G Protein Diversity: A Distinct Class of \( \alpha \) Subunits is Present in Vertebrates and Invertebrates

M Strathmann, and MI Simon

*PNAS* 1990;87;9113-9117
doi:10.1073/pnas.87.23.9113

This information is current as of December 2006.
G protein diversity: A distinct class of \( \alpha \) 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

**ABSTRACT** Heterotrimeric guanine nucleotide-binding proteins (G proteins) are integral to the signal transduction pathways that mediate the cell's response to many hormones, neurotransmitters, 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 study of genes that encode the \( \alpha \) 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 \( \alpha \) subunits, Gaq and Gal11, 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 \( \alpha \) subunits. Multiple mRNAs as large as 7.5 kilobases hybridize to Gaq specific probes and are expressed at various levels in many different tissues. Ga11 is encoded by a single 4.0-kilobase message which is expressed ubiquitously. Amino acid sequence comparisons suggest that Gaq and Gal11 represent a third class of \( \alpha \) subunits. A member of this class was found in *Drosophila melanogaster*. This \( \alpha \) subunit, DGaq1, is 76% identical to Gaq. The presence of the Gaq class in both vertebrates and invertebrates points to a role that is central to signal transduction in multicellular organisms. We suggest that these \( \alpha \) 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 \( \alpha \), \( \beta \), and \( \gamma \) subunits. The \( \alpha \) subunit is believed to confer receptor and effector specificity on the heterotrimer. When the G protein is activated by interaction with receptor, the \( \alpha \) subunit exchanges bound GDP for GTP. The intrinsic GTPase activity of the \( \alpha \) 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. 1–5). 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 \( \alpha \) 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 \( \alpha \) subunits have demonstrated the involvement of the Gaq and Gal11 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 Gaq (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 \( \alpha \) subunits. We found evidence for extensive diversity in the mammalian G protein \( \alpha \) subunit family (15). A small screen uncovered five novel sequences termed Ga10 through Ga14. In this paper we present the cDNA sequences of two \( \alpha \) subunits that define another class of G proteins. This class, termed Gaq, is distinguished by amino acid sequence homology and includes Gaq, Ga11, and Gal14. We argue that the Gaq class appeared early in evolution; it is found in both vertebrates and invertebrates. We present the sequence of a member of the Gaq 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:

\[
\begin{align*}
\text{oMP19} & \quad \text{CGGATCCCARTTGATHCAYTGYTT} \\
\text{oMP20} & \quad \text{GGATTTCRTCYTTYTTRTNAGRAA} \\
\text{oMP21} & \quad \text{GGATTCRTCYTTYTTRTYAARA} \\
\text{GQ112} & \quad \text{CTGAGGAGTAGACTACAYTGT} \\
\text{GQ3} & \quad \text{GACACGAATGACAGGATGTGCT} \\
\text{GQ113} & \quad \text{CCTCAAGCCACATTGAGTCA} \\
\end{align*}
\]

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 + GQ113, 53°C. Each oligonucleotide was used in the PCR at 10 ng/\( \mu \)l. Thirty-five cycles were performed on approximately 5 ng DNA.

Abbreviation: PCR, polymerase chain reaction.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M55412 for Gaq and M55411 for Ga11.)

9113
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 Gaα, Ga11 or Ga11 3' untranslated regions were made by PCR amplifying the cDNA clones Ga7 and G23 (Fig. 1) with oligonucleotide pairs GQ112 + GQ3 (Gaα) and GQ112 + G113 (Ga11). The amplified products were electrophoresed on low melting point agarose gels (Seaplace, 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 Gaα (ATTCGCTAAGCGCTACTAGA) and Ga11 (CTCGCCTAGTGCCACC). The oligonucleotides were end-labeled with [γ-32P]ATP as described (ref. 16, p. 10.59). Blots were washed at room temperature twice for 5 min in 0.9 M NaCl/0.9 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. A cDNA clone (Fig. 1A) was isolated from an oligo(dT)-primed mouse brain library in the α cloning vector mALSMM (unpublished results); 10^5 clones were screened at a density of 10^6 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 Gaα, Ga11, Ga23, and Gaα (bovine Ga11 and Ga12; and human Gaα). These cDNAs were com-

**Fig. 1.** Sequence of Gaα and Ga11. The sequence of Gaα (A) was obtained from two clones, Ga7 and Ga7. These clones were isolated from a random-primer mouse brain cDNA library by using G69 as a probe. These clones contain the following nucleotides: Ga7, 36 to 966; Ga7, 144 to 1423; and G69, 120 to 629. The Ga11 sequence (B) is contained in two clones, G23 and G115-59. G23, which extends from nucleotide 490 to nucleotide 1307, was isolated from an oligo(dT)-primed mouse brain cDNA library by using the Ga11 PCR fragment (15) as a probe. G115-59 was isolated from the random-primer 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 Gaα 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 NaH₂PO₄ (6× SSPE)/0.1% SDS/5× Denhardt’s solution containing denatured salmon sperm DNA at 100 μg/ml (1× Denhardt’s solution = 0.2% bovine serum albumin/0.2% Ficoll/0.2% polyvinylpyrrolidone). The filters were washed in 6× SSC/0.1% SDS three times for 5 min at room temperature and once at 30°C for 15 min. To obtain the entire sequence of Gaq and Ga11, a random hexanucleotide or oligo(dt)-primed mouse brain cDNA library in μLMSb or λZAPII (Stratagene) was screened by using standard techniques. DGaq 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 Gaq and Ga11. The clone G69, which encodes part of Gaq, 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 Gaq coding sequence (Fig. 1A). The cloned PCR product corresponding to Ga11, 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. 1B shows the sequence of Ga11 derived from overlapping cDNA fragments. The deduced amino acid sequence of Ga11 is 88% identical to that of Gaq. Almost all of the amino acid differences between Gaq and Ga11 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 Gaq and Ga11 are compared with the other α subunits, a number of noteworthy differences emerge. Fig. 1C shows a comparison of the N-terminal sequences of Gaq and Ga11 with those found for other α subunits. The methionine predicted by homology to be the first codon in Gaq and Ga11 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 Ga11 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, Gaq and Ga11 are not substrates for myristoylation (Fig. 1C); this may affect their membrane association properties. However, Gaq 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. 1D 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 within this region affect the GTPase activity of the subunit (18, 19). Gaq [also named GαQ (20, 21)] differs in this region from the other α subunits. The slow rates of guanine nucleotide exchange and GTP hydrolysis exhibited by Gaq may be due in part to these sequence changes (22). Gaq and Ga11 show differences in this region as well (Fig. 1D). 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 Gaq, Gaq2, and Ga11 lack this cysteine. These proteins are refractory to modification by pertussis toxin (2, 22). Fig. 1E shows that Gaq and Ga11 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 Gaq and Ga11. To determine the relative size of the mRNAs corresponding to Gaq and Ga11 and their distributions, Northern hybridization was performed with RNA purified from a variety of mouse tissues. Both Gaq and Ga11 are ubiquitously expressed (Fig. 2). Using a probe specific to the 3′ untranslated region of each cDNA clone, we found that Gaq has multiple messages, whereas Ga11 shows a single band of approximately 4 kb. The three largest Gaq transcripts (approximately 5, 6, and 7.5 kb) were evident (Fig. 2A) 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 Gaq message and the variation in signal strength in some of the tissues led us to use another method to probe for the presence of Gaq and Ga11 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 Gaq11 message (Fig. 2C) was found to be expressed in all tissues that were examined. The Gaq message (Fig. 2C) 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...
reveals three classes of α subunits: Ga, Gαi, and Gαq (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, Gaolf can substitute for Gaq in coupling the β-adrenergic receptor to adenyl cyclase (26). Also, the three Ga subtypes open atrial potassium channels (27). The Gα class includes all the known α subunits that are susceptible to pertussis toxin modification.

Gαq and Ga11 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 Gαq and Ga11 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 Ga14 is a member of the Gαq class. Ga2 and Ga13, however, appear to define yet another class of α subunits.

**Representation of the Gαq Class in Drosophila.** The ubiquitous tissue distribution of Gαq and Ga11 is consistent with a signal transducing role that is basic to a variety of cell types and signaling pathways. If members of the Gαq 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 Gαq, Gαi, and Gαo have been found in a variety of organisms, including Drosophila (29-34). Fig. 3 suggests that the Gαq class may have diverged from the Gαi class before members of the Gq class, Gαi, and Gαq, diverged from each other. Thus it seemed possible that Drosophila would have a Gaq homologue. These arguments encouraged us to screen D. melanogaster α subunits for a member of the Gq 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 Gαq 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 Gαq homologue (DGαq) is compared with Gαi and Ga11 in Fig. 4. DGαq is 76% identical (88% similar) to Gαq. 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. 1D). In addition, DGαq 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 Gαq and Ga11 demonstrates the existence of ubiquitously expressed highly homologous G protein α subunits. During the characterization of these cDNA clones, Pang and Sternweiss used affinity chromatography with immobilized βy 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 DGαq. One of the peptides also showed a sequence identical to that of Ga11. Furthermore, Pang and Sternweiss (36) found that the 42-kDa proteins have the properties that we would ascribe to Gαq and Ga11; 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 the 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). Gαq is thought to be a candidate for the G protein that might mediate pertussis toxin-insensitive activation of phospholipase C. However, Gαq 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, Gαq and Ga11 are ubiquitously expressed. Hence they
are good candidates for this role. Furthermore, there are a variety of isomers of phospholipase C; many cells contain multiple related but distinct phospholipase C gene products (39). Members of the Go class could be involved in the activation of specific phospholipase C isozymes. Speculation about the possible interaction of the Go class with phospholipase C is strengthened by the recent finding that DAG Go message is localized mainly to the Drosophila eye and ocellus structures (40). Earlier work indicated that regulation of phospholipase C in invertebrate eye 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 Go11 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 Go2 and Go11 exist in rat brain extracts, and experiments with specific probes in our laboratory have detected Go2 and Go11 messengers 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 Go2 and Go4 by in vitro assays (27, 42). Go2 and Go11 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 βγ dimers. There is evidence to suggest that all βγ 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 Piao for Drosophila RNA and cDNA libraries, Don Soto and Carol Lee for sequencing assistance, and T. Wilkie and Narasimham Gautam for reading the manuscript. This work was supported by Grant GM34236 from the National Institutes of Health.