Diversity of the G-Protein Family: Sequences from Five Additional alpha Subunits in the Mouse

Michael Strathmann, Thomas M. Wilkie, and Melvin I. Simon

PNAS 1989;86;7407-7409
doi:10.1073/pnas.86.19.7407

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)
Gz:   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).

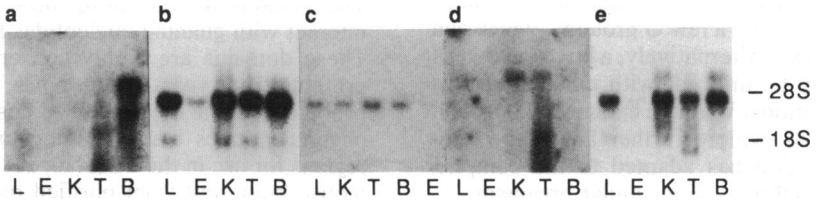

FIG. 2. Northern blot analysis of the five newly discovered α-subunit cDNAs. Poly(A)$^+$ RNA (1 μg) from liver (L), eye (E), kidney (K), testis (T), and brain (B) was fractionated on formaldehyde agarose gels (1%) by electrophoresis and transferred to nitrocellulose. The following radioactive probes 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.