several tissues. Fig. 2 shows that Gα11 (Fig. 2*b*), Gα12 (Fig. 2*c*), and Gα14 (Fig. 2*e*) were expressed in all tissues examined. Although the sizes of the major transcripts on these blots were similar, they were not identical. Furthermore, Gα11 and Gα14 were distinguished by their hybridization to minor transcripts of unique size and tissue distribution. Expression of Gα10 (Fig. 2*a*) was observed primarily in brain, while Gα13 (Fig. 2*d*) was most abundant in kidney and testis. The expression patterns of Gα10 and Gα13 are distinct compared to the other α-subunit genes presented in this work and previously reported (4, 15). These experiments confirm that transcripts corresponding to the new α-subunit sequences exist and rule out the possibility that the PCR products were derived from pseudogenes by amplification of contaminating chromosomal DNA.

Table 1 shows the distribution of G-protein α-subunit sequences in the PCR fragment population from mouse brain cDNA. This distribution varies with the different PCR prim-

*The sequences reported in this paper for Gα10–14 have been deposited in the GenBank data base (accession nos. M26743, M26740, M26741, M26742, M26739, respectively).

FIG. 2. Northern blot analysis of the five newly discovered α-subunit cDNAs. Poly(A)$^+$ RNA (1 μg) from liver (L), eye (E), kidney (K), testis (T), and brain (B) was fractionated on formaldehyde agarose gels (1%) by electrophoresis and transferred to nitrocellulose. The following radioactive probes were used: Gα10 (*a*), Gα11 (*b*), Gα12 (*c*), Gα13 (*d*), Gα14 (*e*). The same filter was used for *a* and *b*, a second filter was used for *d* and *e*, and a third filter was used for *c*. The positions of 28S and 18S RNA were determined by ethidium bromide staining.

Table 1. Distribution of α-subunit sequences in PCR-amplified mouse brain cDNA

	oMP19 × -20	oMP19 × -21	oMP41 × -20	oMP41 × -21
G_s	—	—	5	—
$G_i1/3a$	2	—	13	2
$G_i1/3b$	—	—	—	—
G_i2	—	—	—	—
G_o	16	9	15	6
G_z	3	4	10	2
Gα10	—	—	2	8
Gα11	4	1	3	2
Gα12	—	—	4	—
Gα13	—	—	—	2
Gα14	—	—	—	—

The G-protein α subunits known to be expressed in brain (ref. 4; this work) are listed on the left. G_i1 and G_i3 could not be distinguished by amino acid sequence but were expected to have distinct DNA sequences. In the screen of brain PCR products, only clones of type $G_i1/3a$ were obtained, but clones of type $G_i1/3b$ were found in the screen of spermatid PCR products; one of these is presumably G_i1 and the other is G_i3. Mouse brain cDNA was amplified as described in Fig. 1 by using the combination of primers indicated above each column. Random clones were analyzed by DNA sequencing. The total number of clones corresponding to each α subunit is listed. Every cloned PCR product generated with oligonucleotide pairs oMP19 × -20 and oMP19 × -21 was an α-subunit sequence. The PCR products that were generated with the oligonucleotide pairs oMP41 × -20 and oMP41 × -21 contained α-subunit sequences in only 50% and 25% of the clones, respectively. The remaining clones contained the oligonucleotide primers, but the rest of the sequence, when translated, shared no homology with α subunits. Frequently, these clones did not contain open reading frames.

ers and probably does not reflect the mRNA distribution in the tissue but rather the properties of the PCR reaction. The data do not represent a comprehensive screen of α-subunit sequences in mouse brain. Indeed, only one of three known G_i (inhibitory G protein) sequences was recovered. Nevertheless, four new α-subunit sequences were found suggesting that the number of unique α subunits is greater than the 13 described here. Other G proteins could be involved in signaling events unique to small subpopulations of neuronal cells. Consequently, many novel α-subunit messages may be quite rare with respect to the known α-subunit messages in RNA obtained from the entire brain. Discovery of a greater variety of G proteins may simply require a more extensive search.

The diversity of the G protein family may reflect the extent to which these proteins mediate the numerous signaling processes in a multicellular organism. The list of receptors and effectors that couple through G proteins is growing rapidly. In simple organisms, these proteins have been implicated in a diverse set of responses, including cell cycle control (16, 17), the regulation of stage-specific gene expression (18), and neuromodulation (19). Given the multiplicity of cellular interactions in a complex eukaryote, there may exist a wealth of information-processing systems that exploit G-protein-mediated signal transduction.

We thank Ryn Miake-Lye, Gregg Gundersen, and Anthony Bellvé for RNA; Narasimhan Gautam for valuable discussions, George Komatsoulis for computer analysis; and Mike Lochrie and Andy Pakula for critical reading of the manuscript. This work was supported by the National Science Foundation and the National Institutes of Health.

1. Casey, P. J. & Gilman, A. G. (1988) *J. Biol. Chem.* **263**, 2577–2580.
2. Neer, E. J. & Clapham, D. E. (1988) *Nature (London)* **333**, 129–134.
3. Gilman, A. G. (1987) *Annu. Rev. Biochem.* **56**, 615–650.
4. Lochrie, M. A. & Simon, M. I. (1988) *Biochemistry* **27**, 4957–4965.
5. Lerea, C., Somers, D. E., Hurley, J. B., Klock, I. B. & Bunt-Milam, A. H. (1986) *Science* **234**, 77–80.
6. Yatani, A., Imoto, Y., Codina, J., Hamilton, S. L., Brown, A. M. & Birnbaumer, L. (1988) *J. Biol. Chem.* **263**, 9887–9895.
7. Bellvé, A. R., Cavicchia, J. C., Millette, C. F., O'Brien, D. A., Bhatnagar, Y. M. & Dym, M. (1977) *J. Cell Biol.* **74**, 68–85.
8. Masters, S. B., Stroud, R. M. & Bourne, H. R. (1986) *Protein Eng.* **1**, 47–54.
9. Provost, N. M., Somers, D. E. & Hurley, J. B. (1988) *J. Biol. Chem.* **263**, 12070–12076.
10. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
11. Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M. & Caskey, C. T. (1988) *Science* **239**, 1288–1291.
12. Sullivan, K. A., Liao, Y., Alborzi, A., Beiderman, B., Chang, F., Masters, S. B., Levinson, A. D. & Bourne, H. R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6687–6691.
13. Sewell, J. L. & Kahn, R. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4620–4624.
14. Jones, D. T. & Reed, R. R. (1989) *Science* **244**, 790–795.
15. Jones, D. T., Barbosa, E. & Reed, R. R. (1988) *Cold Spring Harbor Symp. Quant. Biol.* **53**, 349–353.
16. Dietzel, C. & Kurjan, J. (1987) *Cell* **50**, 1001–1010.
17. Miyajima, I., Nakafuku, M., Nakayama, N., Brenner, C., Miyajima, A., Kaibuchi, K., Arai, K., Kaziro, Y. & Matsumoto, K. (1987) *Cell* **50**, 1011–1019.
18. Kumagai, A., Pupillo, M., Gundersen, R., Miake-Lye, R., Devreotes, P. N. & Firtel, R. A. (1989) *Cell* **57**, 265–275.
19. Volterra, A. & Siegelbaum, S. A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7810–7814.

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 6477–6481, September 1990
Biochemistry

Alternative splicing produces transcripts encoding two forms of the α subunit of GTP-binding protein G_o

(guanine nucleotide-binding protein/signal transduction)

MICHAEL STRATHMANN, THOMAS M. WILKIE, AND MELVIN I. SIMON

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125

Contributed by Melvin I. Simon, June 4, 1990

ABSTRACT **The α subunit of the guanine nucleotide-binding protein G_o ("o" for other) is believed to mediate signal transduction between a variety of receptors and effectors. cDNA clones encoding two forms of G_o α subunit were isolated from a mouse brain library. These two forms, which we call $G_oA\alpha$ and $G_oB\alpha$, appear to be the products of alternative splicing. $G_oA\alpha$ differs from $G_oB\alpha$ over the C-terminal third of the deduced protein sequence. Both forms are predicted to be substrates for ADP-ribosylation by pertussis toxin. $G_oA\alpha$ transcripts are present in a variety of tissues but are most abundant in brain. The $G_oB\alpha$ transcript is expressed at highest levels in brain and testis. It is possible that $G_oA\alpha$ and $G_oB\alpha$ have different functions.**

Guanine nucleotide-binding proteins (G proteins) form a large family of signal-transducing molecules. They are found as heterotrimers made up of α, β, and γ subunits. Members of the G protein family have been most extensively characterized by the nature of the α subunit, which binds guanine nucleotide, is capable of hydrolyzing GTP, and interacts with specific receptor and effector molecules (for reviews, see refs. 1–3). A variety of specific G protein α subunits and their corresponding cDNAs and genes have been identified. Some of the α subunits, such as stimulatory G protein (G_s) α subunit, are ubiquitous—i.e., splice variants of the G_s α subunit gene (4) are found in every tissue that has been examined. In reconstitution experiments, G_s α subunit was shown to interact with a specific subset of receptors and to activate adenylate cyclase. Other G proteins are cell-type specific—i.e., they are found in specialized cells and appear to transduce signals from only a small subset of receptors. For example, there are two transducins, one found only in rod photoreceptor cells and the other found only in cone cells (5, 6). Still other members of the α-subunit family are restricted to a subset of tissues.

In the course of the purification of inhibitory G protein (G_i) subunits, a new G protein was discovered and named G_o ("o" for other). The G_o heterotrimer is found abundantly in bovine brain as a membrane-associated protein (7, 8). Homologues of the G_o α subunit have been found in a variety of organisms from *Drosophila* to man (2, 9–12). In general, the G_o protein is localized in neural tissue and is an abundant membrane-bound protein in brain extracts. Evidence for the presence of G_o in some other tissues has also been obtained, but the highest concentration of both protein and mRNA appears to be in brain. The purified G_o α subunit has been used to reconstitute signal transducing systems, and it has been proposed to be the G protein that mediates a variety of processes that are sensitive to pertussis toxin inhibition (see ref. 13 for review). Perhaps the clearest evidence for the specific involvement of G_o comes from studies on ion channels in neuronal cells and heart atria. There is evidence from patch-clamp studies that the GTP-bound G_o α subunit can gate potassium and calcium channels (14–17). It also has been suggested that the pertussis toxin-sensitive activation of phospholipase C may be mediated by G_o α subunit (18, 19). Recently, the GAP 43 protein in nerve growth cones was shown to stimulate guanine 5'-[γ-thio]triphosphate binding to G_o (20). Thus, there are indications that a large variety of receptors and effectors could interact with the G_o α subunit.

The cDNA corresponding to G_o α subunit has been isolated and characterized from a number of different organisms (13). The gene itself is large, containing at least 10 exons that cover >90 kilobase pairs (kbp) of DNA in humans (21). During the course of studies to ascertain the diversity of the G protein α subunit family, we discovered cDNA clones encoding a variant form of mouse brain G_o α subunit. The sequences of the two forms are identical at the N terminus but show considerable variation at the C terminus*. These two forms of mouse brain G_o α subunit, which we call $G_oA\alpha$ and $G_oB\alpha$, are likely the products of alternative splicing, and they may have different functions.

MATERIALS AND METHODS

Polymerase Chain Reaction (PCR). PCR (22) was performed as described (23). Mouse brain and spermatid cDNA were made from poly(A)$^+$ RNA with random hexanucleotide primers by using Moloney murine leukemia virus reverse transcriptase. Conditions were those supplied by the manufacturer (BRL). The oligonucleotides used for PCR amplification of the cDNA were as follows: oMP19, CGGATCCAARTGGATHCAYTGYTT; oMP20, GGAATTCRTCYTTYTTRTTNAGRAA; oMP21, GGAATTCRTCYTTYTTRTTYAARAA; CT60, CATGCACGAATCCCTGAAGC; CT112, CCGCATGCACGAGTCTCTCAT; CT113, CCCGKAGRTTKTTGGCRATGA; CT114, ATGGGATGTACGCTGAGCGCA; GO1, TCGTCCTCGTGGAGCACCTG; and Ta29, GGGATCCNGTRTCNGTNGCRCANGT, in which R = A or G, Y = C or T, K = G or T, H = A, C, or T, and N = A, C, G, or T. PCR was performed on a Perkin–Elmer/Cetus thermal cycler. During each cycle, samples were denatured for 1.0 min at 94°C, extended for 1.0 min at 72°C, and annealed for 0.5 min at the following temperatures: oMP19 + oMP20 + oMP21, at 42°C; CT60 + Ta29, at 50°C; oMP19 + Ta29, at 42°C; GO1 + CT114, at 60°C; CT60 + CT113, at 55°C; and CT112 + CT113, at 55°C. Each oligonucleotide was used in the PCR at 10 ng/μl; 35 cycles were performed on approximately 5 ng of cDNA in a 50-μl reaction volume. The buffer and *Thermus aquaticus* (*Taq*) polymerase were supplied by Cetus.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: PCR, polymerase chain reaction; G protein, guanine nucleotide-binding protein; G_i, G_s, and G_o, inhibitory, stimulatory, and "other" G proteins.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M36777 for $G_oA\alpha$ and M36778 for $G_oB\alpha$).

Northern Analysis. Total RNA and poly(A)$^+$ RNA were run on 1% agarose gels and transferred to GeneScreen (DuPont) as described (24). A probe specific to both $G_oA\alpha$ and $G_oB\alpha$ was made by PCR amplifying clone G_o11 (Fig. 1*A*) with oligonucleotides GO1 and CT114 as described above. Probes specific to $G_oA\alpha$ or $G_oB\alpha$ were made by PCR amplifying G_o11 and 718-52-4 (Fig. 1*B*) with oligonucleotide pairs CT112 + CT113 ($G_oA\alpha$) and CT60 + CT113 ($G_oB\alpha$). The amplified products were run on low-melting-point agarose gels (Seaplaque, FMC), excised, and labeled by random priming as

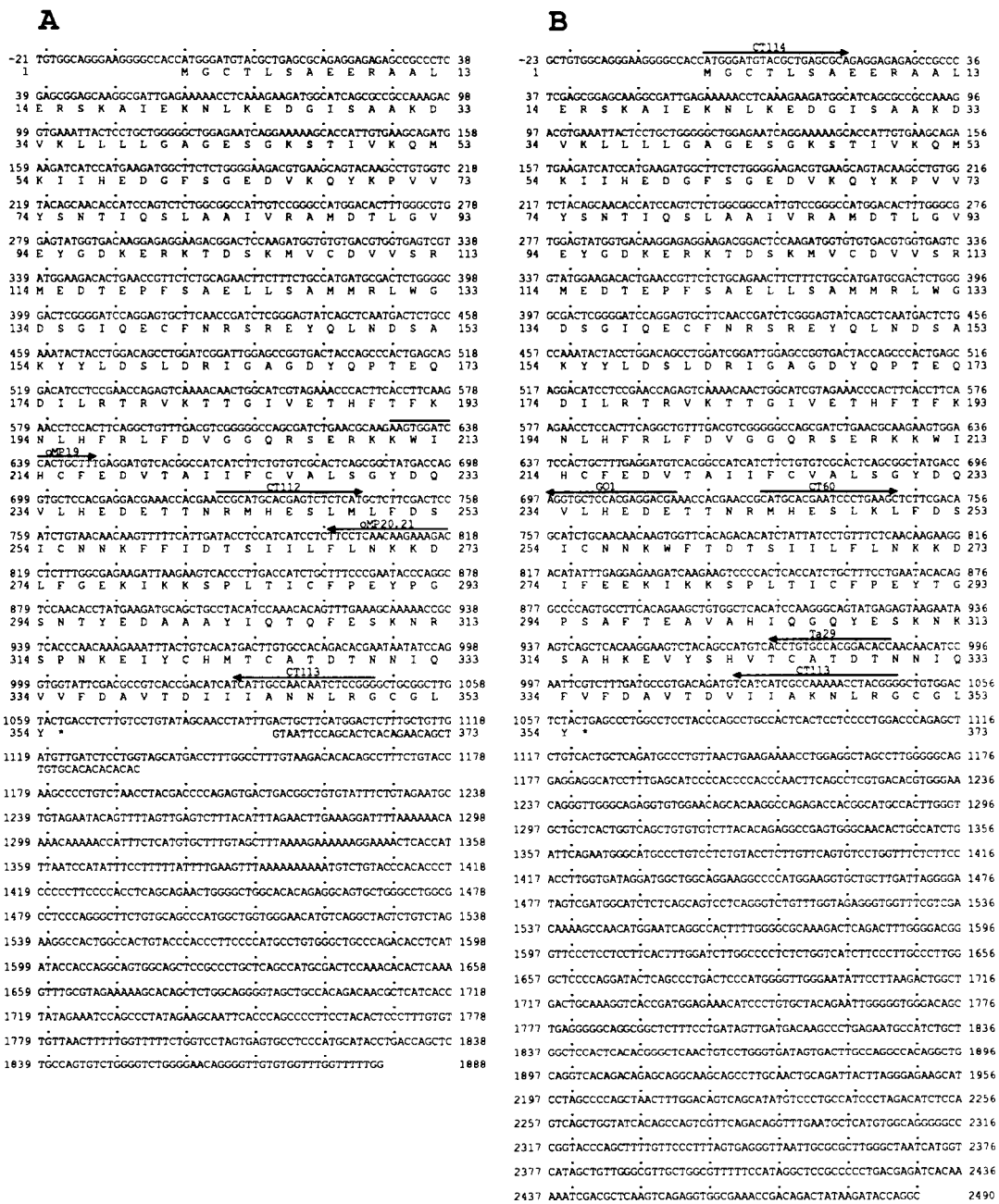


FIG. 1. Sequences of $G_oA\alpha$ (*A*) and $G_oB\alpha$ (*B*). (*A*) The sequence of $G_oA\alpha$ was obtained from two clones, G_o11 and G_o12. These clones were isolated from a random-primed mouse brain cDNA library. G_o11 extends from nucleotide −21 to 1133. G_o12 includes nucleotides 616–1888. G_o11 differs from G_o12 beginning at base 1093. These changes are shown below the G_o12 sequence. (*B*) The $G_oB\alpha$ sequence is contained in two clones, 718-52-4 and 718-52-5. These clones were isolated from the same library as the $G_oA\alpha$ clones. Clone 718-52-5 extends from nucleotide −23 to 903. Clone 718-52-4 extends from nucleotide 822 to 2490. Oligonucleotides mentioned in the text are indicated above the sequence.

described (23). The RNA blots were hybridized as described (23).

PCR Northern Analysis. PCR was performed on cDNA from total RNA as described above. The degenerate oligonucleotides oMP19, oMP20, and oMP21 were used for 30 cycles of amplification. The PCR products were electrophoresed through a 2% agarose gel, blotted to GeneScreen, and hybridized according to the manufacturer's instructions. These blots were probed with radiolabeled oligonucleotides specific to Gα11 (CTCGCTTAGTGCCACC), $G_oA\alpha$ (GAGCATGAGAGACTCG), and $G_oB\alpha$ (GAGCTTCAGGGATTCG). The oligonucleotides were end-labeled with [γ-^{32}P]-ATP as described (24). Blots were washed at room temperature twice for 5 min in 0.90 M NaCl/0.09 M sodium citrate (6× SSC)/0.1% SDS (24), twice for 5 min in 6× SSC, and finally for 1 min in 6× SSC at the melting temperature of each oligonucleotide.

Isolation of cDNAs and Nucleotide Sequencing. A random-hexanucleotide-primed mouse brain cDNA library in the cloning vector mλJSM8 (M.S. and M.I.S., unpublished data) was screened by standard techniques. Sequencing was performed by using the Sequenase kit marketed by United States Biochemical.

RESULTS

Cloning $G_oA\alpha$ and $G_oB\alpha$ cDNAs. PCR was used to screen RNA from both mouse brain and mouse spermatids for messages that correspond to different G protein α subunits (24). Degenerate oligonucleotides that code for amino acid sequences conserved in α subunits were used for amplification. These conserved amino acid sequences are shown in Fig. 1*A* (oMP19, oMP20, oMP21). The DNA sequences of clones derived from the PCR were determined. Two different classes of DNA sequences were found to encode amino acid sequences that correspond to the G_o α protein. Multiple clones of these two classes, termed PCRG-8 and PCRG-12, were obtained from both brain and spermatid cDNA. The derived amino acid sequence of PCRG-8 is identical to that found for the G_o α subunit cDNA from rat olfactory tissue (25). However, the sequence of PCRG-12 is identical to PCRG-8 only for the first 89 nucleotides and then diverges. To obtain a specific probe for the PCRG-12 gene sequence, we amplified mouse spermatid cDNA with oMP19 and another degenerate oligonucleotide, Ta29, which encodes conserved amino acids near the C terminus of α subunits (2). The PCR products were subcloned and screened with the PCRG-12-specific oligonucleotide, CT60 (Fig. 1*B*). The DNA sequence of one positive clone, 718-52, was found to be identical to PCRG-12, and it extended the sequence 168 nucleotides downstream. A hybridization probe was prepared from clone 718-52 by PCR with CT60 and Ta29 as primers. This probe and PCRG-8 were used to screen a randomly primed mouse brain cDNA library. Several cDNA clones were purified and sequenced. The complete DNA sequences of the mouse $G_oA\alpha$ cDNA and the mouse $G_oB\alpha$ cDNA are presented in Fig. 1. The amino acid sequence of mouse $G_oA\alpha$ differs from that of the rat G_o α subunit (25) at two positions (Thr-102 → Ala and Gly-166 → Ala). However, the DNA sequence of the $G_oB\alpha$ clone differs markedly from that of $G_oA\alpha$ in the region corresponding to the C-terminal third of the reading frame and the 3′ untranslated region. The DNA sequence is identical in the portion of the clone that corresponds to the N-terminal two-thirds of the protein and the 5′ untranslated region.

Another variant of the G_o α subunit cDNA clone was obtained whose sequence was identical to that of the $G_oA\alpha$ clones throughout the translated portion of the cDNA but varied from the $G_oA\alpha$ cDNA in the 3′ untranslated region of the sequence. The DNA sequence of the variable region is shown in Fig. 1*A*.

The nucleotide sequence of the $G_oA\alpha$ cDNA clone and that of the $G_oB\alpha$ cDNA clone are identical up to position 737. Thereafter some sequence conservation exists until a point corresponding to nucleotide 1064, which marks the end of the reading frame. In the 3′ untranslated region, the sequence of $G_oA\alpha$ and $G_oB\alpha$ differ dramatically from each other. It seems likely that $G_oA\alpha$ and $G_oB\alpha$ reflect different splice products derived from the same gene. Further support for this idea is provided by the intron–exon map of the G_o α subunit gene reported by Kaziro (21), which shows the same distribution of introns and exons as in the G_i α subunit gene. If the $G_oA\alpha$ and $G_oB\alpha$ gene products were the result of differential mRNA splicing, then we would expect the divergence point to be at nucleotide 723. This corresponds to an asparagine residue in both protein sequences. The first change in DNA sequence occurs at nucleotide 737 of the $G_oA\alpha$ sequence. On the basis of this interpretation, Fig. 2 shows the distribution of amino acid sequences corresponding to the nucleotide sequences of the $G_oA\alpha$ and $G_oB\alpha$ clones: after the point of divergence, there are 25 differences between the amino acid sequences, and 17 of them are concentrated within a stretch of 32 amino acids.

Distribution of G_o Gene Product. Tissue distribution of $G_oA\alpha$ and $G_oB\alpha$ mRNAs was analyzed first by Northern blots. Poly(A)$^+$ and total RNA from a variety of mouse tissues were separated by electrophoresis and hybridized to each of three probes that were obtained from the cDNA clones. These probes were made by PCR amplification of the 5′-end sequences shared by both $G_oA\alpha$ and $G_oB\alpha$ or by amplification of regions unique to either $G_oA\alpha$ or $G_oB\alpha$. Hybridization of the probe that is common to both splice products was compared with hybridization of the probes specific to $G_oA\alpha$ or $G_oB\alpha$ (Fig. 3). G_o α subunit mRNA was most abundant in brain (Fig. 3*A*), and on a longer exposure, six distinct transcripts were visible (Fig. 3*B*). These transcripts were also expressed at lower levels in testis, heart, and lung but were not observed in other tissues when assayed by Northern blot. The $G_oA\alpha$ probe specifically hybridized to the three smallest transcripts that were identified by the common probe. $G_oA\alpha$ was most abundant in brain but also was expressed in testis and heart (Fig. 3*C*). In contrast, $G_oB\alpha$ specifically hybridized to one transcript at equal intensity in brain and testis (Fig. 3*D*). On longer exposures, the largest transcripts could be seen to hybridize to both probes (data not shown) and may represent unspliced or partially spliced precursors derived from nuclear RNA. Thus, each transcript that hybridizes to the common probe is also recognized by either the GoAα- or GoBα-specific probe. The relative levels of $G_oA\alpha$ and $G_oB\alpha$ transcripts in a specific tissue can be estimated by comparing the relative intensities of the $G_oB\alpha$-specific message and the $G_oA\alpha$ transcripts (Fig. 3*B*). $G_oA\alpha$ was much more abundant than $G_oB\alpha$ in brain, whereas the levels of the two transcripts were similar in testis.

We used a technique based on PCR amplification that is more sensitive than Northern analysis (T.M.W., unpublished data) to get a clearer picture of the distribution and relative abundance of $G_oA\alpha$ and $G_oB\alpha$ mRNAs. cDNA from various tissues was amplified by PCR by using the degenerate oligonucleotides oMP19 in combination with oMP20 and oMP21. The amplified products were electrophoresed and stained with ethidium bromide. All tissues had an amplified band of the same size, indicating that they all express G protein α subunit mRNA. A second assay (Fig. 4) also was used to demonstrate that each tissue supported the PCR reaction. Southern blots of the PCR products were hybridized with an oligonucleotide probe that is specific to the G protein α subunit designated Gα11. This protein is expressed ubiquitously (24). All tissues expressed roughly the same amount of

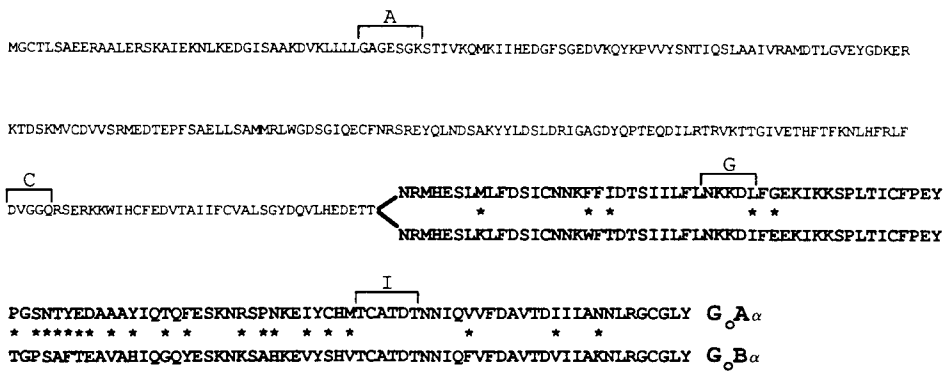


FIG. 2. Comparison of $G_oA\alpha$ and $G_oB\alpha$ amino acid sequences. Differences in amino acid residues are indicated by stars between the two sequences. A, C, G, and I represent domains that have been implicated in guanine nucleotide interactions (2).

Gα11 with the exception of uterus, which might reflect a lower level of Gα11 in this RNA source. Southern blots of the PCR products were also hybridized with oligonucletoide probes specific to either $G_oA\alpha$ or $G_oB\alpha$ (Fig. 4). $G_oA\alpha$ was most abundant in brain, followed by testis, heart, skeletal muscle, and uterus. $G_oA\alpha$ was also expressed at low levels in intestine, kidney, and lung and at exceedingly low levels in spleen and thymus; it was undetectable in liver. These cDNA samples were also amplified by PCR with oligonucleotide primers specific to $G_oA\alpha$. The results showed the same relative distribution pattern of $G_oA\alpha$ transcripts in the different tissues (data not shown). In contrast, $G_oB\alpha$ was expressed primarily in brain and testis and to a lesser extent in lung, but other tissues expressed little or no $G_oB\alpha$ mRNA. The expression pattern of $G_oA\alpha$ and $G_oB\alpha$ was corroborated by PCR analysis using a second set of independently isolated RNA samples. The relative levels of either $G_oA\alpha$ or $G_oB\alpha$ expressed in brain and testis and assayed by PCR agreed with the Northern analysis.

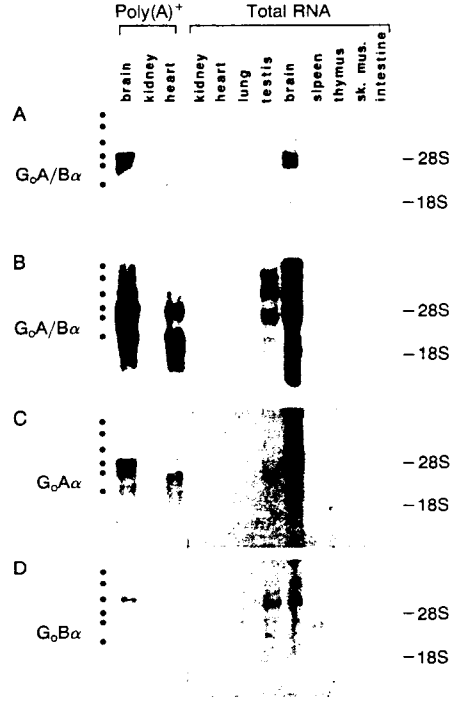


FIG. 3. Northern analysis of $G_oA\alpha$ and $G_oB\alpha$. Probes specific to both $G_oA\alpha$ and $G_oB\alpha$ [$G_o(A/B)\alpha$] (*A* and *B*), to $G_oA\alpha$ (*C*), and to $G_oB\alpha$ (*D*) were hybridized to poly(A)$^+$ and total RNA from various mouse tissues. Total RNA (20 μg) was loaded per lane. The poly(A)$^+$ RNA was loaded as follows: brain, 0.5 μg; kidney, 5 μg; and heart, 5 μg. *B* is a longer exposure of *A*.

DISCUSSION

The G_o α subunit gene is a complex locus including more than 10 exons and extending over 90 kbp of DNA in humans (21). We have demonstrated the existence in mouse of transcripts that code for two forms of the G_o α protein, $G_oA\alpha$ and $G_oB\alpha$. The sequences of these transcripts are identical for over 700 nucleotides at their 5′ ends. They diverge near an intron/exon junction, suggesting that they are products of alternative splicing. Kaziro and coworkers have now confirmed this interpretation by determining the organization of exons that represent both forms (Y. Kaziro, personal communication).

Heterogeneity of proteins antigenically related to the G_o α subunit has been observed. Goldsmith *et al.* (27) resolved on two-dimensional gels $G_oA\alpha$ and another protein from bovine brain, termed G_o^*, that reacted with four G_o α subunit-specific antisera. However, only one of the peptide antisera used in this study was raised against a sequence not shared by $G_oA\alpha$ and $G_oB\alpha$. This sequence, the carboxyl-terminal decapeptide of $G_oA\alpha$, differs from that of $G_oB\alpha$ at only one residue. Consequently, this antisera may recognize $G_oB\alpha$. In addition, G_o^* can be ADP-ribosylated by pertussis toxin. $G_oB\alpha$ is

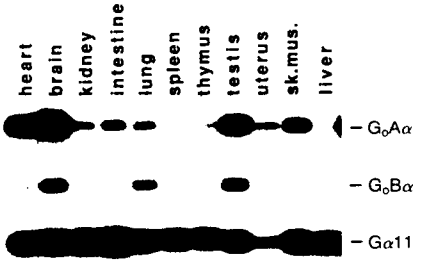


FIG. 4. PCR Northern analysis. PCR was performed on various mouse tissues with the degenerate oligonucleotides of oMP19, oMP20, and oMP21 (Fig. 1). The amplified products were hybridized with radiolabeled oligonucleotides specific to $G_oA\alpha$, $G_oB\alpha$, and Gα11.

predicted to be a substrate for pertussis toxin because the fourth residue from the C terminus is cysteine. A cysteine residue at this position is implicated as the site of ADP-ribosylation in all α subunits that are sensitive to pertussis toxin (1, 2). Thus, it is possible that G_o^* is the $G_oB\alpha$ protein.

Milligan *et al.* (28) identified, in rat myometrical membranes, a protein that reacted with two G_o α subunit-specific peptide antisera but not with an antiserum against partially purified G_o α subunit or with a peptide antiserum against $G_oA\alpha$ amino acids 22–35. Although we do not have sequence data to suggest differential splicing at the N terminus of G_o α subunit, it is interesting that the N-terminal 39 amino acids of the *Drosophila* homolog are switched by an alternative splicing mechanism (9–11).

The 3′ untranslated region of $G_oA\alpha$ also appears to undergo alternative splicing. We sequenced two clones that diverge 29 nucleotides past the stop codon (Fig. 1*A*). Murtagh *et al.* (29) found two forms of the $G_oA\alpha$ message in bovine retina that diverge 31 nucleotides past the stop codon. The significance of these alternately spliced 3′ untranslated regions is unknown. However, the untranslated regions of specific G protein α subunit isotypes are highly conserved across species (13, 30), revealing a selective pressure to maintain these sequences. Hence, the untranslated regions may affect gene regulation at the transcriptional or posttranscriptional level (31).

The functional significance of two forms of G_o α subunit is not clear. Masters *et al.* (32) have shown that the C-terminal 40% of $G_s\alpha$ subunit contains the structural elements required for specific interactions with effector and receptor. Perhaps $G_oA\alpha$ and $G_oB\alpha$ interact with different sets of receptors and/or effectors. Indeed, amino acids 315–327, by analogy to transducin, are important for receptor interactions (33). Within this region, $G_oA\alpha$ and $G_oB\alpha$ differ at 5 positions. Between amino acids 292 and 307, the two sequences differ at 11 positions. However, the contribution of this region to G protein function is not defined.

The list of genes known to encode G proteins is growing. Many novel α (24), β (ref. 34; E. von Weizsäcker and M.I.S., unpublished data), and γ (35) subunit sequences have been cloned recently. Alternative splicing provides another means to increase G protein diversity and thereby to expand the variety of G protein-mediated signaling events in a multicellular organism.

We thank Carol Lee for technical assistance, Ryn Miake-Lye for RNA, Narasimhan Gautam for discussions, and T. H. Chang for critical reading of the manuscript. We are grateful to Y. Kaziro for communicating results before publication. This work was supported by National Institutes of Health Grant GM 34236.

1. Casey, P. J. & Gilman, A. G. (1988) *J. Biol. Chem.* **263,** 2577–2580.
2. Lochrie, M. A. & Simon, M. I. (1988) *Biochemistry* **27,** 4957–4965.
3. Ross, E. M. (1989) *Neuron* **3,** 141–152.
4. Bray, P., Carter, A., Simons, C., Guo, V., Puckett, C., Kamholz, J., Spiegel, A. & Nirenberg, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8893–8897.
5. Lochrie, M. A., Hurley, J. B. & Simon, M. I. (1985) *Science* **228,** 96–99.
6. Lerea, C., Somers, D. E., Hurley, J. B., Klock, I. B. & Bunt-Milam, A. H. (1986) *Science* **234,** 77–80.
7. Sternweis, P. C. & Robishaw, J. D. (1984) *J. Biol. Chem.* **259,** 13806–13813.
8. Neer, E. J., Lok, J. H. & Wolf, L. G. (1984) *J. Biol. Chem.* **259,** 14222–14229.
9. Yoon, J., Shortridge, R. D., Bloomquist, B. T., Schneuwly, S., Perdew, M. H. & Pak, W. L. (1989) *J. Biol. Chem.* **264,** 18536–18543.
10. de Sousa, S. M., Hoveland, L. L., Yarfitz, S. & Hurley, J. B. (1989) *J. Biol. Chem.* **259,** 18544–18551.
11. Thambi, N. C., Quan, F., Wolfgang, W. I., Spiegel, A. & Forte, M. (1989) *J. Biol. Chem.* **259,** 18552–18560.
12. Schmidt, C. J., Garen-Fazio, S., Chow, Y. & Neer, E. J. (1989) *Cell Reg.* **1,** 125–134.
13. Serventi, I. M., Price, S. R., Murtagh, J. J. & Moss, J. (1990) in *ADP-Ribosylating Toxins and G Proteins*, eds. Moss, J. & Vaughan, M. (Am. Soc. Microbiol., Washington, DC), pp. 325–348.
14. Hescheler, J., Rosenthal, W., Trautwein, W. & Schultz, G. (1987) *Nature (London)* **325,** 445–447.
15. Ewald, D. A., Sternweis, P. C. & Miller, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3633–3637.
16. Harris-Warrick, R. M., Hammond, C., Paupardin-Tritsch, D., Homburger, V., Rouot, B., Bockaert, J. & Gerschenfeld, H. M. (1988) *Neuron* **1,** 27–32.
17. Logothetis, D. E., Kim, D., Northup, J. K., Neer, E. J. & Clapham, D. E. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5814–5818.
18. Banno, Y., Nagao, S., Katada, T., Nagata, K., Ui, M. & Nozawa, Y. (1987) *Biochem. Biophys. Res. Commun.* **146,** 861–869.
19. Moriarty, T. M., Padrell, E., Carty, D. J., Omri, G., Landau, E. M. & Iyengar, R. (1990) *Nature (London)* **343,** 79–82.
20. Strittmatter, S. M., Valenzuela, D., Kennedy, T. E., Neer, E. J. & Fishman, M. C. (1990) *Nature (London)* **344,** 836–841.
21. Kaziro, Y. (1990) in *ADP-Ribosylating Toxins and G Proteins*, eds. Moss, J. & Vaughan, M. (Am. Soc. Microbiol., Washington, DC), pp. 189–206.
22. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
23. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
24. Strathmann, M., Wilkie, T. M. & Simon, M. I. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7407–7409.
25. Jones, D. T. & Reed, R. R. (1987) *J. Biol. Chem.* **262,** 14241–14249.
26. Gautam, N., Baetscher, M., Aebersold, R. & Simon, M. I. (1989) *Science* **244,** 971–974.
27. Goldsmith, P., Backlund, P. S., Rossiter, K., Carter, A., Milligan, G., Unson, C. G. & Spiegel, A. M. (1988) *Biochemistry* **27,** 7085–7090.
28. Milligan, G., Tanfin, Z., Goureau, O., Unson, C. & Harbon, S. (1989) *FEBS Lett.* **244,** 411–416.
29. Murtagh, J. J., Price, S. R., Van Muers, K. P., Angus, C. W., Moss, J. & Vaughan, M. (1988) *Trans. Assoc. Am. Physicians* **101,** 235–241.
30. Bray, P., Carter, A., Guo, V., Puckett, C., Kamholz, J., Spiegel, A. & Nirenberg, M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5115–5119.
31. Yaffe, D., Nudel, U., Mayer, Y. & Neuman, S. (1985) *Nucleic Acids Res.* **13,** 3723–3737.
32. Masters, S. B., Sullivan, K. A., Miller, R. T., Beiderman, B., Lopez, N. G., Ramachandran, J. & Bourne, H. R. (1988) *Science* **241,** 448–451.
33. Hamm, H. E., Deretic, D., Arendt, A., Hargrave, P. A., Koenig, B. & Kofmann, K. P. (1988) *Science* **241,** 832–835.
34. Levine, M. A., Smallwood, P. M., Moen, P. T., Jr., Helman, L. J. & Ahn, T. G. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2329–2333.

Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 9113–9117, December 1990
Biochemistry

G protein diversity: A distinct class of α subunits is present in vertebrates and invertebrates

(GTP-binding protein/signal transduction/*Drosophila*)

MICHAEL STRATHMANN AND MELVIN I. SIMON

Division of Biology 147-75, California Institute of Technology, Pasadena, CA 91125

Contributed by Melvin I. Simon, August 21, 1990

ABSTRACT **Heterotrimeric guanine nucleotide-binding proteins (G proteins) are integral to the signal transduction pathways that mediate the cell's response to many hormones, neuromodulators, and a variety of other ligands. While many signaling processes are guanine nucleotide dependent, the precise coupling between a variety of receptors, G proteins, and effectors remains obscure. We found that the family of genes that encode the α subunits of heterotrimeric G proteins is much larger than had previously been supposed. These novel alpha subunits could account for some of the diverse activities attributed to G proteins. We have now obtained cDNA clones encoding two murine α subunits, $G\alpha_q$ and $G\alpha 11$, that are 88% identical. They lack the site that is ordinarily modified by pertussis toxin and their sequences vary from the canonical Gly-Ala-Gly-Glu-Ser (GAGES) amino acid sequence found in most other G protein α subunits. Multiple mRNAs as large as 7.5 kilobases hybridize to $G\alpha_q$ specific probes and are expressed at various levels in many different tissues. $G\alpha 11$ is encoded by a single 4.0-kilobase message which is expressed ubiquitously. Amino acid sequence comparisons suggest that $G\alpha_q$ and $G\alpha 11$ represent a third class of α subunits. A member of this class was found in *Drosophila melanogaster*. This α subunit, $DG\alpha_q$, is 76% identical to $G\alpha_q$. The presence of the G_q class in both vertebrates and invertebrates points to a role that is central to signal transduction in multicellular organisms. We suggest that these α subunits may be involved in pertussis toxin-insensitive pathways coupled to phospholipase C.**

The G proteins are a family of guanine nucleotide-binding proteins that relay signals from cell surface receptors to intracellular effectors. Members of this family are heterotrimers composed of α, β, and γ subunits. The α subunit is believed to confer receptor and effector specificity on the heterotrimer. When the G protein is activated by interaction with receptor, the α subunit exchanges bound GDP for GTP. The intrinsic GTPase activity of the α subunit restores it to the basal state in which GDP is bound. This form of signal transduction is basic to the mechanisms that cells use in responding to hormones, neurotransmitters, and a variety of other ligands (for reviews see refs. 1–3). The process is highly conserved in evolution. Indeed, G proteins are central to intercellular communication among even simple eukaryotes. For example, G proteins are involved in the yeast mating-type pathway (4, 5), and several G protein α subunits are differentially expressed during development in the slime mold *Dictyostelium discoideum* (6, 7).

We are interested in how G protein-mediated signal transduction has adapted to the diverse signaling requirements of complex multicellular organisms. Reconstitution studies and the use of pertussis and cholera toxins to modify specific G protein α subunits have demonstrated the involvement of the $G\alpha_s$ and $G\alpha_i$ subtypes in gating of specific ion channels (8, 9) and in the regulation of adenylyl cyclase in a variety of organisms (10). In the highly specialized visual system in mammals, biochemical experiments have led to the elucidation of the role of $G\alpha_{t1}$ (rod transducin) in regulating phosphodiesterase and subsequently in controlling the levels of cyclic GMP (11). There are, however, many processes that are refractory to toxin inhibition but nonetheless appear to be mediated by guanine nucleotide-binding proteins (12, 13). To understand the extent of involvement of the G protein system and the nature of the specificity required for function, we have examined the diversity of the G protein family in complex organisms.

Recently, we developed an approach involving the polymerase chain reaction (PCR, ref. 14) to detect novel sequences that share highly conserved domains common to all G protein α subunits. We found evidence for extensive diversity in the mammalian G protein α subunit family (15). A small screen uncovered five novel sequences termed $G\alpha 10$ through $G\alpha 14$. In this paper we present the cDNA sequences of two α subunits that define another class of G proteins. This class, termed G_q, is distinguished by amino acid sequence homology and includes $G\alpha_q$, $G\alpha 11$, and $G\alpha 14$. We argue that the G_q class appeared early in evolution; it is found in both vertebrates and invertebrates. We present the sequence of a member of the G_q class in *Drosophila*.*

MATERIALS AND METHODS

PCR. PCR was performed as described previously (15). cDNA was made from poly(A)$^+$ RNA with random hexanucleotide primers by using reverse transcriptase from Moloney murine leukemia virus. Conditions were those described by the supplier (BRL). The oligonucleotides used for PCR amplification of the cDNA were as follows:

oMP19, CGGATCCAARTGGATHCAYTGYTT;
oMP20, GGAATTCRTCYTTYTTRTTNAGRAA;
oMP21, GGAATTCRTCYTTYTTRTTYAARAA;
GQ112, CTGAAGGAGTACAAYCTGGT;
GQ3, GACACGAATGACAGGAGTGCT; and
G113, CCTCAAGCCACATTGAGTCA,

in which R = A or G, Y = C or T, H = A, C, or T, and N = A, C, G, or T. PCR was performed on a Perkin Elmer Cetus thermal cycler. During each cycle, samples were denatured for 1.0 min at 94°C, extended for 1.0 min at 72°C, and annealed for 0.5 min at the following temperatures: oMP19 + oMP20 + oMP21, 42°C; GQ112 + GQ3, 53°C; and GQ112 + G113, 53°C. Each oligonucleotide was used in the PCR at 10 ng/μl. Thirty-five cycles were performed on approximately 5

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PCR, polymerase chain reaction.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M55412 for $G\alpha_q$ and M55411 for $G\alpha 11$.

ng of cDNA in a 50-μl reaction volume. The buffer and *Taq* polymerase were supplied by Cetus.

Northern Analysis. Total RNA and poly(A)$^+$ RNA were run on 1% agarose gels and transferred to GeneScreen (DuPont) as described (15). Probes specific to $G\alpha_q$ or $G\alpha 11$ 3′ untranslated regions were made by PCR amplifying the cDNA clones Gq7 and G23 (Fig. 1) with oligonucleotide pairs GQ112 + GQ3 ($G\alpha_q$) and GQ112 + G113 ($G\alpha 11$). The amplified products were electrophoresed on low melting point agarose gels (Seaplaque, FMC), excised, and labeled by random priming as described (15). The RNA blots were hybridized as described (15).

PCR Northern. PCR was performed on cDNA prepared with reverse transcriptase (supplied by BRL) from total RNA as described above. The degenerate oligonucleotides oMP19, oMP20, and oMP21 were used for 35 cycles of amplification. The PCR products were electrophoresed through a 2% agarose gel, blotted to GeneScreen, and hybridized according to the manufacturer's instructions. These blots were probed with radiolabeled oligonucleotides specific to $G\alpha_q$ (ATTCGCTAAGCGCTACTAGA) and $G\alpha 11$ (CTCGCTTAGTGCCACC). The oligonucleotides were end-labeled with [γ-^{32}P]ATP as described (ref. 16, p. 10.59). Blots were washed at room temperature twice for 5 min in 0.90 M NaCl/0.90 M sodium citrate (6× SSC)/0.1% SDS, twice for 5 min in 6× SSC, and finally they were washed for 1 min in 6× SSC at the melting temperature of each oligonucleotide.

Isolation of cDNAs. The G69 cDNA clone (Fig. 1*A*) was isolated from an oligo(dT)-primed mouse brain library in the λ cloning vector mλJSM8 (unpublished results); 10^5 clones were screened at a density of 10^4 plaques per 150-mm plate. Nitrocellulose filters (Millipore) were prepared by standard techniques (ref. 16, p. 12.30). Appropriate restriction enzymes were used to excise the coding regions of cDNA clones encoding rat $G\alpha_s$, $G\alpha_{i1}$, $G\alpha_{i2}$, $G\alpha_{i3}$, and $G\alpha_o$; bovine $G\alpha_{t1}$ and $G\alpha_{t2}$; and human $G\alpha_z$. These cDNAs were com-

A

```
                                                    CGCGCCGGCGGGGGCTGCAGCGAGGCACTTCGGAAGA -1
ATGACTCTGGAGTCCATCATGGCGTGCTGCCTGAGCGAGGAGGCCAAGGAAGCCCGGAGG 59
M  T  L  E  S  I  M  A  C  C  L  S  E  E  A  K  E  A  R  R   20
ATCAACGACGAGATCGAGCGGCACGTGCGCAGGGACAAGCGCGACGCCCGCGGGAGCTC 119
I  N  D  E  I  E  R  H  V  R  R  D  K  R  D  A  R  R  E  L   40
AAGCTGCTGCTGCTGGGGACAGGGGAGAGTGGCAAGAGCACCTTCATCAAGCAGATGAGG 179
K  L  L  L  L  G  T  G  E  S  G  K  S  T  F  I  K  Q  M  R   60
ATCATCCACGGGTCGGGCTACTCTGACGAAGACAAGCGCGGCTTCACCAAGCTGGTGTAT 239
I  I  H  G  S  G  Y  S  D  E  D  K  R  G  F  T  K  L  V  Y   80
CAGAACATCTTCACGGCCATGCAGGCCATGATCAGAGCGATGGACACGCTCAAGATCCCA 299
Q  N  I  F  T  A  M  Q  A  M  I  R  A  M  D  T  L  K  I  P  100
TACAAGTATGAACACAATAAGGCTCATGCACAATTGGTTCGAGAGGTTGATGTGGAGAAG 359
Y  K  Y  E  H  N  K  A  H  A  Q  L  V  R  E  V  D  V  E  K  120
GTGTCTGCTTTTGAGAATCCATATGTAGATGCAATAAAGAGCTTGTGGAATGATCCTGGA 419
V  S  A  F  E  N  P  Y  V  D  A  I  K  S  L  W  N  D  P  G  140
ATCCAGGAGTGCTACGACAGACGACGGGAATATCAGTTATCTGACTCTACCAAATACTAT 479
I  Q  E  C  Y  D  R  R  R  E  Y  Q  L  S  D  S  T  K  Y  Y  160
CTGAATGACTTGGACCGTGTAGCCGACCCTTCCTATCTGCCTACACAACAAGACGTGCTT 539
L  N  D  L  D  R  V  A  D  P  S  Y  L  P  T  Q  Q  D  V  L  180
AGAGTTCGAGTCCCCACTACAGGGATCATCGAATACCCCTTTGACTTACAAAGTGTCATT 599
R  V  R  V  P  T  T  G  I  I  E  Y  P  F  D  L  Q  S  V  I  200
TTCAGAATGGTCGATGTAGGGGGCCAAAGGTCAGAGAGAAGAAAATGGATACACTGCTT 659
F  R  M  V  D  V  G  G  Q  R  S  E  R  R  K  W  I  H  C  F  220
GAAAATGTCACCTCCATCATGTTTCTAGTAGCGCTTAGCGAATATGATCAAGTTCTTGTG 719
E  N  V  T  S  I  M  F  L  V  A  L  S  E  Y  D  Q  V  L  V  240
GAGTCAGACAATGAGAACCGCATGGAGGAGAGCAAAGCACTCTTTAGAACAATTATCACC 779
E  S  D  N  E  N  R  M  E  E  S  K  A  L  F  R  T  I  I  T  260
TACCCCTGGTTCCAGAACTCCTCTGTGATTCTGTTCTTAAACAAGAAAGATCTTCTAGAG 839
Y  P  W  F  Q  N  S  S  V  I  L  F  L  N  K  K  D  L  L  E  280
GAGAAAATCATGTATTCCCACCTAGTCGACTACTTCCCAGAATATGATGGACCCCAGAGA 899
E  K  I  M  Y  S  H  L  V  D  Y  F  P  E  Y  D  G  P  Q  R  300
GATGCCCAGGCAGCTCGAGAATTCATCCTGAAAATGTTCGTGGACCTGAACCCCGACAGT 959
D  A  Q  A  A  R  E  F  I  L  K  M  F  V  D  L  N  P  D  S  320
GACAAAATCATCTACTCCCACTTCACGTGCGCCACAGATACCGAGAACATCCGCTTCGTC 1019
D  K  I  I  Y  S  H  F  T  C  A  T  D  T  E  N  I  R  F  V  340
TTTGCAGCCGTCAAGGACACCATCCTGCAGCTGAACCTGAAGGAGTACAATCTGGTCTAA 1079
F  A  A  V  K  D  T  I  L  Q  L  N  L  K  E  Y  N  L  V  *  360
CCGTGCCTCCCAGAAACTGGTTCTTCCCTCCCTGTGGGTTGTTGAAGATAAACAAGAGGG 1139
ACTGTATTTCTGCGGAAAACAATTTGCATAATACTAATTTATTGCCGTCCTGGACTCTGT 1199
GAGCGTGTCCACAGAGCTGTAGTAAATATTATGATTTTATTTAAACTATTCAGAGGAAAC 1259
AGGATGCTGAAGTACAGTCCCAGCACATTTCCTCTTTTTTTTTTTTTAGGCAAACCTTG 1319
ATGTATTTTAAATTTTCAGTCATTCACTCACAGTATAAAAGCACTCCTGTCATTCGTGTC 1379
CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCATTCC                    1423
```

B

```
TTCGGGGAACCGGGAAGAGGTAGGGCCGGGCCGGCGGCAGCCGGAGGACCGCGACG   -1
ATGACTCTGGAGTCCATGATGGCGTGTTGCCTGAGCGACGAGGTGAAGGAGTCGAAGCGC  59
M  T  L  E  S  M  M  A  C  C  L  S  D  E  V  K  E  S  K  R   20
ATCAACGCGGAGATCGAGAAACAGCTGCGGCGGGACAAGCGCGACGCCCGGCGCGAGCTC 119
I  N  A  E  I  E  K  Q  L  R  R  D  K  R  D  A  R  R  E  L   40
AAGCTGCTGCTACTTGGCACTGGCGAGAGCGGGAAGAGTACCTTCATCAAGCAGATGCGC 179
K  L  L  L  L  G  T  G  E  S  G  K  S  T  F  I  K  Q  M  R   60
ATCATCCACGGGGCCGGCTACTCGGAGGAGGACAAGCGCGGCTTCACCAAGTTGGTGTAC 239
I  I  H  G  A  G  Y  S  E  E  D  K  R  G  F  T  K  L  V  Y   80
CAGAACATCTTTACCGCCATGCAGGCCATGGTGCGCGCCATGGAGACGCTCAAGATCCTC 299
Q  N  I  F  T  A  M  Q  A  M  V  R  A  M  E  T  L  K  I  L  100
TACAAGTATGAGCAGAACAAGGCCAATGCACTCCTGATCCGGGAGGTCGATGTGGAGAAG 359
Y  K  Y  E  Q  N  K  A  N  A  L  L  I  R  E  V  D  V  E  K  120
GTCACAACTTTTGAGCACCAGTATGTGAATGCCATCAAGACGCTGTGGAGTGACCCTGGT 419
V  T  T  F  E  H  Q  Y  V  N  A  I  K  T  L  W  S  D  P  G  140
GTCCAGGAGTGTTACGATCGCAGGCGGGAGTTCCAGCTATCTGACTCGGCTAAGTACTAC 479
V  Q  E  C  Y  D  R  R  R  E  F  Q  L  S  D  S  A  K  Y  Y  160
TTGACGGACGTGGACCCCATCGCCACAGTAGGCTACCTGCCCACCCAGCAGGATGTGCTG 539
L  T  D  V  D  R  I  A  T  V  G  Y  L  P  T  Q  Q  D  V  L  180
CGGGTACGCGTGCCCACCACTGGCATCATCGAGTACCCGTTTGACCTGGAGAACATCATC 599
R  V  R  V  P  T  T  G  I  I  E  Y  P  F  D  L  E  N  I  I  200
TTCAGGATGGTGGATGTGGGCGGGCAGAGGTCAGAGCGCAGGAAGTGGATCCATTGCTTT 659
F  R  M  V  D  V  G  G  Q  R  S  E  R  R  K  W  I  H  C  F  220
GAGAACGTGACTTCCATCATGTTCTTGGTGGCACTAAGCGAGTATGACCAACTCCTGGTG 719
E  N  V  T  S  I  M  F  L  V  A  L  S  E  Y  D  Q  V  L  V  240
GAGTCAGACAATGAGAACCGCATGGAGGAGAGCAAGGCCCTGTTCCGCACAATCATCACC 779
E  S  D  N  E  N  R  M  E  E  S  K  A  L  F  R  T  I  I  T  260
TACCCCTGGTTCCAGAACTCGTCTGTCATTCTCTTCCTCAACAAGAAGGACCTTCTAGAA 839
Y  P  W  F  Q  N  S  S  V  I  L  F  L  N  K  K  D  L  L  E  280
GACAAGATCCTGCACTCACACTTGGTCGATTACTTCCCTGAGTTTGATGGGCCACAGAGG 899
D  K  I  L  H  S  H  L  V  D  Y  F  P  E  F  D  G  P  Q  R  300
GATGCACAGGCCGCACGCGAGTTCATTCTGAAGATGTTTGTGGACCTGAACCCTGACAGC 959
D  A  Q  A  A  R  E  F  I  L  K  M  F  V  D  L  N  P  D  S  320
GACAAAATCATCTACTCCCACTTCACATGTGCCACCGACACCGAGAACATCCGCTTTGTG 1019
D  K  I  I  Y  S  H  F  T  C  A  T  D  T  E  N  I  R  F  V  340
TTCGCAGCTGTGAAGGACACCATCCTGCAGCTGAACCTGAAGGAGTACAACCTGGTCTGA 1079
F  A  A  V  K  D  T  I  L  Q  L  N  L  K  E  Y  N  L  V  *   359
CCGCGGCATGGCTGCCACCAGGGCCAGCTTCCTCCGCATCCGGTCACCCAAACCGAGGGA 1139
CCTGGGCACACGGGGGCTGGGGCACATCTCTCTTCCCGCTCACTTCCTGTCCCTGTCCCT 1199
CTTCCTCAAGGGCAGAGTGGCCTCATTTCGGCCTTGACTCAATGTGGCTTGAGGTTTTTT 1259
TTTTTCTTAGAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA             1307
```

C

```
Gαq  MTLESIMACCLSEEAK..(36)
Gα11 MTLESMMACCLSDEVK..
Gαil        MGCTLSAEDK..
Gαo         MGCTLSAEER..
```

D

```
Gαq/Gα11..LLLLGTGESGKSTFIKQ..(58)
Gαz      ..LLLLGTSNSGKSTIVKQ..
other Gα..LLLLGAGESGKSTIVKQ..
```

E

```
Gαq/Gα11..QLNLKEYNLV (359)
Gαil     ..KNNLKDCGLF
Gαz      ..QNNLKYIGLC
Gαs      ..RMHLRQYELL
```

FIG. 1. Sequence of $G\alpha_q$ and $G\alpha 11$. The sequence of $G\alpha_q$ (*A*) was obtained from two clones, Gq3 and Gq7. These clones were isolated from a random-primed mouse brain cDNA library by using G69 as a probe. These clones contain the following nucleotides: Gq3, −36 to 966; Gq7, 144 to 1423; and G69, 120 to 629. The $G\alpha 11$ sequence (*B*) is contained in two clones, G23 and G11-5′-9. G23, which extends from nucleotide 490 to nucleotide 1307, was isolated from an oligo(dT)-primed mouse brain cDNA library by using the $G\alpha 11$ PCR fragment (15) as a probe. G11-5′-9 was isolated from the random-primed library mentioned above. This clone extends from nucleotide −56 to nucleotide 957. The N termini (*C*), GTP-binding domains (*D*, ref. 2), and the C termini (*E*) of several α subunits are compared. The position of the last amino acid in each domain of $G\alpha_q$ is given in parenthesis.

bined in equimolar ratios and labeled by random priming as described (15). This probe was hybridized to the filters at 50°C in 0.90 M NaCl/6 mM EDTA/60 mM NaH_2PO_4 (6× SSPE)/0.1% SDS/5× Denhardt's solution containing denatured salmon sperm DNA at 100 μg/ml (1× Denhardt'solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone). The filters were washed in 6× SSC/0.1% SDS three times for 5 min at room temperature and once at 50°C for 15 min. To obtain the entire sequence of $G\alpha_q$ and Gα11, a random hexanucleotide or oligo(dT)-primed mouse brain cDNA library in mλJSM8 or λZAPII (Stratagene) was screened by using standard techniques. $DG\alpha_q$ was cloned from a *Drosophila melanogaster* head-specific library (see *Results*) kindly provided by B. Hamilton (Biology Division, California Institute of Technology, Pasadena, CA).

RESULTS

Nucleotide Sequence of $G\alpha_q$ and Gα11. The clone G69, which encodes part of $G\alpha_q$, was isolated from a mouse brain cDNA library by low-stringency hybridization. It cross-hybridized to a probe consisting of a mixture of cDNAs encoding all the previously known G protein α subunits (see *Materials and Methods*). Further screening yielded overlapping cDNA clones containing the entire $G\alpha_q$ coding sequence (Fig. 1*A*). The cloned PCR product corresponding to Gα11, which included 180 base pairs of sequence (15), was used as a probe to screen a mouse brain cDNA library. Several clones were purified and sequenced. Fig. 1*B* shows the sequence of Gα11 derived from overlapping cDNA fragments. The deduced amino acid sequence of Gα11 is 88% identical to that of $G\alpha_q$. Almost all of the amino acid differences between $G\alpha_q$ and Gα11 are concentrated in the N-terminal half of the protein. Of 42 amino acid differences, 38 are found in the N-terminal region composed of amino acids 1–200, while there are only 4 amino acid changes in a stretch of polypeptide encompassed by amino acids 201–359.

When $G\alpha_q$ and Gα11 are compared with the other α subunits, a number of noteworthy differences emerge. Fig. 1*C* shows a comparison of the N-terminal sequences of $G\alpha_q$ and Gα11 with those found for other α subunits. The methionine predicted by homology to be the first codon in $G\alpha_q$ and Gα11 is preceded by other methionines in frame. The six additional amino acids found in these two α subunits are highly conserved, suggesting that they are functionally significant. The nucleotide sequences of the two cDNAs diverge upstream of these codons, indicating that there may be no further extension of this reading frame. Also in this 5′ region, the Gα11 cDNA contains a stop codon in frame with the downstream coding sequence

The N termini of some α subunits are *N*-myristoylated on a glycine at the second position (Gly-2) (17). On the basis of their deduced amino acid sequences, $G\alpha_q$ and Gα11 are not substrates for myristoylation (Fig. 1*C*); this may affect their membrane association properties. However, $G\alpha_s$ is not myristoylated, yet this α subunit is membrane associated. Perhaps other forms of post-translational modification will prove to be responsible for anchoring these hydrophilic proteins to the membrane.

Fig. 1*D* compares the amino acid sequences in the region of the "GAGE-box." This domain is highly conserved among all α subunits and has been implicated in GTP binding. Mutations in this region affect the GTPase activity of the α subunit (18, 19). $G\alpha_z$ [also named $G\alpha_x$ (20, 21)] differs in this region from the other α subunits. The slow rates of guanine nucleotide exchange and GTP hydrolysis exhibited by $G\alpha_z$ may be due in part to these sequence changes (22). $G\alpha_q$ and Gα11 show differences in this region as well (Fig. 1*D*). Consequently, these α subunits may display unusual kinetic properties when the proteins are characterized (see *Discussion*).

The C-terminal region of the α subunit is of considerable interest. The cysteine residue that lies four amino acids from the end of most mammalian α subunits can be ADP-ribosylated by pertussis toxin, thereby inactivating the G protein. Among the known α subunits, only $G\alpha_s$, $G\alpha_{olf}$, and $G\alpha_z$ lack this cysteine. These proteins are refractory to modification by pertussis toxin (2, 22). Fig. 1*E* shows that $G\alpha_q$ and Gα11 also lack the cysteine residue at this position. Thus it is likely that the proteins corresponding to these cDNA clones will be insensitive to pertussis toxin modification.

Distribution of $G\alpha_q$ and Gα11. To determine the relative size of the mRNAs corresponding to $G\alpha_q$ and Gα11 and their distributions, Northern hybridization was performed with RNA purified from a variety of mouse tissues. Both $G\alpha_q$ and Gα11 are ubiquitously expressed (Fig. 2). Using a probe specific to the 3′ untranslated region of each cDNA clone, we found that $G\alpha_q$ has multiple messages, whereas Gα11 shows a single band of approximately 4 kb. The three largest $G\alpha_q$ transcripts (approximately 5, 6, and 7.5 kb) were evident (Fig. 2*A*) in other RNA preparations, suggesting that they are not the products of a single degraded message (data not shown). The very large size of the $G\alpha_q$ message and the variation in signal strength in some of the tissues led us to use another method to probe for the presence of $G\alpha_q$ and Gα11 specific RNA. This technique, termed PCR Northern analysis (T. M. Wilkie and M.I.S., unpublished), utilizes oligonucleotide primers designed to amplify a pool of α subunits by PCR. Specific sequences in the pool are detected by oligonucleotide hybridization. The Gα11 message (Fig. 2*C*) was found to be expressed in all tissues that were examined. The $G\alpha_q$ message (Fig. 2*C*) was also found in all tissues, although the relative levels appeared to be lower in RNA samples derived from intestine and testes and higher in brain and lung.

Analysis of the Relationships Among the G Protein α Subunits. Itoh *et al.* (23) have described the relationships between the α subunits on the basis of amino acid identity. These comparisons suggest evolutionary relationships that may exist among the G proteins. An expansion of this analysis

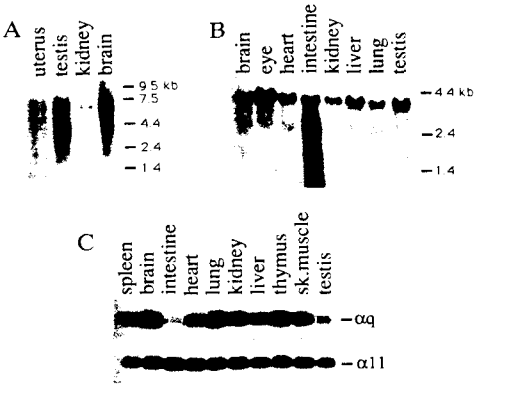


FIG. 2. Northern and PCR Northern analysis of $G\alpha_q$ and Gα11. Probes specific to $G\alpha_q$ (*A*) and Gα11 (*B*) were hybridized to poly(A)$^+$ RNA (*A*) and total RNA (*B*) from various mouse tissues. Loads were 20 μg of total RNA per lane and 4 μg of poly(A)$^+$ RNA per lane; however, ethidium bromide staining revealed that the amount of ribosomal RNA contamination varied among the different tissues. kb, Kilobases. PCR Northern analysis (*C*) was performed on various mouse tissues by using the degenerate oligonucleotides oMP19, oMP20, and oMP21. The amplified products were hybridized with radiolabeled oligonucleotides specific to $G\alpha_q$ and Gα11.

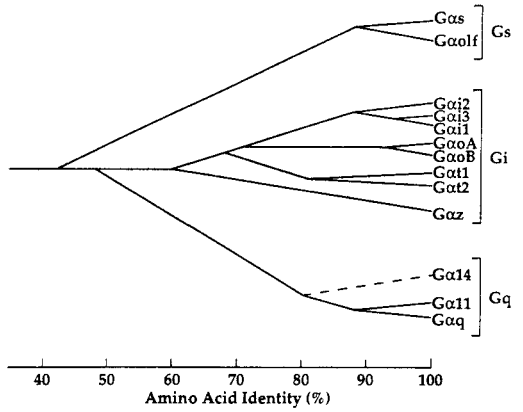


FIG. 3. Relationships among mammalian G protein α subunits. The α subunits are grouped by amino acid identity (2). Branch junctions approximate values calculated for each pair of sequences. $G\alpha_o$A and $G\alpha_o$B represent slice variants of the $G\alpha_o$ gene (24, 25). A broken line extends from Gα14 because the entire sequence of this α subunit is not available.

reveals three classes of α subunits: G_s, G_i, and G_q (Fig. 3). Within a class, α subunits display not only primary sequence relationships but also, to some extent, functional similarities. Thus in reconstitution studies some α subunits show significant levels of "crosstalk." For example, $G\alpha_{olf}$ can substitute for $G\alpha_s$ in coupling the β-adrenergic receptor to adenylyl cyclase (26). Also, the three $G\alpha_i$ subtypes open atrial potassium channels (27). The G_i class includes all the known α subunits that are susceptible to pertussis toxin modification.

$G\alpha_q$ and Gα11 form a separate class (Fig. 3). They show less than 50% amino acid sequence identity with any of the other α subunits. It is likely that $G\alpha_q$ and Gα11 will display significant crosstalk, since they differ at only four residues over the C-terminal 144 amino acids. This C-terminal domain is thought to contain the structural elements required for specific interactions with effector and receptor (28). Comparison of these amino acid sequences with the partial sequences uncovered in the PCR screen described earlier (15) reveals that Gα14 is a member of the G_q class. Gα12 and Gα13, however, appear to define yet another class of α subunits.

Representation of the G_q Class in *Drosophila*. The ubiquitous tissue distribution of $G\alpha_q$ and Gα11 is consistent with a signal transducing role that is basic to a variety of cell types and signaling pathways. If members of the G_q class are involved in a central "housekeeping" pathway, then this class should be represented not only in mammals but also in other distantly related organisms. Homologues of $G\alpha_s$, $G\alpha_i$, and $G\alpha_o$ have been found in a variety of organisms, including *Drosophila* (29–34). Fig. 3 suggests that the G_q class may have diverged from the G_i class before two members of the G_i class, G_o and $G\alpha_i$, diverged from each other. Thus it seemed possible that *Drosophila* would have a $G\alpha_q$ homologue. These arguments encouraged us to screen *D. melanogaster* α subunits for a member of the G_q class. We used the same PCR technique described previously (15) with mixed degenerate oligonucleotides corresponding to two highly conserved amino acid hexamer sequences found in all of the G proteins thus far studied. We screened clones derived from amplification of RNA isolated from 0- to 24-hr embryos of *Drosophila* and from total *Drosophila* adult RNA. We initially obtained a PCR product from both RNA sources which when sequenced differed from the murine $G\alpha_q$ at 3 of the deduced 51 amino acids. This PCR product was used to probe a *Drosophila* head cDNA library, and a clone was isolated and sequenced. The deduced amino acid sequence of the *Drosophila* $G\alpha_q$ homologue ($DG\alpha_q$) is compared with $G\alpha_q$ and Gα11 in Fig. 4. $DG\alpha_q$ is 76% identical (88% similar) to $G\alpha_q$. This α subunit lacks the putative N-terminal six additional amino acids; however, it has the characteristic changes in the sequence of the GTP-binding domain (GAGE box, Fig. 1*D*). In addition, $DG\alpha_q$ is predicted to be insensitive to pertussis toxin on the basis of its C-terminal amino acid sequence.

DISCUSSION

The analysis of cDNA clones encoding $G\alpha_q$ and Gα11 demonstrates the existence of ubiquitously expressed highly homologous G protein α subunits. During the characterization of these cDNA clones, Pang and Sternweiss used affinity chromotography with immobilized $\beta\gamma$ subunits (35) to purify novel α subunits (36). They found α subunits with an apparent molecular mass of 42 kDa. The amino acid sequences of tryptic fragments from these proteins were identical to the deduced amino acid sequence of $G\alpha_q$. One of the peptides also showed a sequence identical to that of Gα11. Furthermore, Pang and Sternweiss (36) found that the 42-kDa proteins have the properties that we would ascribe to $G\alpha_q$ and Gα11; they are not substrates for pertussis toxin modification, and they have unusual nucleotide binding kinetics. Using peptide antisera, they showed that the proteins are abundant in brain and lung extracts and present at lower levels in a variety of tissues.

There are many examples in the literature of effects of GTP and GTP analogs on signal transducing processes that are difficult to account for on the basis of the previously characterized G protein α subunits. Perhaps the most prominent among these is the observation that in many systems inositol trisphosphate and diacylglycerol release mediated by G protein activation of phospholipase C is pertussis toxin resistant (37). $G\alpha_z$ is thought to be a candidate for the G protein that might mediate pertussis toxin-insensitive activation of phospholipase C. However, $G\alpha_z$ expression appears to be restricted to neural tissues and platelets and it does not appear to be present in all tissues that show pertussis toxin-insensitive activation of phospholipase C (38). On the other hand, $G\alpha_q$ and Gα11 are ubiquitously expressed. Hence they

```
DGqa        MECCLSEEAKEQKRINQEIEKQLRRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGSGYSDEDKRGYIKLVFQNIFMAMQSM   84
Gqa   MTLESI•A••••••••AR•••D•••RHV•••••••••••••••••••••••••••••••••••••••••••••FT•••Y••••T•••A•   90
Ga11  MTLESM•A••••D•V••S•••A•••••••••••••••••••••••••••••••••••••••••••••A•••E•••••FT•••Y••••T•••A•   90

      IKAMDMLKISYGQGEHSELADLVMSIDYETVTTFEDPYLNAIKTLWDDAGIQECYDRRREYQLTDSAKYYLSDLARIEQADYLPTEQDIL   174
      •R•••T•••P•KYEHNKAH•Q••REV•V•K•SA••N••VD•••S••N•P•••••••••••••S••T••••N••D•VADPS••••Q••V•   180
      VR••ET•••L•KYEQNKAN•L•IREV•V•K•••••HQ•V•••••••S•P•V•••••••F••S•••••••T•VD••ATVG•••Q••V•   180

      RARVPTTGILEYPFDLDGIVFRMVDVGGQRSERRKWIHCFENVTSIIFLVALSEYDQILFESDNENRMEESKALFRTIITYPWFQNSSVI   264
      •V•••••••I••••••QSVI•••••••••••••••••••••••••M••••••••••V•V•••••••••••••••••••••••••••••••   270
      •V•••••••I••••••EN•I•••••••••••••••••••••••••M••••••••••V•V•••••••••••••••••••••••••••••••   270

      LFLNKKDLLEEKIMYSHLVDYFPEYDGPKQDHAAAKQFVLKKYLACNPDPERQCYSHFTTATDTENIKLVFCAVKDTIMQNALKEFNLG   353
      ••••••••••••••••••••••••••••QR•AQ••RE•I••MFVDL•••SDKII•••••C•••••••RF••A••••••L•LN•••Y••V   359
      ••••••••••D••LH•••••••••F•••QR•AQ••RE•I••MFVDL•••SDKII•••••C•••••••RF••A••••••L•LN•••Y••V   359
```

FIG. 4. Amino acid comparison of the *Drosophila* $G\alpha_q$ homologue ($DG\alpha_q$) with $G\alpha_q$ and Gα11. Dots represent identities with the $DG\alpha_q$ sequence.

are good candidates for this role. Furthermore, there are a variety of isoforms of phospholipase C; many cells contain multiple related but distinct phospholipase C gene products (39). Members of the G_q class could be involved in the activation of specific phospholipase C isozymes. Speculation about the possible interaction of the G_q class with phospholipase C is strengthened by the recent finding that DGα_q message is localized mainly to the *Drosophila* eye and ocellus structures (40). Earlier work indicated that regulation of phospholipase C in invertebrate eyes is central to the phototransduction cascade (41). There are a variety of other roles that are possible for the α subunits described here—e.g., ion channel activation and phospholipase A2 activation.

cDNA clones with deduced amino acid sequences that are identical to Gα11 have been found in RNA preparations from human tissue (T. T. Amatruda and M.I.S., unpublished results). The partial amino acid sequences found by Pang and Sternweiss (36) suggest that both Gα_q and Gα11 exist in rat brain extracts, and experiments with specific probes in our laboratory have detected Gα_q and Gα11 messages in a variety of tissues and cloned cell lines. These results suggest that both α subunits are expressed together in at least some cells.

The diversity of the G protein family continues to grow. As more α subunits are cloned, more classes will emerge. Within each class, highly homologous members are likely to display apparent crosstalk in reconstitution experiments. Indeed, it is difficult to discriminate between Gα_{i1} and Gα_{i3} by *in vitro* assays (27, 42). Gα_q and Gα11 may behave in similar fashion, since they are 97% identical over the domains apparently responsible for receptor and effector specificity. However, the strict evolutionary conservation of amino acid sequence differences that identify these α subunits argues that they are not redundant (2). The distinctions may become apparent as we begin to use more sophisticated assays of G protein function that include different combinations of $\beta\gamma$ dimers. There is evidence to suggest that all $\beta\gamma$ subunits are not equivalent (43, 44). Perhaps the diversity afforded by combinatorial associations of α, β, and γ subunits allows a cell to "fine tune" G protein function to the specific requirements of a particular signaling response.

We thank Paul Sternweiss and David Hyde for communicating their results prior to publication. We are grateful to Tom Wilkie for help with the PCR analysis of mRNA, Bruce Hamilton and Mike Palazzolo for *Drosophila* RNA and cDNA libraries, Don Seto and Carol Lee for sequencing assistance, and T. Wilkie and Narasimhan Gautam for reading the manuscript. This work was supported by Grant GM34236 from the National Institutes of Health.

1. Gilman, A. G. (1987) *Annu. Rev. Biochem.* **56,** 615–649.
2. Lochrie, M. A. & Simon, M. I. (1988) *Biochemistry* **27,** 4957–4965.
3. Kaziro, Y. (1990) in *ADP-Ribosylating Toxins and G Proteins*, eds. Moss, J. & Vaughan, M. (Am. Soc. Microbiol., Washington, DC), pp. 184–201.
4. Dietzel, C. & Kurjan, J. (1987) *Cell* **50,** 1000–1010.
5. Miyajima, I., Nakafuku, M., Nakayama, N., Brenner, C., Miyajima, A., Kaibuchi, K., Arai, K.-I., Kaziro, Y. & Matsumoto, K. (1987) *Cell* **50,** 1011–1019.
6. Kumagi, A., Papillo, M., Gunderson, R., Miake-Lye, R., Devreotes, P. & Firtel, R. A. (1989) *Cell* **57,** 265–275.
7. Firtel, R., Van Haastert, P. J. M., Kimmel, A. R. & Devreotes, P. (1989) *Cell* **58,** 235–239.
8. Yatani, A., Codina, J., Brown, A. M. & Birnbaumer, L. (1987) *Science* **235,** 207–211.
9. Yatani, A., Imoto, Y., Codina, J., Hamilton, S., Brown, A. M. & Birnbaumer, L. (1988) *J. Biol. Chem.* **263,** 9887–9895.
10. Graziano, M. P., Freissmuth, M. & Gilman, A. G. (1989) *J. Biol. Chem.* **264,** 409–418.
11. Stryer, L. (1986) *Annu. Rev. Neurosci.* **9,** 87–119.
12. Masters, S. B., Martin, M. W., Harden, T. K. & Brown, J. H. (1985) *Biochem. J.* **227,** 933–937.
13. Martin, T. F. J., Bajjalieh, S. M., Lucas, D. O. & Kowalehyk, J. A. (1986) *J. Biol. Chem.* **261,** 10141–10149.
14. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
15. Strathmann, M., Wilkie, T. M. & Simon, M. I. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7407–7409.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
17. Buss, J. E., Mumby, S. M., Casey, P. J., Gilman, A. G. & Sefton, B. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7493–7497.
18. Masters, S. B., Miller, R. T., Chi, M., Chang, F., Beiderman, B., Lopez, N. G. & Bourne, H. R. (1989) *J. Biol. Chem.* **264,** 15467–15474.
19. Graziano, M. P. & Gilman, A. G. (1989) *J. Biol. Chem.* **264,** 15475–15482.
20. Fong, H. K. W., Yoshimoto, K. K., Eversole-Cire, P. & Simon, M. I. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 3066–3070.
21. Matsuoka, M., Itoh, H., Kozasa, T. & Kaziro, Y. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5384–5388.
22. Casey, P. J., Fong, H. K. W., Simon, M. I. & Gilman, A. G. (1990) *J. Biol. Chem.* **265,** 2383–2390.
23. Itoh, H., Toyama, R., Kozasa, T., Tsukamoto, T., Matsuoka, M. & Kaziro, Y. (1988) *J. Biol. Chem.* **263,** 6656–6664.
24. Hsu, W. H., Rudolph, U., Sanford, J., Bertrand, P., Olate, J., Nelson, C., Moss, L. G., Boyd, A. E., III, Codina, J. & Birnbaumer, L. (1990) *J. Biol. Chem.* **265,** 11220–11226.
25. Strathmann, M., Wilkie, T. M. & Simon, M. I. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6477–6481.
26. Jones, D. T., Masters, S. B., Bourne, H. R. & Reed, R. R. (1990) *J. Biol. Chem.* **265,** 2671–2676.
27. Yatani, A., Mattera, R., Codina, J., Graf, R., Okabe, K., Padrell, E., Iyengar, R., Brown, A. M. & Birnbaumer, L. (1988) *Nature (London)* **336,** 680–682.
28. Masters, S. B., Sullivan, K. A., Miller, R. T., Beiderman, B., Lopez, N. G., Ramachandran, J. & Bourne, H. R. (1988) *Science* **241,** 448–451.
29. Provost, N. M., Somers, D. E. & Hurley, J. B. (1988) *J. Biol. Chem.* **263,** 12070–12076.
30. Quan, F., Wolfgang, W. J. & Forte, M. A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 4321–4325.
31. Yoon, J., Shortridge, R. D., Bloomquist, B. T., Schneuwly, S., Perdew, M. H. & Pak, W. L. (1989) *J. Biol. Chem.* **264,** 18536–18543.
32. de Sousa, S. M., Hoveland, L. L., Yarfitz, S. & Hurley, J. B. (1989) *J. Biol. Chem.* **259,** 18544–18551.
33. Thambi, N. C., Quan, F., Wolfgang, W. I., Spiegel, A. & Forte, M. (1989) *J. Biol. Chem.* **259,** 18552–18560.
34. Schmidt, C. J., Garen-Fazio, S., Chow, Y. & Neer, E. J. (1989) *Cell Regul.* **1,** 125–134
35. Pang, I.-H. & Sternweiss, P. C. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7814–7818.
36. Pang, I.-H. & Sternweiss, P. C. (1990) *J. Biol. Chem.*, in press.
37. De Vivo, M. & Gershengorn, M. C. (1990) in *ADP-Ribosylating Toxins and G Proteins*, eds. Moss, J. & Vaughan, M. (Am. Soc. Microbiol., Washington, DC), pp. 267–293.
38. Hinton, D. R., Blanks, J. C., Fong, H. K. W., Casey, P. J., Hildebrandt, E. & Simon, M. I. (1990) *J. Neurosci.* **10,** 2763–2774.
39. Rhee, S. G., Suh, P.-G., Ryu, S.-H. & Lee, S. Y. (1989) *Science* **244,** 546–556.
40. Lee, Y., Dobbs, M. B., Verardi, M. L. & Hyde, D. R. (1990) *Neuron*, in press.
41. Devary, O., Heichal, O., Blumenfeld, A., Cassel, D., Suss, E., Barash, S., Rubinstein, C. T., Minke, B. & Selinger, Z. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6939–6943.
42. Linder, M. I., Ewald, D. A., Miller, R. J. & Gilman, A. G. (1990) *J. Biol. Chem.* **265,** 8243–8251.
43. Cerione, R. A., Gierschik, P., Staniszewski, C., Benovic, J. L., Codina, J., Somers, R. L., Birnbaumer, L., Spiegel, A. M., Lefkowitz, R. J. & Caron, M. G. (1987) *Biochemistry* **26,** 1485–1491.
44. Casey, P. J., Graziano, M. P. & Gilman, A. G. (1989) *Biochemistry* **28,** 611–616.

Chapter 4

Gα12 and Gα13 define a fourth class of G protein alpha subunits

Michael P. Strathmann & Melvin I. Simon

Division of Biology 147-75
California Institute of Technology
Pasadena, California 91125
Telephone: (818) 356-3944
FAX: (818) 796-7066

Contributed by Melvin I. Simon

Abbreviations: PCR, polymerase chain reaction

Proceedings of the National Academy of Sciences USA, in press

Heterotrimeric GTP-binding proteins (G proteins) are central to the signaling processes of multicellular organisms. We have explored the diversity of the G protein subunits in mammals and found evidence for a large family of genes that encode tha alpha subunits. Amino acid sequence comparisons show that the different alpha subunits fall into at least three classes. These classes have been conserved in animals separated by considerable evolutionary distances; they are present in mammals, *Drosophila* and nematodes. We have now obtained cDNA clones encoding two murine alpha subunits, Gα12 and Gα13, that define a fourth class. The translation products are predicted to have molecular weights of 44 kD and to be insensitive to ADP-ribosylation by pertussis toxin. They share 67% amino acid sequence identity with each other and less than 45% identity with other alpha subunits. Their transcripts can be detected in every tissue examined, though the relative levels of the Gα13 message appear to be somewhat variable.

G proteins are heterotrimers composed of alpha, beta, and gamma subunits (for reviews see refs. 1-3). The alpha subunits belong to a much larger group of GTPases, including EF-Tu and Ras, which share similar structural elements (4,5). In all of these GTPases, a cycle of guanine nucleotide exchange and hydrolysis enables the protein to exist in two distinct states. This cycle allows G proteins to transiently relay signals from cell surface receptors to intracellular effectors. The receptors comprise a diverse family of proteins characterized by their transmembrane structure; they have seven membrane-spanning domains with highly conserved amino acid sequences. Upon interaction with the appropriate agonist, the receptor serves to accelerate the exchange of GDP for GTP on the G protein alpha subunit. This exchange is believed to be accompanied by dissociation of the alpha and beta-gamma subunits, allowing alpha (and in some cases beta-gamma) to interact with effectors. The intrinsic GTPase activity terminates the signal, returning the alpha subunit to its basal GDP bound state.

More than 100 different receptors in mammals are thought to couple through G proteins to a variety of effectors (6). The diverse alpha subunits, which mediate receptor function, can be classified on the basis of their amino acid sequence similarity. G_S and G_i subtypes have been implicated in the regulation of adenylate cyclase and the gating of certain ion channels (7). In the highly specialized visual system, biochemical experiments have elucidated the role of transducin (G_{t1}) in regulating phoshodiesterase and subsequently in controlling the levels of cyclic GMP during the photoreceptor cascade (8). These systems are the best characterized examples of G protein mediated effector activation. However, many other effector systems have been shown to be regulated by G proteins.

Pertussis toxin has proven to be an important tool in the dissection of G protein mediated pathways. Certain alpha subunits can be ADP-ribosylated by this toxin, thereby uncoupling the G protein from receptors. There are, however, some signaling

processes that are refractory to toxin inhibition but seem to be mediated by G proteins. In particular, phospholipase C is involved in both pertussis toxin sensitive and insensitive pathways (9). Recently, using molecular biological techniques we found a novel class of alpha subunits termed G_q that were predicted to be insensitive to pertussis toxin (10). Two groups independently purified the corresponding alpha subunits and have now demonstrated that these proteins can activate one of the phospholipase C (PLC) isotypes, PLC-β (11,12,13).

We are interested in understanding how G protein mediated signal transduction has adapted to the complex signaling processes that define multicellular organisms. To this end, we have examined G protein diversity in the mouse. Using a method based on the polymerase chain reaction (PCR), we found evidence for extensive diversity among the G protein alpha subunits (14). A small screen uncovered five novel sequences termed Gα10 through Gα14. Analysis of amino acid sequence conservation suggested that the known alpha subunits could be grouped into three distinct classes, G_s, G_i and G_q (10). Gα11 and Gα14 belong to the G_q class. In this paper, we present the cDNA sequences of Gα12 and Gα13, which define the fourth class of alpha subunits.

Materials and Methods

Polymerase Chain Reaction cDNA was made from 5μg total RNA with random hexanucleotide primers using MMLV reverse transcriptase in a 50μl reaction volume. Conditions were those supplied by the manufacturer (Bethesda Research Labs). The reactions were heated to 100°C for 5 min. The cDNA was then diluted to 400μl with water containing 10μg/ml DNase-free RNase A and incubated at 37°C for 30 min. 2μl of the cDNA mixture was used in a 20μl PCR reaction. The oligonucleotides used for PCR amplification of the cDNA were as follows:

oMP19, CGGATCCAARTGGATHCAYTGYTT;

oMP20, GGAATTCRTCYTTYTTRTTNAGRAA

oMP21, GGAATTCRTCYTTYTTRTTYAARAA;

CT56, CGGATCCARRTGGHTNSARTGYTT

in which R = A or G, Y = C or T, S = C or G, H = A, C, or T, and N = A, C, G or T. PCR was performed for 35 cycles on a Perkin Elmer Cetus thermal cycler. During each cycle, samples were denatured for 0.5 min at 94°C, annealed for 0.5 min at 40°C and extended for 0.5 min at 72°C. Each oligonucleotide was used in the PCR at 10 ng/μl. The buffer and taq polymerase were supplied by Cetus.

PCR Northern Analysis PCR was performed on cDNA prepared from total RNA as described above. To control for possible contamination of the RNAs by chromosomal DNA, each preparation of RNA was treated as described above except reverse transcriptase was not added to the cDNA synthesis mixture. These controls were then subjected to PCR alongside the other samples. The PCR products were electrophoresed through a 2% agarose gel, blotted to Hybond-N (Amersham) and hybridized according to the manufacturer's instructions. These blots were probed with radiolabeled oligonucleotides specific to Gα11 (CTCGCTTAGTGCCACC), Gα12 (CTCGCTCGAGGACACCATGAAC) and Gα13 (TTCACTTGAAGAGACAAGGA AA). The oligonucleotides were end-labeled with γ^{32}P-ATP as described (15). Blots were washed three times at room temperature for 5 min in 0.90 M NaCl/0.90 M sodium citrate (6X SSC)/ 0.1% SDS followed by a 1 min wash in the same solution at the calculated melting temperature of each oligonucleotide.

Isolation of cDNAs To obtain cDNAs encoding Gα12 and Gα13, a random-hexanucleotide primed mouse brain cDNA library in λZAPII (Stratagene) was screened using standard techniques. $6x10^5$ clones were plated at a density of $3x10^5$ plaques per 150 mm petri dish (16). Nylon filters (Hybond-N, Amersham) were

prepared and hybridized according to the manufactures instructions. Probes were made from the Gα12 and Gα13 PCR fragments isolated previously. These fragments were labeled using the Multiprime DNA Labelling System (Amersham).

Sequence Analysis rG12-5 and 3-G13 cDNA clones (Fig. 1) were subcloned into the plasmid pMOB (17). In order to obtain sequence priming sites spaced throughout the clones, the γδ transposon was randomly inserted into the cDNAs as described (17). A subset of these insertions was isolated, with transposons located every 300-400 bases along the length of each clone. Supercoiled plasmid DNA was sequenced using the Sequenase kit marketed by United States Biochemical.

Results

Sequence of Gα12 and Gα13 Gα12 and Gα13 were initially isolated as PCR products. Mouse brain cDNA was amplified with a set of degenerate oligonucleotide primers corresponding to amino acid domains conserved in all alpha subunits (14). The resulting mixture of PCR products was characterized by sequencing individual clones.

To obtain full-length cDNAs encoding Gα12 and Gα13, approximately $6x10^5$ clones from a random-primed mouse brain cDNA library were screened at high stringency using the individual PCR products as probes. 25 positives were obtained with the Gα12 probe, whereas four were found using the Gα13 probe. In order to determine the extent of the 5'-sequence present in these Gα12 clones, PCR was performed on the primary positives with oligonucleotide primers specific to the cloning vector and the known Gα12 sequence (16). Those clones yielding the largest amplified fragments were purified. The complete sequences of two cDNAs encoding Gα12 and Gα13 are shown in Figure 1. The predicted translation products are 379 and 377 amino acids in length, respectively. Both proteins are predicted to have a

molecular weight of 44 kD. A small open reading frame (12 amino acids) is located just upstream of the Gα12 coding region; however, the first methionine lies in a poor context for translation initiation (18).

The amino termini of Gα12 and Gα13 are quite distinctive when compared to other mammalian alpha subunits (Fig. 2). A number of proteins, including $G\alpha_i$ and $G\alpha_o$, are known to be N-myristylated on a glycine at the second position (Gly-2, ref.19). This modification is essential for anchoring the src oncogene product to the membrane (20). A comparison of the known N-myristylated proteins yields the consensus sequence MGXXXS (2). Thus Gα12 and Gα13 are not predicted to be modified in this way. Perhaps these two alpha subunits undergo a different form of post-translational modification.

The crystal structures of ras and EF-Tu, as well as genetic studies of the effects of specific mutations on GTPase function, have revealed several domains in these proteins that are critical for guanine nucleotide interactions (5). Analogous regions in G protein alpha subunits are readily identifiable (2). These regions are indicated in figure 2 as A,C,G and I from Halliday's nomenclature (21). The function of the I-region is not well characterized in the alpha subunits; however, it is very highly conserved. All of the previously known mammalian alpha subunits contain the sequence TCATDT except $G\alpha_s$ and $G\alpha_{olf}$, which differ at a single position (Fig.2). A mutation (Thr-144 to Ile) in the analogous region in ras results in a 25-fold reduction in GTP affinity (22). The fact that both Gα12 and Gα13 differ markedly from the consensus sequence in this region suggests that they may interact with guanine nucleotides in a manner distinct from some of the more well characterized alpha subunits.

The C-terminus of several alpha subunits can be ADP-ribosylated by pertussis toxin. The modification occurs at a cysteine 4 residues from the end of the protein.

Those mammalian alpha subunits that lack this cysteine, $G\alpha_s$, $G\alpha_{olf}$, $G\alpha_z$, $G\alpha_q$ and $G\alpha 11$ are all refractory to modification by pertussis toxin (2,12,23). The predicted amino acid sequences of $G\alpha 12$ and $G\alpha 13$ also lack this cysteine. Consequently, the translation products are likely to be insensitive to pertussis toxin.

Distribution of $G\alpha 12$ and $G\alpha 13$ Messages Northern analysis indicates that the $G\alpha 12$ message is approximately 4 kb whereas the transcript encoding $G\alpha 13$ is >6 kb (14). As a sensitive assay for the distribution of these transcripts in a wide variety of tissues, we used PCR northern analysis (T.M. Wilkie and M.I.S., unpublished). With this technique, multiple alpha subunit sequences are amplified by PCR from cDNA using degenerate oligonucleotide primers. These primers correspond to amino acid domains that are conserved in all alpha subunits. An individual sequence is detected by Southern hybridization of the amplified pool with a labeled oligonucleotide specific for that sequence. The method provides a sensitive means of detecting the presence and relative abundance of a specific alpha subunit sequence. The results of this analysis are shown in Figure 3. Both $G\alpha 12$ and $G\alpha 13$ can be detected in every tissue examined. $G\alpha 12$ is expressed at relatively constant levels in most tissues, though it is noticeably lower in intestine. $G\alpha 13$ appears to be expressed at somewhat higher levels in eye, kidney, liver, lung and testis. $G\alpha 11$ has been shown by northern hybridization to be expressed at fairly constant levels in a variety of tissues (10); this pattern is reproduced by the PCR northern analysis. Controls for contaminating chromosomal DNA were negative (data not shown, see Material and Methods). It should be noted that PCR northern analysis can minimize the effect of partially degraded RNA preparations which upon northern analysis will indicate artificially low expression levels. This effect is particularly pronounced with large messages and may explain the lower relative expression level originally observed for $G\alpha 13$ in liver (14).

Relationships Among the Alpha Subunits The previously known mammalian alpha subunits can be grouped by amino acid identity into three classes, G_s, G_i and G_q (10). Members of one class, in general, are less than 50% identical to members of the other classes. These relationships represent not only primary sequence homologies but also, to some extent, functional similarities. Thus within a class, members often display significant levels of "crosstalk" in reconstitution experiments. For example, both $G\alpha_s$ and $G\alpha_{olf}$ can couple β-adrenergic receptors to adenylate cyclase (24). Also, the three types of $G\alpha_i$ proteins, both purified and recombinant forms, can open atrial potassium channels (25).

Gα12 and Gα13 define a fourth class of alpha subunits (Fig. 4). They are less than 45% identical to members of the other three classes while they share 67% amino acid identity with each other.

Discussion

Cellular responses elicited by numerous different ligands can be affected by pertussis toxin. The ability to inhibit a particular pathway with this toxin has been widely used to implicate a G protein mediator. Only the Gi class includes alpha subunits known to be substrates for ADP-ribosylation by pertussis toxin. In fact, the majority of genes encoding alpha subunits yield products that appear to be insensitive to this modification. The ubiquitously expressed Gα12 and Gα13 cDNAs reveal a fourth class of alpha subunits that are predicted to encode pertussis toxin resistant proteins.

Phospholipase C (PLC) catalysed release of inositol triphosphate and diacylglycerol was inferred to involve a pertussis toxin resistant G protein that has been called "G_p" (9). Purified activated alpha subunits corresponding to the $G\alpha_q$

and/or Gα11 clones have been shown to stimulate PLC-β activity and to have some of the properties previously ascribed to G_p (13). However, there are a variety of PLC isoforms; many cells contain multiple related but distinct PLC gene products (28). Perhaps Gα12 and Gα13 couple to these other PLC isozymes or to other phospholipases that are activated through G-protein mediated pathways.

The relationships among the alpha subunits depicted in Figure 3 developed early in the evolution of the animal kingdom. Representatives of the G_s, G_i and G_q classes have been found in *Drosophila* and nematodes (10, 29-36, Kim, U.-J. and Simon, M.I., unpublished). It seems likely that the G12 class will be found in these organisms as well. In fact, Parks and Wieschaus recently showed that the Drosophila gene concertina (*cta*) encodes a G protein alpha subunit (37). *cta* is clearly more related to Gα12 and Gα13 (56% amino acid sequence identity) than to the other alpha subunits (35-44% identity). Mutations in *cta* affect gastrulation and have a maternal effect; mothers homozygous for these mutations will survive, but their offspring do not develop properly. If the functions of specific alpha subunits are conserved in evolution, then we may expect a member of the G12 class, perhaps as yet unidentified, to influence early developmental processes in more complex organisms.

The application of molecular biological techniques to the study of G proteins has proven to be a powerful tool in the analysis of signal transduction in complex multicellular organisms. The perception that the diverse signaling requirements of mammals need only a small set of G proteins is being replaced by an image of greater complexity. In addition to multiple classes of alpha subunits, extensive diversity of beta and gamma subunits has also been found (38-42). If combinatorial associations of the three subunits can occur in a functionally relevant manner, then the potential for G protein diversity is enormous.

We are grateful to Carol Lee for sequencing assistance, Joyce Ito for help in preparing the manuscript, and Bruce Hamilton, David Horowitz, Michael Lochrie and Tom Wilkie for many discussions. This work was supported by Grant GM34236 from the National Institutes of Health.

References

1. Gilman, A.G. (1987) *Ann. Rev. Biochem.* **56**, 615-649.

2. Lochrie, M.A. and Simon, M.I. (1988) *Biochemistry* **27**, 4957-4965.

3. Simon, M.I., Strathmann, M.P. and Gautam, N. (1991) *Science*, in press.

4. Bourne, H.R., Sanders, D.A. and McCormick, F. (1990) *Nature* **348**, 125-132.

5. Bourne, H.R., Sanders, D.A. and McCormick, F. (1991) *Nature* **349**, 117-127.

6. Raymond, J.R., Hnatowich, M., Caron, M.G. and Lefkowitz, R. (1990) in *ADP-Ribosylating Toxins and G-Proteins*, eds. Moss, J. and Vaughan, M. (Am. Soc. Microbiol., Washington, DC), pp. 163-176.

7. Brown A. and Birnbaumer, L. (1990) *Annu. Rev. Physiol.* **52**, 197-213.

8. Stryer, L. (1986) *Ann. Rev. Neurosci.* **9**, 87-119.

9. De Vivo, M. and Gershengorn, M.C. (1990) in *ADP-Ribosylating Toxins and G Proteins,* eds. Moss, J. and Vaughan, M. (Am. Soc. Microbiol., Washington, D.C.), pp. 267-293.

10. Strathmann, M.P. and Simon, M.I. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 9113-9117.

11. Taylor, S.J., Smith, J.A. and Exton, J.H. (1990) *J. Biol. Chem.* **265**, 17150-17156.

12. Pang, I.-H., and Sternweis, P.C. (1990) *J. Biol. Chem.***265**, 18707-18712.

13. Smrcka, A.V., Hepler, J.R., Brown, K.O. and Sternweis, P.C. (1991) *Science* **251**, 804-807.

14. Strathmann, M., Wilkie, T.M. and Simon, M.I. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7407-7409.

15. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

16. Hamilton, B.A., Palazzolo, M.P. and Meyerowitz, E.M. (1991) *Nucl. Acids Res.*, in press.

17. Strathmann, M.P., Hamilton, B.A., Mayeda, C.A., Simon, M.I., Meyerowitz, E.M. and Palazzolo, M.J. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1247-1250.

18. Kozak, M. (1989) *J. Cell Biol.* **108**, 229-241.

19. Buss, J.E., Mumby, S.M., Casey, P.J., Gilman, A.G. and Sefton, B.M. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7493-7497.

20. Kamps, M.P., Buss, J. and Sefton, B. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4625-4628.

21. Halliday, K.R. (1983-1984) *J. Cyclic Nucleotide Protein Phosphorylation Res.* **9**, 435-448.

22. Feig, L.A., Pan, B.-T., Roberts, T.M. and Cooper, G.M. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4607-4611.

23. Casey, P.J., Fong, H.K.W., Simon, M.I. and Gilman, A.G. (1990) *J. Biol. Chem.* **265**, 2383-2390.

24. Jones, D.T., Masters, S.B., Bourne, H.R. and Reed, R.R. (1990) *J. Biol. Chem.* **265**, 2671-2676.

25. Yatani, A., Mattera, R., Codina, J., Graf, R., Okabe, K., Padrell, E., Iyengar, R., Brown, A.M. and Birnbaumer, L. (1988) *Nature* **336**, 680-682.

26. Hsu, W.H., Rudolph, U., Sanford, J., Bertrand, P., Olate, J., Nelson, C., Moss, L.G., Boyd, A.E. III, Codina, J. and Birnbaumer, L. (1990) *J. Biol. Chem.* **265**, 11220-11226.

27. Strathmann, M., Wilkie, T.M. and Simon, M.I. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6477-6481.

28. Rhee, S.G., Suh, P.-G., Ryu, S.-H. and Lee, S.Y. (1989) *Science* **244**, 546-556.

29. Provost, N.M., Somers, D.E. and Hurley, J.B. (1988) *J. Biol. Chem.* **263**, 12070-12076.

30. Quan, F., Wolfgang, W. J. and Forte, M. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 4321-4325.

31. Yoon, J., Shortridge, R.D., Bloomquist, B.T., Schneuwly, S., Perdew, M.H. and Pak, W.L. (1989) *J. Biol. Chem.* **264**, 18536-18543.

32. de Sousa, S.M., Hoveland, L.L., Yarfitz, S. and Hurley, J.B. (1989) *J. Biol. Chem.* **259**, 18544-18551.

33. Thambi, N.C., Quan, F., Wolfgang, W.I., Spiegel, A. and Forte, M. (1989) *J. Biol. Chem.* **259**, 18552-18560.

34. Schmidt, C.J., Garen-Fazio, S., Chow, Y. and Neer, E.J. (1989) *Cell Regulation* **1**, 125-134.

35. Lee, Y., Dobbs, M.B., Verardi, M.L. and Hyde, D.R. (1990) *Neuron* **5**, 889-898.

36. Lochrie, M.A., Mendel, J.E., Sternberg, P. and Simon, M.I. (1991) *Cell Regulation*, in press.

37. Parks, S. and Wieschaus, E. (1991) *Cell* **64**, 447-458.

38. Levine, M.A., Smallwood, P.M., Moen, P.T., Helman, L.J. and Ahn, T.G. (1990) *Proc. Natl. Acad. Sci USA.* **87**, 2329-2333.

39. von Weizsäcker, E., Strathmann, M.P. and Simon, M.I. (1991) *Mol. Cell. Biol.*, submitted.

40. Gautam, N., Baetscher, M., Aebersold, R. and Simon, M.I. (1989) *Science* **244**, 971-974.

41. Robishaw, J., Kalman, V.K., Moomaw, C.R. and Slaughter, C.A. (1989) *J. Biol. Chem.* **264**, 15758-15761.

42. Gautam, N., Northup, J., Tamir, H. and Simon, M.I. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7973-7977.

Figure 1. The nucleotide sequence and predicted amino acid sequence of Gα12 and Gα13. The complete sequences of two cDNA clones encoding Gα12 and Gα13 (rG12-5 and 3-G13, respectively) are shown above. These clones were isolated from a random primed mouse brain cDNA library. Sequence was obtained from two other clones (rG12-1 and rG12-2) which begin at bases 5 and -100 respectively of the Gα12 sequence. 200 bases from the 5'-ends of both clones were sequenced and found to be identical to rG12-5.

Gα12

```
                                                          TCGGGGCGGCGGCGG -121
CGGACGCAGCCTGAGGGAGCGGCGGGCGGGGCGGCGCCTGGCCCGGCAGGGCCGCCGCCGG -61
CTGAGGGGCGTCCATGGCGCGCGGGCCCCAGCGGGGCGCGGCCGCGGCCTGAGGGCGGCC -1

ATGTCCGGGGTGGTGCGGACCCTTAGCCGCTGCTTGCTGCCGGCCGAGGCCGAGCCCGGC 60
M  S  G  V  V  R  T  L  S  R  C  L  L  P  A  E  A  G  A  R  20

GAGCGCAGGGCGGGCGCGGCGCGCGACGCGGAGCGCGAGGCCCGACGGCGCAGCCGCGAC 120
E  R  R  A  G  A  A  R  D  A  E  R  E  A  R  R  R  S  R  D  40

ATCGACGCGCTGCTGGCCCGCGAGCGGCGCGCGGTGCGGCGGCTGGTCAAGATCCTGCTG 180
I  D  A  L  L  A  R  E  R  R  A  V  R  R  L  V  K  I  L  L  60

CTGGGCGCCGGCGAGAGCGGCAAGTCCACCTTCCTCAAGCAGATGCGCATCATCCACGGC 240
L  G  A  G  E  S  G  K  S  T  F  L  K  Q  M  R  I  I  H  G  80

CGGGAGTTCGACCAGAAGGCGCTGCTGGAGTTCCGCGACACCATCTTCGACAACATCCTT 300
R  E  F  D  Q  K  A  L  L  E  F  R  D  T  I  F  D  N  I  L  100

AAGGGTTCGAGGGTTCTTGTGGACGCTCGAGACAAGCTCGGCATTCCCTGGCAGCACTCT 360
K  G  S  R  V  L  V  D  A  R  D  K  L  G  I  P  W  Q  H  S  120

GAGAACGAGAAGCACGGGATGTTTCTGATGGCCTTCGAGAACAAGGCAGGGCTGCCTGTG 420
E  N  E  K  H  G  M  F  L  M  A  F  E  N  K  A  G  L  P  V  140

GAGCCTGCCACCTTCCAGCTCTACGTGCCAGCCCTGAGTGCCCTCTGGAGAGACTCGGGG 480
E  P  A  T  F  Q  L  Y  V  P  A  L  S  A  L  W  R  D  S  G  160

ATCAGGGAAGCCTTCAGCCGCAGAAGCGAGTTCCAGCTGGGTGAATCAGTGAAGTACTTC 540
I  R  E  A  F  S  R  R  S  E  F  Q  L  G  E  S  V  K  Y  F  180

CTGGATAACTTGGACCGGATTGGCCAGCTGAACTACTTCCCCAGTAAGCAAGACATCCTG 600
L  D  N  L  D  R  I  G  Q  L  N  Y  F  P  S  K  Q  D  I  L  200

CTGGCTAGAAAGGCCACCAAGGGAATCGTGGAACATGACTTCGTTATAAAGAAAATCCCA 660
L  A  R  K  A  T  K  G  I  V  E  H  D  F  V  I  K  K  I  P  220

TTTAAGATGGTGGATGTGGGCGGCCAGAGGTCACAGCGCCAGAAGTGGTTCCAGTGCTTC 720
F  K  M  V  D  V  G  G  Q  R  S  Q  R  Q  K  W  F  Q  C  F  240

GACGGCATCACATCTATCCTGTTCATGGTGTCCTCGAGCGAGTATGACCAGGTCCTCATG 780
D  G  I  T  S  I  L  F  M  V  S  S  S  E  Y  D  Q  V  L  M  260

GAGGACAGGCGCACCAACCGGCTGGTGGAGTCCATGAACATCTTCGAGACCATCGTCAAC 840
E  D  R  R  T  N  R  L  V  E  S  M  N  I  F  E  T  I  V  N  280

AACAAGCTCTTCTTCAACGTCTCCATCATCCTCTTCCTCAACAAGATGGACCTCCTGGTG 900
N  K  L  F  F  N  V  S  I  I  L  F  L  N  K  M  D  L  L  V  300

GAGAAGGTGAAGTCTGTGAGCATTAAGAAGCACTTCCCAGATTTCAAGGGCGACCCGCAC 960
E  K  V  K  S  V  S  I  K  K  H  F  P  D  F  K  G  D  P  H  320

CGGCTGGAGGACGTCCAGCGCTACCTGGTGCAGTGCTTCGACAGGAAGCGCAGGAACCGC 1020
R  L  E  D  V  Q  R  Y  L  V  Q  C  F  D  R  K  R  R  N  R  340

AGCAAGCCCCTGTTCCACCACTTCACCACCGCCATAGACACCGAGAACATCCGCTTCGTG 1080
S  K  P  L  F  H  H  F  T  T  A  I  D  T  E  N  I  R  F  V  360

TTTCATGCTGTGAAGGACACGATCCTGCAGGAGAACCTGAAAGACATCATGCTGCAGTGA 1140
F  H  A  V  K  D  T  I  L  Q  E  N  L  K  D  I  M  L  Q      379

GAGACAGCCCCTCCCTCTGCTGTCCTCTCGGCCTCTGGGGCTGTGGCCGCCCACGTGGTT 1200
CTGGTGTGTCTGCCGCCTGCATGGACCTGCAGTGGGCATCTCGAATCCACACCATCTAAT 1260
ACCTGGCTCAAGAGTGCTGGACTACCAGCCACACCAGCGAGCTCCGGGCAAGAGGACGCG 1320
GGAGTGTCAGTGATAGCTACTGAATCGGGGACTGTGAGACTACTGAATCCACACCATCAC 1380
TCTGCTGTTGATCTGTAAAACCTGGGGACACACGTTTCGCCTCTCTACAGAAACTCTTCC 1440
TTGTTGACACAGTTAACCAAGGTTCAGGTGTACGCAGACACACGCGTGCTCTCATTAATG 1500
ACAGAAACCCAACCCCAGCTGGCCTTGCAGCCGGCTCTCCCCTCTGCCGTGCAGGTTCTC 1560
ACTGAGTCACACTAACTCGGGAGCTTGGCCTCACACTAGGCTGTAAAGAGCATGCGATGC 1620
TGCTAAGCTCACCGTGTCCAAGCCCAGGCCTCAGTGCCACCTCCCAGAGTGAACACACAG 1680
CCTTAGCATCGTGCCAGCACCCTCCGAGTCTCGCAAAGGCGCAGGCGGGGTCAGGTGTCC 1740
ACACT                                                          1745
```

Gα13

```
                                                          TCCCTCCC -121
CGGGCGCGCTCGGGCCGCCGCTGCGCTCCCCGCCCTCCAGCCGCCTTGCCGGAGCCGCCC -61
GCTGCCGGAGGAGGAGGTGGAGGAGCCGCAGGGGCCCGCCGAGGCGGCGGCGGCGGCAAG -1

ATGGCGGACTTCCTGCCGTCGCGCTCCGTGCTGTCCGTGTGCTTCCCGGGCTGCGTGCTG 60
M  A  D  F  L  P  S  R  S  V  L  S  V  C  F  P  G  C  V  L  20

ACGAACGGCGAGGCCGAGCAGCAGCGCAAGTCCAAGGAGATCGACAAATGCCTGTCGCGG 120
T  N  G  E  A  E  Q  Q  R  K  S  K  E  I  D  K  C  L  S  R  40

GAGAAGACCTACGTGAAGCGGCTGGTGAAGATCCTGCTGCTGGGCGCGGGCGAGAGCGGC 180
E  K  T  Y  V  K  R  L  V  K  I  L  L  L  G  A  G  E  S  G  60

AAGTCCACCTTCCTGAAGCAGATGCGGATCATCCACGGCCAGGACTTCGACCAGCGCGCG 240
K  S  T  F  L  K  Q  M  R  I  I  H  G  Q  D  F  D  Q  R  A  80

CGCGAGGAGTTCCGCCCCACCATCTACAGCAACGTGATCAAAGGTATGAGGGTGCTGGTA 300
R  E  E  F  R  P  T  I  Y  S  N  V  I  K  G  M  R  V  L  V  100

GATGCCCGAGAGAAGCTTCATATTCCCTGGGGAGATAACAAAAACCAGCTCCATGGAGAC 360
D  A  R  E  K  L  H  I  P  W  G  D  N  K  N  Q  L  H  G  D  120

AAGTTGATGGCATTTGATACCCGCGCCCCCATGGCTGCCCAGGGGATGGTGGAGACTCGA 420
K  L  M  A  F  D  T  R  A  P  M  A  A  Q  G  M  V  E  T  R  140

GTGTTCCTGCAGTATCTTCCTGCTATCAGAGCCTTATGGGAGGACAGTGGTATACAGAAT 480
V  F  L  Q  Y  L  P  A  I  R  A  L  W  E  D  S  G  I  Q  N  160

GCCTACGATCGGCGCCGGGAATTCCAGCTGGGTGAGTCTGTAAAGTATTTCTTGGATAAC 540
A  Y  D  R  R  R  E  F  Q  L  G  E  S  V  K  Y  F  L  D  N  180

TTGGATAAACTTGGAGTACCGGATTACATTCCATCACAGCAAGATATCCTGCTTGCCAGA 600
L  D  K  L  G  V  P  D  Y  I  P  S  Q  Q  D  I  L  L  A  R  200

AGGCCCACCAAAGGCATCCATGAGTACGACTTTGAAATTAAAAATGTTCCTTTCAAAATG 660
R  P  T  K  G  I  H  E  Y  D  F  E  I  K  N  V  P  F  K  M  220

GTTGATGTAGGTGGCCAGAGATCAGAACGGAAACGCTGGTTTGAATGCTTTGACAGTGTG 720
V  D  V  G  G  Q  R  S  E  R  K  R  W  F  E  C  F  D  S  V  240

ACGTCGATACTTTTCCTTGTCTCTTCAAGTGAATTTGACCAGGTGCTTATGGAGGACCGC 780
T  S  I  L  F  L  V  S  S  S  E  F  D  Q  V  L  M  E  D  R  260

CAGACCAATCGCCTTACAGAATCTCTGAACATTTTTGAAACAATTGTCAACAATCGGGTT 840
Q  T  N  R  L  T  E  S  L  N  I  F  E  T  I  V  N  N  R  V  280

TTCAGCAACGTCTCCATAATCCTCTTCTTAAACAAGACAGACTTGCTCGAGGAGAAAGTG 900
F  S  N  V  S  I  I  L  F  L  N  K  T  D  L  L  E  E  K  V  300

CAAGTTGTTAGCATCAAAGACTATTTCCTAGAATTTGAAGGGGACCCCCACTGCTTAAGA 960
Q  V  V  S  I  K  D  Y  F  L  E  F  E  G  D  P  H  C  L  R  320

GACGTCCAAAAGTTTCTGGTGGAATGCTTCCGGGGGAAACGCCGGGACCAGCAGCAGAGG 1020
D  V  Q  K  F  L  V  E  C  F  R  G  K  R  R  D  Q  Q  Q  R  340

CCGTTGTACCACCACTTCACCACCGCGATCAACACAGAGAACATCCGCCTCGTGTTCCGG 1080
P  L  Y  H  H  F  T  T  A  I  N  T  E  N  I  R  L  V  F  R  360

GACGTGAAGGACACGATCCTTCATGACAACCTGAAGCAGCTCATGCTGCAGTGATGACAA 1140
D  V  K  D  T  I  L  H  D  N  L  K  Q  L  M  L  Q      377

GACTCGCTGTTTTAATACCTTGTTGTGATTTTGATTGTTTTCTGTTTGTCTTGTTTTTCT 1200
CTTTAAATGGCAGTTTACAACAGCAGTTAGAAGAATCTCAGTTGTGTGTGTCTAACTTCT 1260
TGGAGAACTTAGTTCATCTGTGGCCTTTAGTTTGTGGCTGACAGCTGCTGAATTAACCCA 1320
TTGCCAATATCTGCTAGAATTTGGGCTTTGATTGGTTTCACTGTGGCTTACCTTCAAGTC 1380
CAATTATAACTTCCTGGCAACCTTACACTAGGTGTTTTTCAGGCCTTGAATTCTCATAAC 1440
TGTCCTGCAGTTCTAAGAATAGGCAGTATCTTTAAGAACTTGTAGGGATGAGAGTATGAA 1500
AATTCAGCTCTAAAGAGAGAGAGTGCCTTACCCAGCAATGACATGCCTCCCTATGTGAGC 1560
CCTGGTCTGGAAGGGCTGTTAGAGAAAAGTTTTAGGGAGGAGTCTCGGCCCACAGAGGCT 1620
TGGCCCCCTACACACACAGCAGTAGTTTGTTTTCCCTCAATTTCTCTAATTTCTGTAGGG 1680
CAGCAGTCTCAGCCCACAGAGGCTTGGCCCCCTACACACACAGCAGTAGTTTGTTTTTCT 1740
CAATCTCTTTACTTTCTGAAGGTAATGTCTTACTGCTTAAGTTTACAAGATGGATCTCAG 1800
CCTGATGCTGATGCGGTCACAAATGGTACATGTGCTATGTTGATCTGAGATGAGCAGTGA 1860
CACGCAGTTTCCTTTGCCATACATGCTGTCACTAACTTAATGTGCCATGTTACAACATCT 1920
GGAGAGGGCAAGCACCCTGGTGACGGACATCACCATAGCAGCCCCAGAGTGTGACTTTCT 1980
GGTCCTGTCTCCTTGGACTTTGTGCCTTGCTGGAGAGAACCACTGTTGGGTCTTCTTCCT 2040
TTTGCCCCACTGTCTCAACCTCAGCAGCTCTGAACTGTGGTGCTCTCGGGCTTCTCTTGT 2100
CCCCGAGAGTTTGGTGGGCAGCATTTCCCTGTGAACTGTTCAAAAAGGCATCCTTTCAGT 2160
ACCCTGTCCCTGGTTAGTGTCTCTCTGTGACAGAGTTTGTATGTGTTGTGTGTGTCATCG 2220
TGTGTGAGACACAGCTTCATTGAGGTACATTCAAATGTCTTAAGGATGCATCTAAGGATA 2280
AGCAATTTTTTTTT                                                 2294
```

Figure 2. Amino acid comparison of Gα12 with Gα13, $G\alpha_s$, $G\alpha_{i1}$ and $G\alpha_q$. Dots represent identities with the Gα12 sequence.


Gα12 MSGVVRTLSRCLLPAEAGARERRAGAARDAEREARRRSRDIDALLARERRAVRRLVKILLLGAGESGKSTFLKQMRIIHGREFDQK-ALLEFRDTIFDNILKGSRVLVDA
Gα13 MADFLPSRSVLSVCFPGCVLTNG•A•QQ•K•KE••KC•S••KTY•K•••••••••••••••••••••••••••QD•••R-•RE•••P••YS•VI••M••••••
Gαs MGCLGNSKTEDQRNE•KAQ•EANKK•EKQ•QKDKQVY•ATHR••••••••••••IV•••••L•VNG•NGESEKATKVQD•KN•LKEAIETI•A•
Gαi1 MGCTLSAEDKA•VE••KM••RN•REDGEKAA•E••L••••••••••••IV•••K•••EAGYSEE-ECKQYKAVVYS•TIQSIIAIIR•
Gαq MTLESIMACCLSEEAK••••INDE•ERHVR•DK•DA••EL•L••••T••••••••I••••••••SGYSDE-DKRG•TKLVYQ••FTAMQAMIR•
Gα12 RDKLGIP--WQHSENEKHGMFLMAFENKAGLP----VEPATFQL•VPALSALWRDSGIREAFSRRSEFQLGESVKYFLDNLDRIGQLNYFPSKQDILLARKATKGIVEHD
Gα13 •E••H••••-•GDNK•QL••DK••••DTR•PMAAQGM••TRV•LQ•L••IR•••E••••QN•YD••R•••••••••••••••KL•VPD•I••Q•••••••RP••••H•Y•
Gαs MSN•VP•VELANP••QFRVDYILSVM•VPNFDFPP-------EF•EH•-K•••E•E•V•ACYE•SN•Y••IDCAQ••••K••V•K•AD•V••D••L•RC•VL•S••F•TK
Gαi1 MGR•K•DFGDAARADDARQL•VL•GAAEE•F--------M•AE•-AGVIKR••K•••VQAC•N•SR•Y••ND•AA•Y•ND••••A•P••I•TQ••V•RT•VK•T••••TH
Gαq M•T•K••YKYE•NKAHAQLVREVDV•KVSAFE----------NP••D•IKS••N•P••Q•CYD••R•Y••SD•T••Y•ND•••VADPS•L•TQ••V•RV•VP•T••I•YP
Gα12 FVIKKIPFKMVDVGGQRSQRQKWFQCFDGITSILFMVSSSEYDQVLMEDRRTNRLVESMNIFETIVNNKLFFNVSIILFLNKMDLLVEKVK--SVSIKKHFPDF------
Gα13 •E••NV••••••••••••E•KR••E•••SV•••••L•••••F••••••••Q••••T••L•••••••••RV•S••••••••••T•••E•••Q--V••••DY•LE•-------
Gαs •QVD•VN•H•F••••••DE•R••I•••NDV•A•I•V•A••S•NM•IR••NQ••••Q•AL•L•KS•W••RWLRTI•V••••••Q•••A•••LAGKSK•EDY••E•ARYTTP
Gαi1 •TF•DLH•••F••••••••E•K••IH••E•V•A•I•C•AL•D••L••A••EEM••MH•••KL•DS•C•••W•TDT••••••••K••FE••I•--KSPLTICY•EY------
Gαq •DLQSVI•R••••••••••E•R••IH••ENV•••M•L•AL••••••V•SDNE••ME••KAL•R••ITYPW•Q•S•V••••••K•••E••IM--YSHLVDY••EY------
Gα12 ------KGDPHRLEDVQRYLVQCFDR--KRRN-RSKPLFHHFTTAIDTENIRFVFHAVKDTILQENLKDIMLQ
Gα13 ------E••••C•R•••KF••E••RG-----DQQQR••Y•••••••N•••••L••RD••••••HD•••QL•••
Gαs EDATPEP•EDP•VTRAKYFIRDE•L•ISTASGDGRHYCYP•••C•V••••••R••NDCR•I•QRMH•RQYE•L
Gαi1 ------A•S-NTY•EAAA•IQCQ•EDLN••KDTKEI--YT•••C•T••K•VQ•••D••T•V•IKN••••CG•F
Gαq ------D•PQRDAQAAREFILKM•VDLNPDSD-KII--YS•••C•T••••••••A••••••••L•••EYN•V

Figure 3. PCR Northern analysis of Gα12, Gα13 and Gα11. RNA from various mouse tissues was reverse transcribed and PCR amplified using the degenerate oligonucleotides oMP19, oMP20, oMP21, and CT56. The amplified products were hybridized with radiolabelled oligonucleotides specific to Gα12, Gα13 and Gα11 (see Materials and Methods).

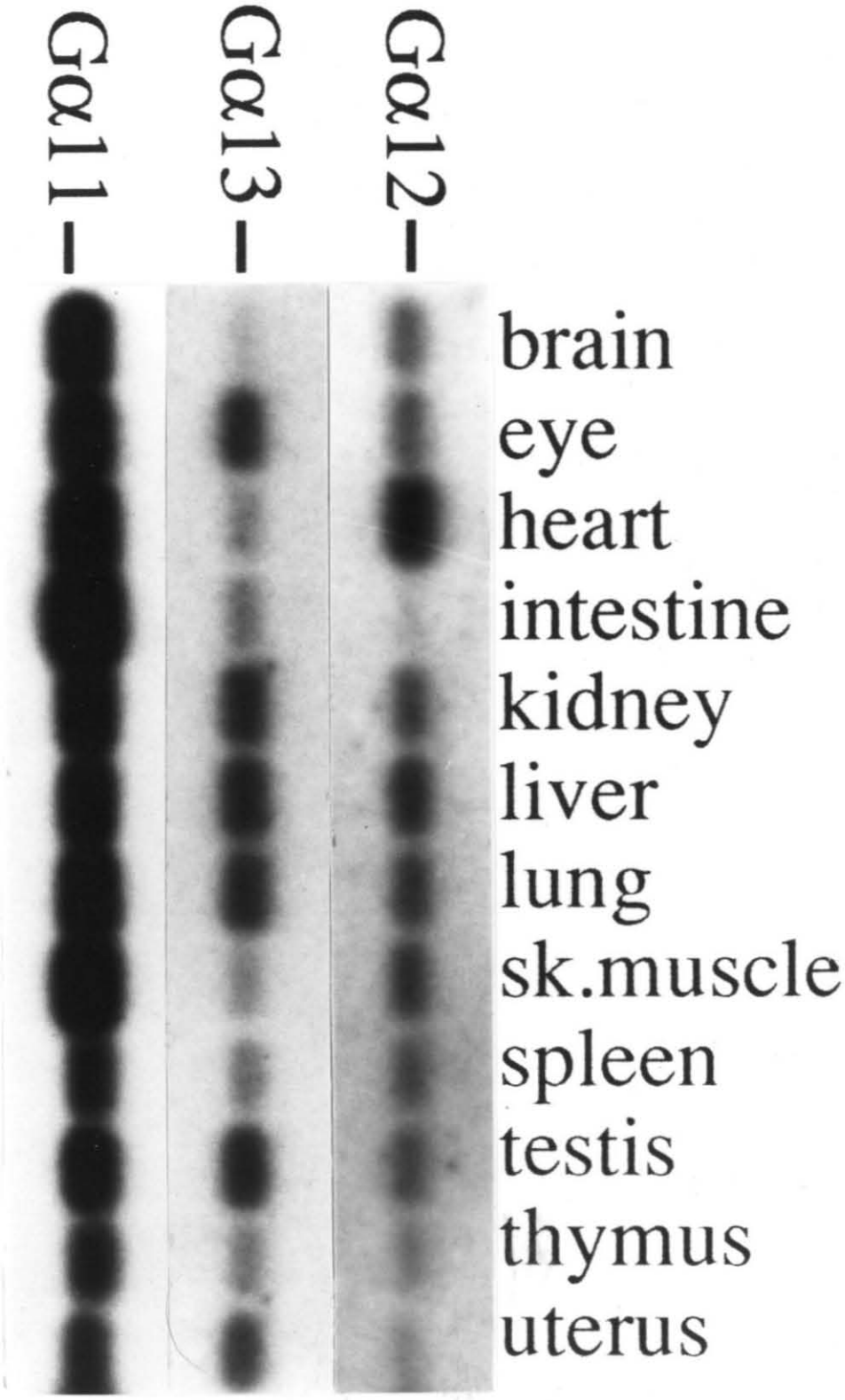
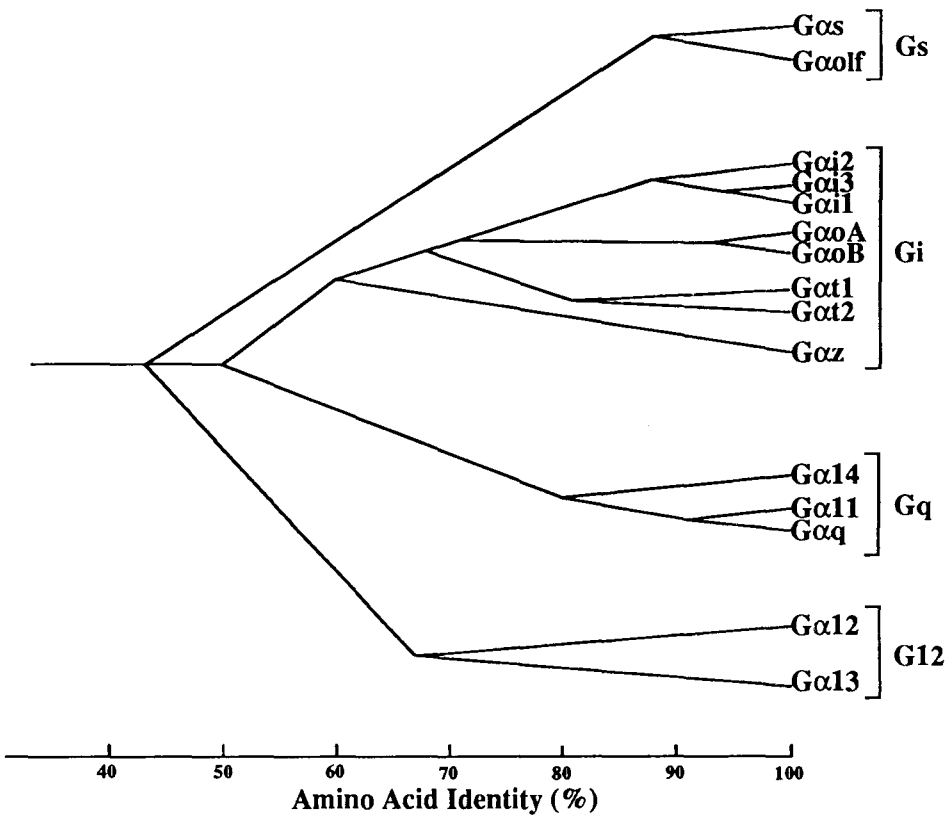

Gα11
Gα13
Gα12
brain
eye
heart
intestine
kidney
liver
lung
sk.muscle
spleen
testis
thymus
uterus

Figure 4. Relationships among mammalian G Protein alpha subunits. Alpha subunits are grouped by amino acid identity (2, 10, 26, 27). Branch junctions approximate values calculated for each pair of sequences.


Gαs
Gαolf
Gs
Gαi2
Gαi3
Gαi1
GαoA
GαoB
Gαt1
Gαt2
Gαz
Gi
Gα14
Gα11
Gαq
Gq
Gα12
Gα13
G12
40
50
60
70
80
90
100
Amino Acid Identity (%)

Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 1247–1250, February 1991
Genetics

Transposon-facilitated DNA sequencing

(γδ/polymerase chain reaction)

MICHAEL STRATHMANN, BRUCE A. HAMILTON, CAROL A. MAYEDA, MELVIN I. SIMON, ELLIOT M. MEYEROWITZ, AND MICHAEL J. PALAZZOLO

California Institute of Technology, Division of Biology, 147-75, Pasadena, CA 91125

Communicated by E. B. Lewis, September 26, 1990

ABSTRACT **We describe here a transposon-based DNA sequencing strategy that allows the introduction of sequencing priming sites throughout a target sequence by bacterial mating. A miniplasmid was designed to select against transposon insertions into the vector. Sites of transposon insertion are mapped by the polymerase chain reaction with bacterial overnight cultures providing the templates. A small set of plasmids with transposons spaced several hundred base pairs apart can then be sequenced. Sequencing primers corresponding to the transposon ends allow sequencing in both directions. Thus, the entire sequence of both strands can be easily determined.**

One of the major problems in DNA sequence analysis of large or even moderately sized fragments is how to position unsequenced regions next to known priming sites. A variety of techniques have been developed for this purpose including random shotgun subcloning, unidirectional deletions and subcloning, and the continued synthesis of additional oligodeoxynucleotide primers (1–4). These methods are expensive or require many molecular manipulations.

A number of strategies employ bacterial transposons to generate priming sites within a target DNA sequence (5–10). Several criteria exist for an efficient transposon-based sequencing strategy: (*i*) Mobilization of the transposon must be relatively simple. (*ii*) Selection for transposon insertions into the plasmid as opposed to the bacterial chromosome must be efficient. (*iii*) The transposon must insert into the target sequence and not into the plasmid vector. (*iv*) The transposition sites must be easily mapped to minimize the number of required sequencing reactions. In this paper we describe a transposon-based strategy that meets these criteria.

We employ γδ, which belongs to the Tn*3* family of transposons (11) and which has been used previously in transposon-facilitated strategies (8, 20). The members of this family contain 38-base-pair (bp) terminal inverted repeats and transpose by a replicative mechanism. Donor and target sequences are joined in an intermediate structure termed a cointegrate. The cointegrate, which contains two copies of the transposon, is rapidly resolved by a site-specific recombination system. The resolvase is encoded by the transposon and acts at the 120-bp *res* site located within the mobile element.

γδ is present on the F factor. Consequently, transposition to a plasmid transiently fuses the F factor and plasmid in a cointegrated structure. This cointegrate can be transferred to a recipient cell by conjugation. Resolution of this structure in the recipient yields the F factor and the plasmid each with a single γδ insertion (Fig. 1).

This paper describes the use of conjugal transfer of a plasmid to introduce γδ insertions into a target sequence (12). The target DNA fragment has been subcloned into a minimal plasmid in which nearly all the plasmid sequences are selectable. Under these conditions recovered transpositions almost always contain a γδ transposon inserted into the target sequences. The sites of insertion can be readily mapped by polymerase chain reaction (PCR) (13). Finally, orientation-specific sequencing primers allow sequence analysis in both directions from the insertion point.

METHODS

Bacterial Mating. Two *Escherichia coli* strains are grown overnight under appropriate antibiotic selection. The donor strain, DPWC (*supE42 ΔrecA*[*Sst II–Eco*RI] *srl*::Tn*10*-[Tet^s], F$^+$), contains the target plasmid, which confers resistance to ampicillin. The recipient, strain JGM, is strain MC1061 (14) that carries Tn*5*seq1 (9) and is F$^-$ and kanamycin resistant. One-tenth milliliter of each overnight culture and 2 ml of LB medium are combined in a sterile 15-ml tube and incubated on a rotary wheel (30 rpm) for 3 hr or longer at 37°C. One-tenth milliliter of a 100-fold dilution of the mating mixture is plated on an LB-agar plate containing both ampicillin at 100 μg/ml and kanamycin at 50 μg/ml. This plate is incubated overnight at 37°C. Generally, this procedure yields 50–500 Amp^{r}Kan^r colonies. Plating nondiluted and 10-fold dilutions of the mating mixture results in bacterial lawns that consist primarily of bacteria resistant only to kanamycin.

Construction of the Miniplasmid Vector. The miniplasmid vector, used as a transposon target, was constructed by the PCR (14) followed by standard recombinant DNA techniques (15). (*i*) Two oligonucleotides were synthesized that would PCR-amplify the β-lactamase gene and the replication origin of the pUC plasmids. One oligonucleotide, N-AMP (5′-ATGAGACAATAACCCTGA-3′), hybridizes just upstream of the β-lactamase gene (near position 4210 of the pBR322 map) (16). The second oligonucleotide, ORI F (5′-GCCCGGGCGTTGCTGGCGTTTTTCC), is located around pBR322 position 2520 and contains a *Sma* I site. A PCR was performed by using this oligonucleotide pair as primers and pBluescriptKS2 (Stratagene) as template. The PCR product was phosphorylated with T4 polynucleotide kinase, ligated, and introduced into *E. coli* to generate plasmid pOAS. The polylinker from pBluescriptKS2 was then introduced into pOAS to generate pMOB (Fig. 2). To accomplish this construction, the polylinker was first PCR-amplified by using the reverse and universal sequencing oligonucleotides that are commercially available (New England Biolabs). Plasmid pOAS was linearized with *Sma* I and ligated to the polylinker.

DNA Sequence Analysis. Templates for DNA sequence analysis were prepared by using rapid-boil plasmid preparations (17). The sequencing reactions were done with the United States Biochemical T7 Sequenase version 2.0 kit according to the enclosed protocols. The sequencing primers are oligonucleotides complementary to transposon sequences adjacent to the inverted repeat ends. GD1 (5′-CAACGAATTATCTCCTT-3′) will sequence outward from the γδ end

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: PCR, polymerase chain reaction.

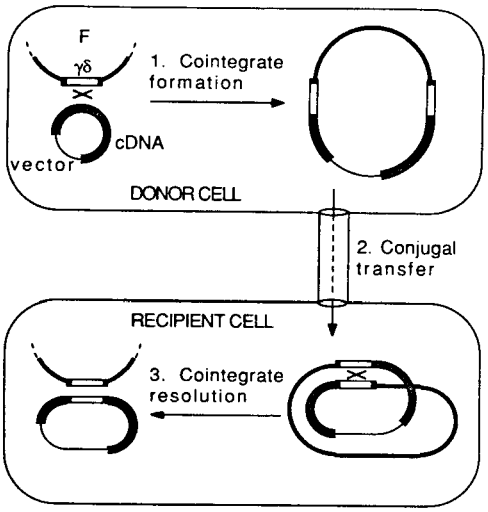


FIG. 1. Transposon mobilization into target DNA sequences. This diagrams the probable mechanism of γδ-transposon insertion into the target sequences by bacterial mating. According to this model (5), a cointegrate is formed in the donor cell between the F factor and the target plasmid. In this cointegrate, both plasmids are flanked by γδ transposons. After transfer, the cointegrate is resolved, leaving a transposon copy in each plasmid.

closest to the *Sac* I site in the transposon (12). GD2 (5'-TCAATAAGTTATACCAT-3') will sequence outward from the opposite end.

PCR Conditions. All PCRs were done in 40-μl total volume with a light mineral oil overlay. For the construction steps, a three-step protocol was used that included the following: 94°C for 45 sec, 55°C for 45 sec, and 72°C for 2 min. This protocol was then repeated for 10 cycles. For the transposon mapping experiments a two-step PCR protocol was used: 92°C for 20 sec and 72°C for 2 min. One-half microliter of a bacterial overnight culture was added to the 40-μl reaction mixture and subjected to the PCR for 40 cycles. For each template two PCRs were done. Both PCRs used an oligonucleotide (GDIR; 5'-TTTCGTTCCATTGGCCCTCAAACCCC-3') complementary to the inverted repeat of the γδ transposon. The second oligonucleotide primer used in one PCR (LREV; 5'-AACAGCTATGACCATGATTACGCCAAG-3') was complementary to a sequence just upstream of the T7 promoter (see Fig. 2). The second oligonucleotide primer used in the other PCR mapping (LUNI; 5'-GTAAAACGACGGCCAGTGAGCGCG-3') is complementary to a region immediately adjacent to the T7 RNA promoter. The PCR buffer was purchased from Cetus.

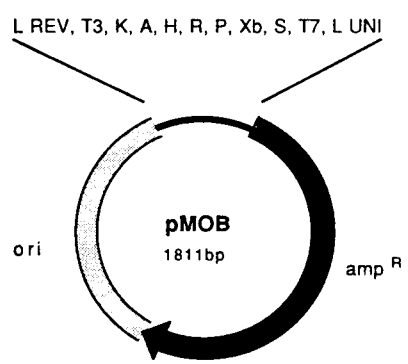


FIG. 2. Map of pMOB, the miniplasmid used for transposon-targeted DNA sequence analysis. This plasmid contains (*i*) a pUC plasmid origin of replication (pUC ori), (*ii*) β-lactamase gene (Amp^r), (*iii*) a multiple cloning site (MCS), (*iv*) bacteriophage RNA polymerase promoters (T3 and T7), (*v*) primer sites for PCR mapping of the transposon insertion sites (LREV, LUNI). The polylinker was derived from a PCR from the plasmid pBluescriptKS2; many restriction sites occur in this polylinker. We have tested for and used only the *Apa* I (A), *Eco*RI (R), *Kpn* I (K), *Pst* I (P), *Hin*dIII (H), *Sac* I (S) and *Xba* I (Xb) sites (these sites are all unique and found only in the MCS region). Because this plasmid was constructed from DNA fragments amplified by PCRs, the DNA sequence may vary somewhat from sequence of the parent plasmids.

RESULTS AND DISCUSSION

Efficiency of γδ Transposition into Target Plasmids. Transposition of γδ from an F factor to a plasmid is believed to produce a cointegrate that can be conjugally transferred. After conjugation the cointegrate is resolved in the recipient cell (Fig. 1).

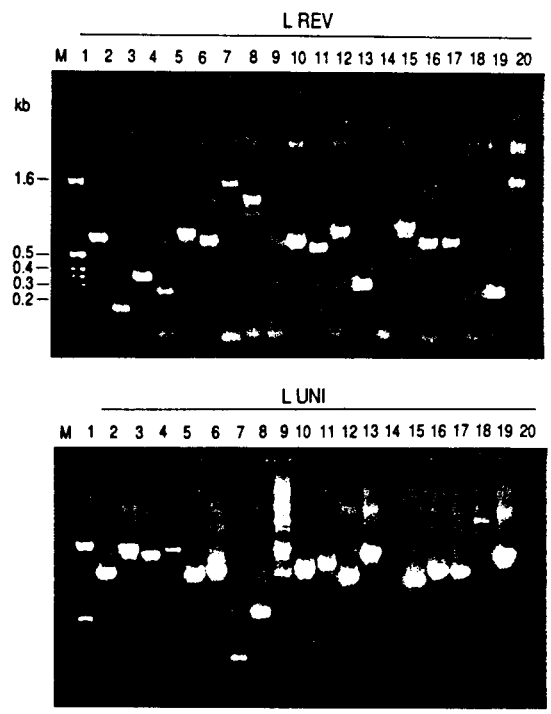


FIG. 3. PCR mapping of transposon-insertion sites. γδ transposons were introduced by bacterial mating into 1.7-kb cDNA (1H2) that was subcloned into pMOB. Bacterial overnight cultures from each of 20 individual transposition events provided the templates in two separate PCRs. In the first set of PCRs LREV and GDIR were the primers; PCR products were then analyzed by agarose gel electrophoresis (*Upper*, labeled LREV). The second set of PCRs were identical, except that LUNI and GDIR oligonucleotides were the primers (*Lower*); results of these PCRs are labeled LUNI. Lanes in each of these gels are matched so that the upper gel shows the PCR products generated by using the first set of primers, and the bottom gels show the analogous products with the second set of primers. Exact sequences of the oligonucleotides and PCR conditions are indicated in text; interpretation of this data is presented in Fig. 4. As mentioned in text, some PCRs gave ambiguous results: for example, multiple PCR products can be seen in lane 9 (*Lower*), whereas no apparent products can be seen in either reaction in lane 14. Size markers (M) are *Hin*fI digests of pBR322.

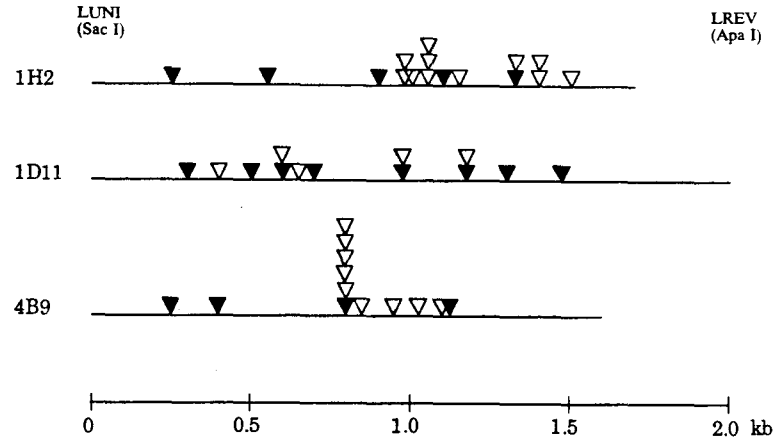


FIG. 4. PCR mapping analysis allows the generation of maps of the γδ-transposon-insertion points. Insertions into three different cDNAs (1H2, 1.7 kb; 1D11, 2.0 kb; and 4B9, 1.6 kb) are shown. These plasmids are all longer versions of cDNA clones originally isolated by Palazzolo *et al.* (19). Each triangle represents the insertion of a single γδ transposon into a given cDNA. PCR mappings that gave ambiguous results were excluded from the map. The filled triangles represent clones selected for DNA sequence analysis. A complete double-stranded analysis was possible from just these selected clones.

Although the mobilized plasmid is expected to contain a single copy of γδ, some experiments (18) indicated that an alternative method of plasmid transfer is possible in certain *E. coli* strains. Examination of transferred plasmids in these strains showed that only 30% of the plasmids contained transposons after transfer. To determine the fraction of mobilized plasmids that contain γδ, *E. coli* DPWC (a donor strain; see *Methods*) was transformed with a 3-kilobase (kb) plasmid vector conferring ampicillin resistance, pBluescriptKS2 (Stratagene), into which a 1.5-kb *Drosophila* cDNA fragment (4B9; ref. 19) had been subcloned. These cells were then mated to JGM (a kanamycin-resistant recipient strain, see *Methods*). Recipient cells that received the plasmid were selected on plates containing ampicillin and kanamycin. Restriction enzyme analysis of 20 mobilized plasmids indicated that, in each instance, the plasmid contained a γδ insertion. Furthermore, restriction fragment length polymorphisms in the restriction enzyme digests suggested that the transposons had inserted at different sites.

Construction of a Miniplasmid Sequencing Vector. An important criterion for the successful application of transposition to DNA sequence analysis is that the transposon be forced into the target DNA sequences and not the plasmid. Restriction analysis of the γδ insertions described above suggested that most insertion events occurred in the vector and not in the insert. We used dideoxynucleotide chain-termination DNA sequencing to identify the transposon-insertion sites in more detail. Of eight clones chosen at random and sequenced, all transposition sites occurred at different locations in the vector.

These experiments suggest that a plasmid, in which insertions into the vector sequences can be selected against, is useful, as then only the transposon insertions in the target can be recovered. Ideally, the vector should contain only an origin of replication, a drug-resistance gene, and a multiple cloning site [the construction of such a miniplasmid, pMOB (Fig. 2), has been described].

To test where transposons insert in this 1.8-kb construct, three different *Drosophila* cDNA molecules (1H2, 1D11, 4B9) (19) were subcloned into this plasmid, and each subclone was separately used as target for γδ transposition. Twenty clones from each of the mating mixtures were selected and analyzed by restriction mapping. All 60 clones contained a γδ transposon, and most insertion sites were in or near the target cDNA fragments.

Analysis of Transposon-Insertion Sites by PCR. The ability to rapidly and simply map the sites of insertion is important for minimizing the labor required to sequence a given target DNA fragment. PCR promised to allow such an identification. For this purpose, we synthesized three oligodeoxynucleotides. One (GDIR) matches the inverted repeat found at each end of γδ. The other two oligonucleotides (LREV and LUNI) flank the cloning site of the miniplasmid (see Fig. 2). Two separate PCRs can be used to determine the point of insertion of a given transposon. In one reaction LREV and GDIR are used as primers, whereas LUNI and GDIR are used in the second reaction. In both cases the same plasmid containing a γδ transposon is the template. Size of the LUNI–GDIR PCR product allows determination of the distance from the transposon-insertion site to the LUNI site at one end of the target fragment, whereas the size of the LREV–GDIR PCR product allows a similar determination of the distance from transposition site to the opposite end of the subclone. Furthermore, the two PCR products should add up to approximately the same size as the fragment subcloned into the miniplasmid.

Such an analysis was performed on the 60 plasmids isolated in the transposon experiments described above. Bacterial overnight cultures provided the templates in two separate PCRs. These reactions were subsequently analyzed by agarose gel electrophoresis. Analysis of 20 transposition events into one clone is shown in Fig. 3. The size of the fragment in each lane delimits the distance of the transposition site from the fixed points in the plasmid.

These experiments allowed us to map the γδ insertion sites for most of the 60 plasmids containing transposons (Fig. 4). Three conclusions can be drawn from these results. (*i*) Forty-two of the 60 transposition events occurred within the cDNA inserts and could be rapidly localized. (*ii*) The insertion sites were sufficiently dispersed within the target to be useful for DNA sequencing. (*iii*) Eighteen of the PCRs gave apparently anomalous results, including multiple PCR products or the apparent absence of PCR products. The transposon insertions that gave such results could not be placed on the transposition maps by this technique and were not further characterized.

DNA Sequence Analysis By Using Transposon Sequences as Priming Sites. Several plasmids were then chosen for DNA sequence analysis of the different cDNA clones. The plasmids that were sequenced were selected because the transposons were spaced ≈300–400 bp apart (Fig. 4). Because end-specific sequences are found immediately adjacent to the inverted repeats, oligonucleotides complementary to these regions can be used to sequence outward from each end of the transposon (see *Methods* for primer sequences). Thus, it is straightforward to simultaneously obtain complete sequence information from both strands of the target sequence from a relatively small number of plasmid–transposon templates. All three plasmids presented in this paper were completely sequenced by a small number of transposon-containing templates (Fig. 4). In addition, we have sequenced six other cDNA clones with inserts from 1.2 to 2 kb. For each clone, a screen of 20–30 $\gamma\delta$ insertions was sufficient to obtain a subset of transposon insertions that were spaced every 300–400 bp along the cDNA insert.

Some Limitations to this Sequencing Strategy. One major limitation to this strategy is the inability of the PCR to map transpositional events that are relatively distant from the fixed points of the plasmid (LREV and LUNI). Specifically, PCRs typically yield anomalous products on templates in which transposons have inserted >3 kb away from the fixed plasmid point. One potential response to this limitation is the use of strand-switching PCR to map the position of unknown transposon insertions relative to known transposon-insertion points. In other words, two plasmids that contain the same initial insert but have transposons in different locations can be mapped relative to each other in a single PCR. This reaction contains both plasmids as templates but uses only the inverted-repeat oligonucleotide (GDIR) as primer. The PCR product should be the DNA sequence between the two transpositional events, and its size will map the position of the unknown site relative to the known one.

To test this strategy we used various combinations of the plasmids containing the 1H2 cDNA and different transpositional events. PCRs containing different pairwise combinations of 1H2 plasmids as templates and only GDIR as primer resulted in PCR products of the sizes predicted by the results presented in Figs. 3 and 4.

We thank K. VijayRaghavan, Susan Celniker, and Christopher Martin for many helpful discussions. We also thank Douglas Berg for λTnSeq5 and Randy Reed for the sequence of GD1 and GD2. M.J.P. is a Lucille P. Markey Scholar of the Lucille P. Markey Charitable Trust. B.A.H. is supported, in part, by a U.S. Public Health Service Predoctoral National Research Service Award (T32 GM07616). This work was also supported by U.S. Public Health Service Program Project Grant GM40499 (E.M.M.) and by U.S. Department of Energy Grant FG03-89ER60891 (M.I.S.).

1. Barnes, W. M., Bevan, M. & Sons, P. H. (1980) *Methods Enzymol.* **65,** 98–122.
2. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
3. Henikoff, S. (1984) *Gene* **28,** 351–359.
4. Strauss, E. C., Kobori, J. A., Siu, G. & Hood, L. E. (1986) *Anal. Biochem.* **154,** 353–360.
5. Guyer, M. (1983) *Methods Enzymol.* **101,** 362–369.
6. Ahmed, A. (1985) *Gene* **39,** 305–310.
7. Adachi, T., Mizuchi, M., Robinson, E. A., Appella, E., O'Dea, M. H., Gellert, M. & Mizuchi, K. (1987) *Nucleic Acids Res.* **15,** 771–784.
8. Liu, L., Whalen, W., Das, A. & Berg, C. M. (1987) *Nucleic Acids Res.* **15,** 9461–9469.
9. Nag, D. K., Huang, H. V. & Berg, D. E. (1987) *Gene* **64,** 135–145.
10. Barrett, B. K. & Berget, P. B. (1989) *DNA* **8,** 287–295.
11. Grindley, N. F. (1988) *The Recombination of Genetic Material,* ed. Low, K. B. (Academic, New York), pp. 283–360.
12. Guyer, M. (1978) *J. Mol. Biol.* **126,** 347–365.
13. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230,** 1350–1354.
14. Casadaban, M. J. & Cohen, S. N. (1980) *J. Mol. Biol.* **138,** 179–207.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Sutcliffe, J. G. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43,** 77–90.
17. Holmes, D. S. & Quigley, M. (1981) *Anal. Biochem.* **114,** 193–197.
18. Goto, N., Shoji, A., Horiuchi, S. & Nakaya, R. (1984) *J. Bacteriol.* **159,** 590–596.
19. Palazzolo, M. J., Hyde, D. R., VijayRaghavan, K., Mecklenburg, K., Benzer, S. & Meyerowitz, E. M. (1989) *Neuron* **3,** 527–539.
20. Strausbaugh, L. D., Bourke, M. T., Sommer, M. T., Coon, M. E. & Berg, C. M. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6213–6217.

Appendix

Diversity among the beta subunits of heterotrimeric GTP-binding proteins: Characterization of a novel beta-subunit cDNA

Elisabeth von Weizsäcker, Michael P. Strathmann

and Melvin I. Simon*

Division of Biology 147-75
California Institute of Technology
Pasadena, California 91125
Telephone: (818) 356-3944
FAX: (818) 796-7066

*Corresponding author.

Molecular and Cellular Biology, submitted.

Abstract

Heterotrimeric guanine nucleotide binding proteins transduce signals from cell surface receptors to intracellular effectors. The alpha subunit is believed to confer receptor and effector specificity on the G protein. This role is reflected in the diversity of genes that encode these subunits. The beta subunit is thought to have a more passive role in G protein function; biochemical data suggests that beta-gamma dimers are shared among the alpha subunits. There is growing evidence, though, for active participation of beta-gamma dimers in some G protein mediated signaling systems. To further investigate this role, we examined the diversity of the beta subunit family in mouse. Using the polymerase chain reaction, we uncovered a new member of this family, Gβ4, which is expressed at widely varying levels in a variety of tissues. The predicted amino acid sequence of Gβ4 is 79% to 89% identical to the three previously known beta subunits. The diversity of beta gene products may be an important corollary to the functional diversity of G proteins.

Introduction

G proteins are guanine nucleotide binding regulatory proteins composed of three subunits, alpha, beta and gamma. They relay signals from a multitude of cell surface receptors to numerous intracellular effectors including adenylyl cyclases, cGMP-phosphodiesterases, phospholipases and ion channels (for reviews see 1, 2 and 3). The alpha subunit binds GDP which is exchanged for GTP upon interaction with the appropriate activated receptor. This exchange is believed to be accompanied by dissociation of the alpha subunit from the beta-gamma dimer. An intrinsic GTPase activity hydrolyses the GTP to GDP, thereby allowing reassociation of the subunits and termination of the signal.

There is a rapidly growing list of genes that encode alpha subunits. Nonetheless, the diversity of alpha subunits is much less than the number of different receptors that are coupled to G proteins. There may be more than 100 receptors that are linked to GTP binding proteins while thus far sixteen different alpha subunits have been found (1). The alpha subunit is believed to confer receptor and effector specificity on the heterotrimer. However, there is accumulating evidence for a role of the beta-gamma dimer in determining the specificity of G protein function (4). Indeed, genetic studies with signal transducing systems in yeast provide evidence for an active role of the beta-gamma subunit in coupling to an effector; the binding of GTP to the alpha subunit may serve to release and activate beta-gamma function (5,6). In mammalian cells, a role for the beta-gamma subunits in the activation of ion channel function has been suggested (7). This effect may be due to the ability of beta-gamma dimers to activate phospholipase A_2. If beta subunits are integrally involved in influencing receptor or effector function, then we might expect a diverse family of these subunits with a relatively wide range of activities.

Initially, two genes specifying highly homologous beta subunits, Gβ1 and Gβ2, were found (8,9,10,11,). These were correlated to two separable polypeptides with gel migration patterns corresponding to a 35 Kd and a 36 Kd polypeptide (12,13). Recently, a transcript encoding a third highly homologous protein, Gβ3, was described (14). All three of these gene products appear to be distributed in a relatively ubiquitous fashion. This raises the problem that in any reconstitution experiment the beta subunits that are used to interact with alpha subunits represent a mixture of gene products. Furthermore, until we can determine the nature and number of the components that make up the beta-gamma portion of the G protein complexes, we will not be able to rigorously assess their functional role. Therefore, in order to examine the diversity among the beta subunits, we used a Polymerase Chain Reaction technique that has, in the past, been successful in revealing the multiplicity of G protein alpha subunits (15). We found a fourth member of the beta subunit family which is expressed at significantly different levels among a variety of tissues.

Materials and Methods

Polymerase Chain Reaction PCR was performed as described previously (15). cDNA was made from PolyA$^+$ RNA with random hexanucleotide or oligo dT primers using MMLV reverse transcriptase. Conditions were those supplied by the manufacturer (Bethesda Research Labs). The oligonucleotides used for PCR amplification of the cDNA were as follows:

E1, AAGGATCCCARGARGCNGARCARCT;

E5, CCGGAATTCCCARTGCATNGCRTADAT;

E8, GACTCGAGTCGACATCGA$(T)_{17}$;

E10, GGAATTCGTCTAATATGGACTCCG;

E11, GGAATTCGTTGCAGGCCTTCCG; E12, GGACACACGGGCTACTTG;

E18,GGCTCCAGCTCTTCACTTGAG; E19, GGCTGTAACACGGATTTCTCC in which R = A or G, D = A,G, or T, and N = A, C, G or T. PCR was performed on a Perkin Elmer Cetus thermal cycler. During each cycle, samples were denatured for 1.0 min at 94°C, extended for 1.0 min at 72°C, and the different oligonucleotide pairs were annealed for 0.5 min at the following temperatures: E1+E5 - 48°C; E8+E10 - 54°C; E11+E12 - 50°C; E18+E19 - 64°C. Each oligonucleotide was used in the PCR at 10 ng/μl. 35 cycles were performed on approximately 5 ng cDNA in a 50 μl reaction volume. Inverse PCR was performed on cDNA that was circularized by intramolecular ligation as follows: 2 μg oligo dT primed cDNA was synthesized with the BRL cDNA synthesis system, diluted into 500 μl ligation mix, ligated with T4 DNA ligase for 2 hours at room temperature, ethanol precipitated and resuspended in 100 μl H_2O. 2μl of the circularized cDNA was used for PCR. The buffer and taq polymerase were supplied by Cetus.

Northern Analysis Total RNA was run on a 1% agarose gel and transferred to Genescreen (Dupont) as described (15). A probe specific to the Gβ4 3'-untranslated regions was made by PCR amplifying the cDNA clone β4–2 with the complement of oligonucleotide E19 (Fig. 1) and the M13 -20 primer (New England Biolabs) which is specific to the cloning vector. The amplified product was run on a low melting point agarose gels (Seaplaque, FMC), excised, and labeled by random priming as described (15). The RNA blot was hybridized as previously described (15).

PCR Northern PCR was performed on cDNA prepared with reverse transcriptase (supplied by BRL) from total RNA as described above. The degenerate oligonucleotides E1 and E5 were used for 35 cycles of amplification. The PCR products were electrophoresed through a 2% agarose gel, blotted to Genescreen and hybridized according to the manufacturer's instructions. These blots were probed with radiolabeled oligonucleotides specific to β4 (E10), β1 (TCACAAACAATATTG

ATCCA) and β2 (TCACAGCTGGGCTTGACCCA). The oligonucleotides were end-labeled with γ^{32}P-ATP as described (16). Blots were hybridized at 42°C and washed at room temperature twice for 5 min in 0.90 M NaCl/0.90 M sodium citrate (6X SSC)/ 0.1% SDS.

Isolation of cDNAs The β4–2 cDNA clone (Fig. 1) was isolated from a random primed mouse brain cDNA library in the lambda cloning vector mλJP3 (M.P.S. and M.I.S., unpublished). 10^6 clones were screened with the β4 E18-E19 PCR fragment using standard techniques (16).

Results and Discussion

To examine beta-subunit diversity, the polymerase chain reaction (PCR) was used to amplify cDNA from mouse brain. A set of degenerate oligonucleotide primers (E1 and E5, Fig. 1) was designed for use in the PCR. These primers correspond to two domains in which the amino acid sequences of Gβ1 and Gβ2 are completely conserved. PCR amplified products were separated on an agarose gel, excised and subcloned into a plasmid vector. Analysis of twenty clones revealed three different beta-like sequences. Two of these sequences were classified as the mouse Gβ1 and Gβ2 homologs since the deduced translation products were identical to the bovine beta subunits and the nucleotide sequences differed by 15% and 7%, respectively. The third sequence, labeled Gβ4, was unique, though clearly related to Gβ1 and Gβ2.

To obtain more of the Gβ4 sequence, a specific primer (E10) was synthesized which, in combination with an oligo dT primer (E8), was used to selectively amplify the 3' end of the Gβ4 cDNA (17). The 5' end of the coding sequence was also obtained by PCR. Circular mouse brain cDNA was generated by intramolecular ligation (see Materials and Methods). This cDNA was used as a template for inverse PCR (18). Two Gβ4 specific primers (E11, E12), directed towards the noncoding

regions, were used to amplify the 3' end and 5' ends of the coding region as well as the connecting noncoding regions. The PCR product, 1 kb in size, was subcloned and sequenced. Finally, two primers based on the 3' and 5' noncoding regions of Gβ4 were designed (E18, E19). PCR of mouse brain cDNA with these primers resulted in a 1.1 kb product which was subcloned and sequenced. It contained the entire coding region of Gβ4.

In addition, the E18-E19 PCR product was used as a probe for screening a randomly primed mouse brain cDNA library. Approximately 10^6 clones were screened. Four positives were obtained and purified. One clone (β4–2) contains a 2 kb insert which includes the entire coding region. Figure 1 shows the nucleotide sequence and the predicted amino acid sequence obtained from β4–2. The open reading frame encodes a protein of 340 amino acids.

Figure 2 compares the amino acid sequence of Gβ4 to those of Gβ1, Gβ2 and Gβ3. All four beta subunits are closely related; Gβ4 is 89% identical to Gβ1 and Gβ2, and 79% identical to Gβ3. The variation evident among the N-terminal 40 amino acids of the different beta subunits is noteworthy since Cys25 of Gβ1 is known to contact Gγ1. Perhaps the surrounding sequence can influence the pairing of beta with gamma, thereby limiting the pairwise combinations of these betas with the growing number of different gamma subunits (19).

Gβ4 belongs to an expanding set of proteins, including the other beta subunits, that conserve a repetitive segmental structure of about 40 amino acids, the WD-40 motif, which is characterized by a tryptophan-aspartate amino acid pair (20). Members of this family include two yeast proteins involved in RNA splicing, PRP4 (21) and PRP17 (1), a protein critical to cell cycle regulation, CDC4 (8), a protein involved in regulating yeast Ras function MSI1 (22), the 12.3 cDNA in chicken which is linked to

the major histocompatibility locus (23) and Enhancer of split E(spl), the product of a *Drosophila* neurogenic locus (24). The function of the WD-40 motif is unknown.

To examine the distribution of the Gβ4 message, Northern analysis was performed on a variety of mouse tissues. A probe was made from the 3' untranslated sequence of the β4–2 cDNA, and it was hybridized to total RNA (Fig. 3). In contrast to Gβ1, Gβ2 and Gβ3 messages, the levels of the Gβ4 transcript vary significantly among the examined tissues. The Gβ4 message is highest in brain, eye, lung, heart, and testis. As a more sensitive assay of tissue distribution, we examined Gβ4, Gβ1 and Gβ2 using PCR Northern analysis (25). The degenerate oligos, E1 and E5, were used to PCR amplify cDNA from numerous tissues. The population of amplified sequences was separated on a gel, transferred to a nylon membrane, and probed with end-labeled oligos specific for each beta. The results of this analysis (Fig. 4) agree with the Northern data. Gβ1 and Gβ2 transcripts show relatively constant levels of expression, whereas Gβ4 varies significantly among the different tissues.

The existence of diverse sets of beta and gamma subunits raises the possibility of combinatorial assembly of G proteins. Perhaps variations in primary sequence of the beta subunits limits the pairwise combinations with gamma subunits. Alternatively, other signals may prevent the "inappropriate" assembly of a particular heterodimer in the cell. However, the availability of a diverse collection of beta-gamma dimers may be essential for a complex multicellular organism. There is evidence to suggest that all beta-gamma dimers are not functionally equivalent (26,27,28). Since beta-gamma is essential for the interactions of the alpha subunit with receptor, a diverse collection of heterodimers may be critical for meeting the specific requirements of the many different receptors that activate a particular alpha subunit. Furthermore, the ability of different beta-gamma dimers to influence the coupling of receptor to alpha subunit could also enable these dimers to modulate the

kinetics of a signaling response. If the beta-gamma dimers released upon stimulation of a particular G protein coupled receptor can influence G proteins that are not coupled to this receptor, then complex signaling networks could exist in which the dynamics of one system are modulated directly by the released beta-gamma dimers from another system. In fact, G_i is postulated to exert its inhibitory action on adenylate cyclase through the effect of free beta-gamma on the stimulatory G protein, G_s (3). Thus there may exist G protein networks which are not only capable of transducing signals vertically across the cell membrane but also capable of distributing information horizontally among G proteins within the cell.

We thank Bruce Birren, Narasimhan Gautam, Ryn Miake-Lye and Thomas Wilkie for discussions and materials. This work was supported by National Institutes of Health Grant #GM34236.

References

1. Simon, M.I., Strathmann, M.P. and Gautam, N. (1991) *Science*, in press.

2. Lochrie, M.A. and Simon, M.I. (1988) *Biochemistry* **27**, 4957-4965.

3. Gilman, A.G. (1987) *Ann. Rev. Biochem.* **56**, 615-649.

4. Neer, E. and Clapham, D.E. (1988) *Nature* **333**, 129-134.

5. Whiteway, M., Hougan, L., Dignard, D., Thomas, D.Y., Bell, G.C., Saari, L., Grant, F.J., Ohara, P. and Mackay, V.L. (1989) *Cell* **56**, 467-477.

6. Blinder, D., Bonvier, S. and Jenness, D.D. (1989) *Cell* **56**, 479-486.

7. Logothetis, D.E., Kurachi,Y., Galper, J., Neer, E.J. and Clapham, D.E. (1987).*Nature*.**325**, 321-326.

8. Fong, H.K.W., Hurley, J.B., Hopkins, R.S., Miake-Lye, R., Johnson, M.S., Doolittle, R.F. and Simon, M.I. (1986) *Proc. Natl. Acad. Sci USA.* **83**, 2162-2166.

9. Sugimoto, K., Toshihide, N., Tanabe, T., Takahashi, H., Noda, M., Minamino, N., Kangawa, K., Matsuo, H., Hirose, T., Imayama, S. and Numa, S. (1985) *FEBS Lett.* **191**, 235-240.

10. Fong, H.K.W., Amatruda, T.T., Birren, B.W. and Simon, M.I. (1987) *Proc. Natl. Acad. Sci USA.* **84**, 3792-3796.

11. Gao, B., Gilman, A.G. and Robishaw, J.D. (1987) *Proc. Natl. Acad. Sci USA.* **84**, 6122-6125.

12. Gao, B., Mumby, S. and Gilman, A.G. (1987) *J. Biol. Chem.* **262**, 17524-17527.

13. Amatruda, T.T., Gautam, N., Fong, H.K.W., Northup, J.K. and Simon, M.I. (1988) *J. Biol. Chem.* **263**, 5008-5011.

14. Levine, M.A., Smallwood, P.M., Moen, P.T., Helman, L.J. and Ahn, T.G.(1990)*Proc. Natl. Acad. Sci USA.* **87**, 2329-2333.

15. Strathmann, M., Wilkie, T.M. and Simon, M.I. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 7407-7409.

16. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

17. Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8998-9002.

18. Triglia, T., Peterson, M.G. and Kemp, D.J. (1988) *Nucleic Acids Res.* **16**, 8186.

19. Bubis, J. and Khorana, H.G. (1990) *J. Biol. Chem.* **265**, 12995-12999.

20. Gautam, N. and Simon, M.I. (1990) in ADP-ribosylating toxins and G proteins, ed. Moss, J. and Vaughan, M. (Am. So. Microbiol., Washington, D.C.), pp. 371-380.

21. Dalrymple, M.A., Peterson-Bjorn, S., Friesen, J.D. and Beggs, J.D. (1989) *Cell* **58**, 811-812.

22. Ruggieri, R., Tanaka, K., Nakafuku, M., Kaziro, Y., Toh-e, A. and Matsumoto, K. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 8778-8782.

23. Guillemot, F., Billault, A. and Auffray, C. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4594-4598.

24. Hartley, D.A., Preiss, A. and Artavanis-Tsakonas, S. (1988) *Cell* **55**, 785-795.

25. Wilkie, T.M., Strathmann, M.P. and Simon, M.I., in preparation.

26. Cerione, R.A., Gierschik, P., Staniszewski, C., Benovic, J.L., Codina, J., Somers, R.L., Birnbaumer, L., Spiegel, A.M., Lefkowitz, R.J. and Caron, M.G. (1987) *Biochemistry* **26**, 1485-1491.

27. Casey, P.J., Graziano, M.P. and Gilman, A.G. (1989) *Biochemistry* **28**, 611-616.

28. Gautam, N., Northup, J., Tamir, H. and Simon, M.I. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7973-7977.

Figure 1. The nucleotide sequence and predicted amino acid sequence of murine Gβ4. The nucleotide sequence was obtained from the clone β4-2 which was isolated from a mouse brain cDNA library. The oligonucleotide primers used for PCR are indicated by arrows above the nucleotide sequence. At position 431, the β4-2 clone contains an adenine instead of a guanine. Another cDNA clone, β4-4, and several PCR products contain a guanine at this position.

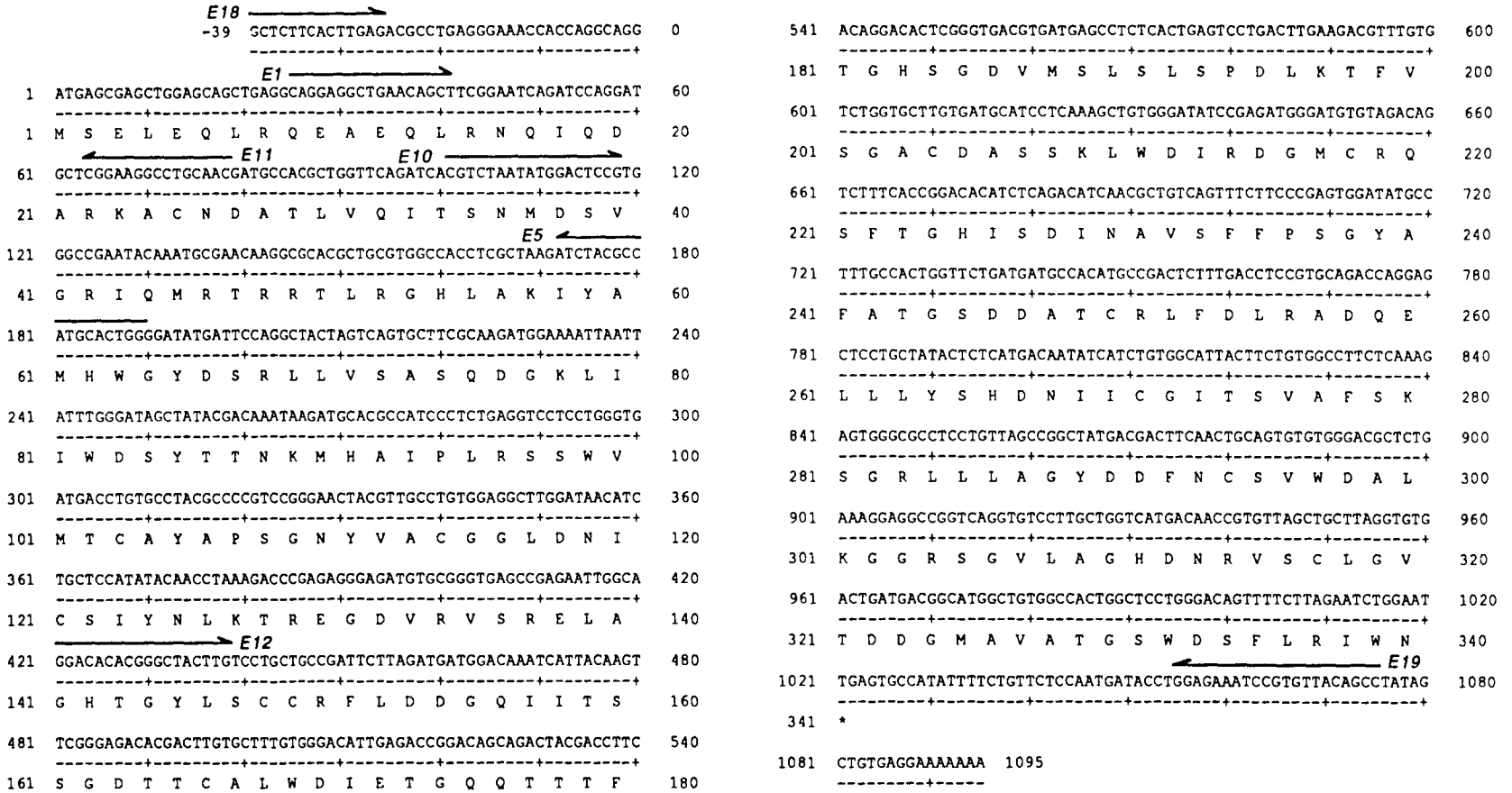
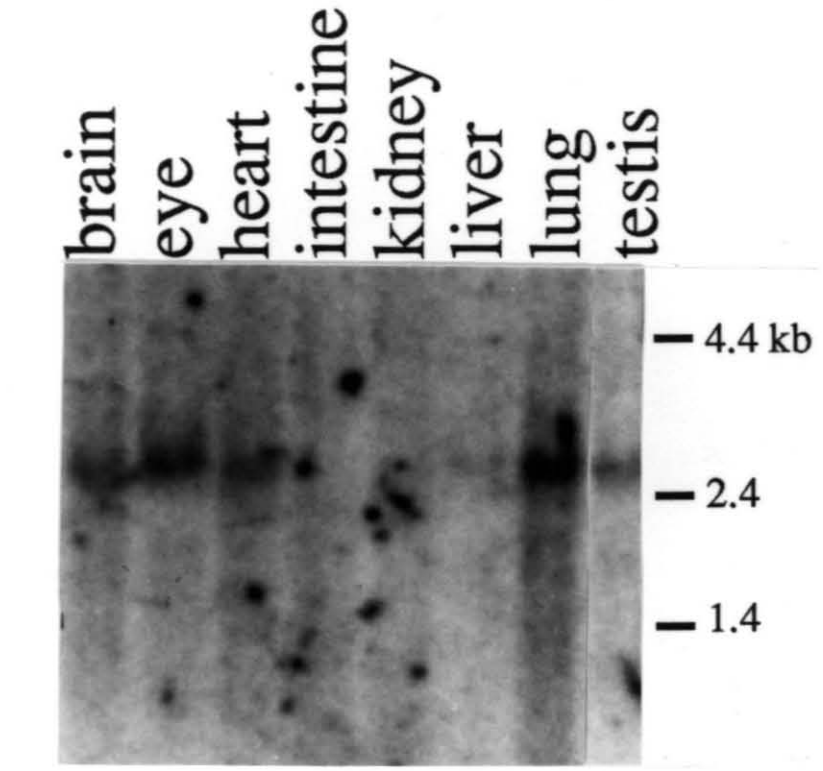

E18
E1
E11
E10
E5
E12
E19

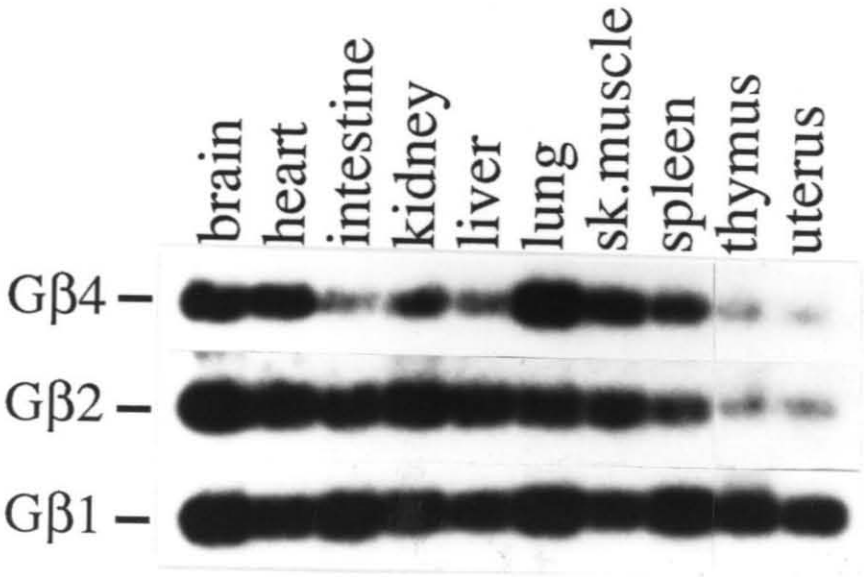
Figure 2. Amino acid sequence comparison of mammalian G protein beta subunits. The sequences are aligned to display the repetitive segmental pattern. Dots represent identities with the Gβ4 sequence. The consensus amino acids in the repeated motif are shown at the bottom of the figure. Aliphatic amino acids are represented by lower case A.

```
Gβ4   MSELEQLRQEAEQLRNQIQDARKACND--------ATLVQITSNMD
Gβ1   ••••D•••••••••K•••R••••••A••••••••••••S•••N•I•
Gβ2   ••••••••••••••••••R••••••G•---------S••T•••AGL•
Gβ3   •G•M••••••••••KK••A••••••A•---------V••AELV•GLE

Gβ4   SVGRI-QMRTRRTLRGHLAKIYAMHWGYDSRLLVSASQDGKLIIWD
Gβ1   P••••-•••••••••••••••••••••T••••••••••••••••••
Gβ2   P••••-•••••••••••••••••••••T••••••••••••••••••
Gβ3   V•••V-•••••••••••••••••••AT••K••••••••••••V••

Gβ4   --SYTTNKMHAIPLR--SSWVMTCAYAPSGNYVACGGLDNICSIYN
Gβ1   --•••••••V•••••••--•••••••••••••••••••••••••••••
Gβ2   --•••••••V•••••••--••••••••••••••F••••••••••••S
Gβ3   --•••••••V•••••••--••••••••••••••F••••••••M•••••

Gβ4   LKTREGDVRVSRELAGHTGYL-SCCRFLDDGQIITSSGDTTCALWD
Gβ1   ••••••N•••••••••••••••-••••••••N••V••••••••••••
Gβ2   ••••••N••••••••P••••••-••••••••N••••••••••••••
Gβ3   ••S•••N•K•••••SA•••••-••••••••NN•V••••••••••••

Gβ4   --IETGQQTTTFT--GHSGDVMSLSLSPDLKTFVSGACDASSKLWD
Gβ1   --•••••••••••••--••T••••••••A••TRL•••••••••A••••
Gβ2   --•••••••••VG•A--••••••••••••A••GR•••••••••I••••
Gβ3   --•••••••K•V•V----•T••C•••AV•••FNL•I•••••••A••••

Gβ4   I--RDGMCR--QSFTGHISDINAVSFFPSGYAFATGSDDATCRLFD
Gβ1   V---•E•••••---•T••••E•••••IC•••N•N•••••••••••••
Gβ2   V----•S•••••---•T•I••E••••••A•••N••••T••••••••••
Gβ3   V---•E•T•----•T••••E•••••IC•••N•E•IC••••••S•••••

Gβ4   --LRADQELLLYSHDNIICGITSVAFSKSGRLLLAGYDDFNCSVWD
Gβ1   --•••••••••MT••••••••••••••S•••••••••••••••••N•••
Gβ2   --•••••••••M•••••••••••••••••R•••••••••••••••NI••
Gβ3   --•••••••••ICF••ES••••••••••••L•••••F••••••••N•••

Gβ4   A-LKGGRS---GVLAGHDNRVSCLGVTDDGMAVATGSWDSFLRIWN
Gβ1   •---AD•A-----••••••••••••••••••••••••••••••••K•••
Gβ2   •-M••D•A-----••••••••••••••••••••••••••••••••K•••
Gβ3   S-M•SE•V----I•S••••••••••••A••••••••••••••K•••
         R          L GH    a         DG   aa GS D   C aWD
```

Figure 3. Northern analysis of Gβ4. Probes specific to the 3-prime untranslated region of Gβ4 (see Materials and Methods) were hybridized to total RNA from various mouse tissues. 20μg total RNA was loaded per lane.

Figure 4. PCR Northern. PCR was performed on various mouse tissues using the degenerate oligonucleotides E1 and E5 (see Materials and Methods). The amplified products were hybridized with radiolabelled oligonucleotides specific to Gβ1, Gβ2, and Gβ4 (see Materials and Methods).


brain
eye
heart
intestine
kidney
liver
lung
testis
4.4 kb
2.4
1.4


brain
heart
intestine
kidney
liver
lung
sk.muscle
spleen
thymus
uterus
Gβ4
Gβ2
Gβ1