Identification by molecular cloning of two forms of the α-subunit of the human liver stimulatory (G_s) regulatory component of adenylyl cyclase

Rafael Mattera*, Juan Codina*, Ann Crozat*+, Vincent Kidd*†, Savio L.C. Woo*† and Lutz Birnbaumer*

*Department of Cell Biology and †Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX 77030, USA

Received 5 August 1986

Two DNA molecules complementary to human liver mRNA coding for the α-subunit of the stimulatory regulatory component G_s of adenylyl cyclase were cloned. One of the two forms is a full-length cDNA of 1614 nucleotides plus a poly(A) tail of 59 nucleotides. The deduced sequence of 394 amino acids encoded by its open reading frame is essentially identical to that of the α-subunits of G_s identified by molecular cloning from bovine adrenals, bovine brain and rat brain. Two independent clones of the other type of cDNA were isolated. Both were incomplete, beginning within the open reading frame coding for the α_s polypeptide. One codes for amino acids 5 through 394 and the other for amino acids 48 through 394 of the above described cDNA of 1614 nucleotides, and both have the identical 3'-untranslated sequence. They differ from the first cDNA, however, in that they lack a stretch of 42 nucleotides (numbers 214 through 255) and have nucleotides 213 (G) and 256 (G) replaced with C and A, respectively. This results in a predicted amino acid composition of another α-subunit of G_s that is shorter by 14 amino acids and contains two substitutions (Asp for Glu and Ser for Gly) at the interface between the deletion and the unchanged sequence. We call the smaller subunit α_{s1} and the larger α_{s2}. This is the first demonstration of a structural heterogeneity in α_s subunits that is due to a difference in amino acid sequence.

G protein *Adenylyl cyclase* *Cholera toxin* *Receptor coupling* *cAMP* *cDNA*

1. INTRODUCTION

Stimulation of adenylyl cyclase by hormones and neurotransmitters depends on activation of a coupling protein commonly called G_s (formerly also N_s). This coupling protein is ADP-ribosylated by cholera toxin, which facilitates its activation by GTP by inhibiting its GTPase activity (for review see [1]). G_s is part of a family of signal transducing G proteins. They are all of an approximate M_r of 80000 [2,3] and heterotrimers of composition $\alpha\beta\gamma$ [3–5]. Although they share common β- and γ-subunits, they differ from each other in their α subunit. This subunit binds GTP and analogs, regulates the effector system and, in the case of α of G_s, is the substrate for cholera toxin (for review see [6]).

Although the β- and γ-subunits of G proteins, of apparent M_r 35000 and 6000, respectively, were discovered only upon analysis of the puriied proteins, the existence and some of the molecular properties of α-subunits of G_s were known previously. Notable among these are not only their function to mediate stimulation of adenylyl cyclase, but also their being a heterogeneous mixture of components that, when ADP-ribosylated

+ Present address: The Population Council, Rockefeller University, New York, USA
† Present address: Department of Cell Biology, University of Alabama, Birmingham, AL, USA

Dedicated to Dr Luis F. Leloir on the occasion of his 80th birthday, 6 September 1986

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

and subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE), migrate with apparent M_r between 42000 and 52000 [7–9]. In fact, most tissues contain at least two forms of α_s, one migrating as a relatively narrow band at 42–45 kDa and the other migrating as a broader band between 46–52 kDa. The M_r values reported for these forms of α_s are approximate, and vary depending on the reporting laboratory [7–17]. The broader band is often reported as a doublet [9,11,17]. This raises the possibility that there exist not only two, but three types of forms of α_s subunits.

Recently, Gilman and collaborators [18,19], as well as research groups from Japan [20,21], succeeded in cloning full-length cDNA's encoding α-subunits of bovine adrenal G_s [19], bovine brain G_s [20] and rat brain G_s [21]. In spite of minor differences in composition, involving a total of only three amino acids, it is clear that the mRNA's from which the cloned cDNA's were derived code for the same type of α-subunit. It has a total of 394 amino acids and a calculated M_r, in the case of bovine adrenal α_s, of 45692. Over-expression of the bovine adrenal cDNA in COS cells showed the product to migrate on SDS-PAGE as a molecule with an apparent M_r of 52000 [19].

We now report the molecular cloning of cDNA's, derived from human liver mRNA, encoding for two α-subunits of G_s, one similar to that obtained previously by others and another which is shorter by 14 amino acids.

2. MATERIALS AND METHODS

A cDNA library, constructed as in [22,23] in the cloning vector λgt11, was made using poly(A) RNA [24,25] from a human liver [26]. 100000 recombinant phages of this library were screened at a density of 20–25000 phages per 150 mm Petri dish using a replicate plaque amplification technique [27] and ^{32}P-labeled oligonucleotide A (5′-CATTTGCTTCACAATGGTGCTTTTACC-3′) as probe, which is part of the antisense strand of the coding sequence of the GTP binding region (amino acids 52–60) of bovine brain α_s cDNA [18]. To this end filters were prehybridized in 6 × SSC, 5 × Denhardt's solution, 300 mM sodium phosphate, pH 6.8, 0.1% SDS, and 0.1 mg/ml sheared herring sperm DNA [28], for 4 h at 37°C, and then hybridized with the same solution plus 0.5 × 10^6 cpm/ml of oligonucleotide A phosphorylated with T_4 polynucleotide kinase using [γ-^{32}P]ATP of 4500 Ci/mmol. The filters were then washed extensively at 4°C with 6 × SSC, dried and subjected to autoradiography for 17 h at −70°C in the presence of two Dupont Cronex lighting plus enhancing screens using Kodak X-Omat AR X-ray films.

Screening of the cDNA library with [^{32}P]cDNA encoding parts of α_s was with hybridization solution containing 3 × SSC, 10 × Denhardt's solution, 0.1% SDS, 0.05 mg/ml sheared herring sperm DNA and 0.5 × 10^6 cpm [^{32}P]cDNA, incubating before and during hybridization at 68°C, performing final washes with 3 × SSC at 68°C.

Phages giving duplicate signals in the above described screening procedure were plaque purified [28]. Their inserts were excised with *Eco*RI, isolated by electrophoresis in 1% low melting point agarose and subcloned into M13 mp19 for sequencing as in [29] using the buffer system of [30]. Southern blots [31] were done onto nylon membranes (Zeta probe, BioRad).

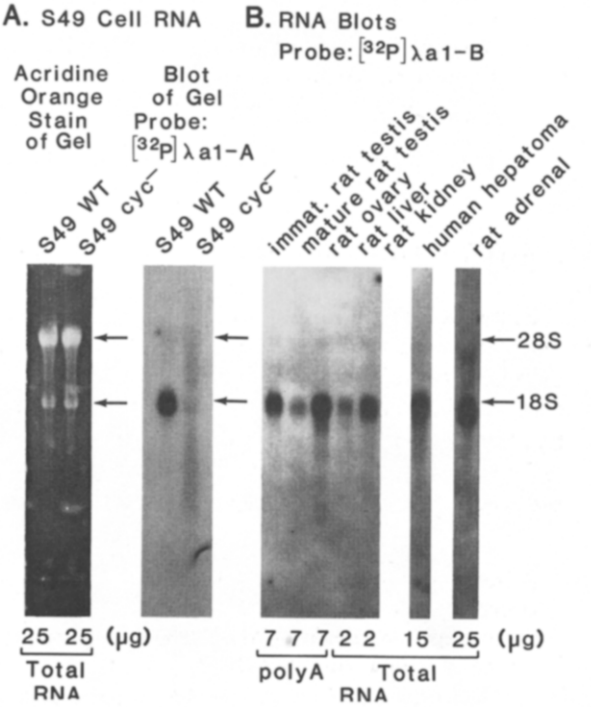


Fig.1. Northern analysis [31,32] of RNA fractions isolated from the indicated cells or tissues. Immature and mature rats were of 10 and 180 days of age.

Total [24] and poly(A) [25] RNA were fractionated for Northern analysis by electrophoresis in 1.5% agarose gels containing 6% formaldehyde [32] and analyzed after transfer onto nitrocellulose filters as in [33].

3. RESULTS

Screening of the λgt11 human liver cDNA library with oligonucleotide A led to the isolation of two recombinant phages: one with an insert of

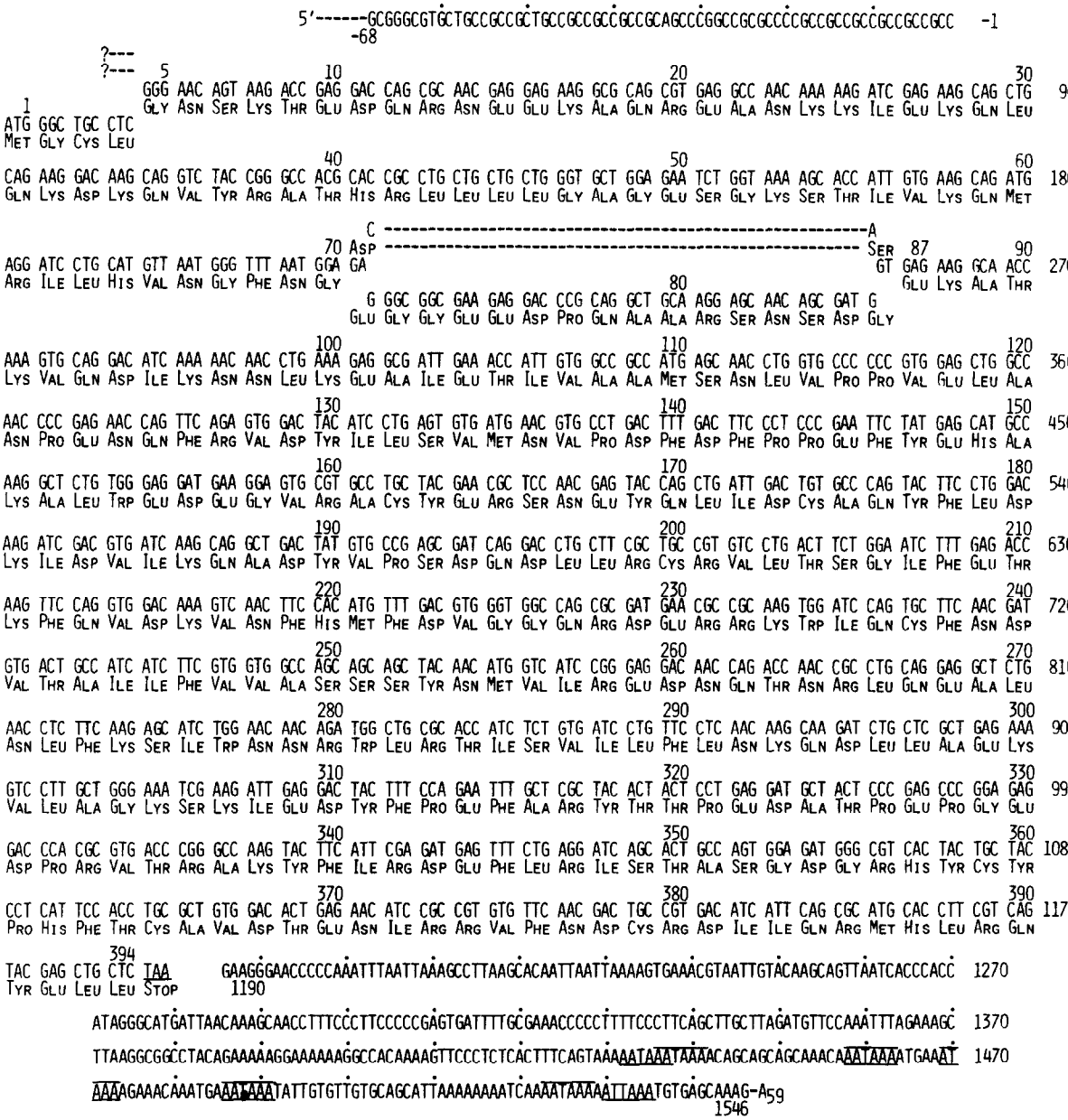


Fig.2. Nucleotide sequence and predicted amino acid composition of the human liver α_s subunits. The CG content of the 68 nucleotide 5'-leader sequence is 94%, that of the 1182 nucleotide coding region is 66% and of the 364 nucleotide 3'-untranslated region, excluding poly(A), is 33%. One ATTAAA and six AATAAA (two overlapping) cleavage and polyadenylation signals are indicated.

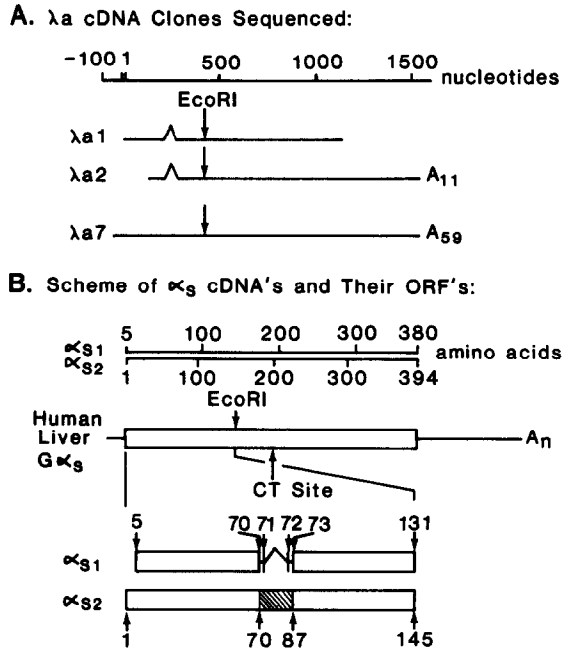
1167 nucleotides (λa1) and another with an insert of 1416 nucleotides (λa2). Each of these inserts had an internal *Eco*RI restriction site, leading to the isolation of four *Eco*RI DNA fragments: λa1-A (789 nucleotides), λa1-B (378 nucleotides), λa2-A (1167 nucleotides) and λa2-B (249 nucleotides), where suffixes A and B refer to the longer and shorter of fragments from each insert, as detected by agarose gel electrophoresis. All fragments were subcloned into M13 mp19 and sequenced. They were found to be partial mRNA copies. Alignment of the resulting nucleotide sequence and analysis for presence of an open reading frame revealed that the two cDNA clones (λa1 and λa2) were derived from a single mRNA species and had an open reading frame encoding a polypeptide of at least 376 amino acids, followed by a 364 nucleotide long 3′-untranslated region terminated by a poly(A) tail of 11 nucleotides.

Northern analysis of total RNA from wild type and *cyc*⁻ S49 murine lymphoma cells (the latter being devoid of RNA sequences coding for α_s [18]), using any one of the *Eco*RI fragments described above (shown in fig.1 for λa1-A), confirmed that cDNA's coding for the α-subunit of G_s had been isolated. Northern blots of other tissues indicated, as expected, the presence of α_s subunit mRNA of a size slightly smaller than 18 S RNA, in agreement with similar analyses reported [18].

The predicted amino acid composition encoded in the two cDNAs cloned from the human liver λgt11 library differed from those reported in [19–21], in that our cDNAs did not code for a stretch of 14 amino acids present in position 72–85 of the 394 amino acid long bovine adrenal, bovine brain and rat brain (glioma) α_s subunits.

Using nick-translated λa2-B cDNA as probe, 46 additional recombinant phages were identified as possible candidates for containing α_s cDNA inserts. Of these, insert λa7 was longer than λa1 or λa2 and like λa1 and λa2 contained one internal *Eco*RI restriction site. On subcloning into M13 mp19 and sequencing, λa7 was determined to be a full-length cDNA encoding for a polypeptide of 394 amino acids, essentially identical to those described for α_s previously, and hence to be derived from an mRNA distinct from the one giving rise to the λa1 and λa2 cDNA clones.

Fig.2 presents the complete nucleotide sequence and the predicted amino acid sequence of the two human liver α_s subunits identified by these experiments, including both the 5′- and the 3′-untranslated sequences. We call α_{s1} the shorter and α_{s2} the longer of the two α subunits of G_s. Where the composition of α_{s1} differs from α_{s2}, i.e. from nucleotide 213 to nucleotide 256 of the longer α_{s2} sequence, the composition of the α_{s1} cDNA is



C. Comparison of Sequences of Various G∝ Subunits in the Region of Difference between Human Liver α_{s1} and α_{s2}:

Human α_{s1} 60–85:	MRILHVNGFNG----------SEKATKVQDIKNNL
Human α_{s2} 60–99:	MRILHVNGFNGEGGEEDPQAARSNSDGEKATKVQDIKNNL
Bovine α_s 60–99:	MRILHVNGFNGEGGEEDPQAARSNSDGEKATKVQDIKNNL
Rat α_o "9–33":	MKIIHEDGF---SGEDVKQ----------YKPVVYSNT
Bovine α_i 53–77:	MKIIHEAGY---SEEECKQ----------YKAVVYSNT
Bovine α_t 49–73:	MKIIHQDGY---SLEECLE----------FIAIIYGNT

Fig.3. A summary of the work presented. (A) Length and relative alignment of cDNA clones sequenced; the scale is in nucleotides. (B) Deduced structure of the two cDNA species characterized. Open box, open reading frame; line, untranslated nucleotide sequences; the scale is in amino acids (α_{s1} upper and α_{s2} lower scale). The region to the left of the *Eco*RI restriction site is expanded to allow for better comparison between α_{s1} and α_{s2}. (C) Comparison of the amino acid composition of α-subunits of various G proteins in the region of difference between α_{s1} and α_{s2}.

shown above that of the α_{s2} cDNA; only one composition is presented where the composition of α_{s1} is the same as that of α_{s2}. Although we are assuming that the open reading frame of the cDNA coding for α_{s1} extends in its 5′-direction for only 12 additional nucleotides more and that the first nucleotide triplet listed (GGG:Gly) is indeed coding for the 5th amino acid of the polypeptide encoded by the mRNA that gave rise to λa1 and λa2, this still needs to be established.

Fig.3 summarizes the work presented here by showing schematically the cDNA clones sequenced (panel A), a drawing of the deduced linear structure of the full-length cDNAs encoding human liver type 1 and type 2 α_s, assuming α_{s1} to be of 380 and α_{s2} of 394 amino acids, and a comparison of the region of difference in the predicted amino acid composition of α_{s1} and α_{s2} to the homologous regions of other signal transducing proteins.

4. DISCUSSION

Direct purification of G_s from liver [4], as well as the ADP-ribosylation studies mentioned in section 1 suggest the existence of at least two, if not three, mature forms of α_2 subunits. One of these forms seems to be of M_r 42000–45000, the other seems to be of M_r 47000–52000 and appears often as a doublet. The functional difference(s) between G_s with one or the other type of α-subunit is unknown. G_s from both, human erythrocytes – which only have the ~42 kDa α-subunit [3,34] – and transformed lung fibroblasts – which have only a single ~52 kDa α-subunit [34] – reconstitute the adenylyl cyclase system of the α_s^- *cyc*$^-$ S49 cell membranes in all of its aspects: nucleotide and aluminum fluoride stimulation and hormonal regulation [34]. The molecular basis for the differences in migration among the different α_s subunits is not known and could be due to post-translational modifications of a single polypeptide or to existence of more than one mRNA species directing their synthesis.

We now report the cloning from a human liver λgt11 cDNA library of two cDNAs. One is full-length and encodes a 394 amino acid α-subunit of G_s (called α_{s2}). The other is an incomplete, yet sufficiently long cDNA to permit the definition of an α-subunit of G_s that is distinct from the 394 amino acid polypeptide. This new form of α_s (called α_{s1}) differs from α_{s2} by lacking 14 amino acids and is likely to be 380 amino acid long.

Together with that of the human form reported on here, there are now four complete amino acid sequences known of the 394 amino acid α_s. It is highly conserved among species and organs, differing only in two amino acids from the bovine adrenal α_s, first reported on by Robishaw et al. [19]. The human polypeptide has Ala^{18} and Ala^{188} substituting for Gly^{18} and Asp^{188} of the bovine adrenal polypeptide. Interestingly, rat brain (glioma) α_s is of the same composition as human liver α_s in positions 18 and 188, but differs in position 139 where it has Asn instead of the Asp found in both the human and the bovine adrenal α_s. The calculated M_r for the human liver 394 amino acid α_{s2} is 45664. Since that of the bovine adrenal α_s is 45692 and, on expression in COS cells, migrates as a polypeptide of apparent M_r 52000 [19], it would appear that the human liver α_{s2} cDNA encodes for a polypeptide that, after full post-translational modification(s), should migrate on SDS-PAGE with an M_r of 50–52000.

Assuming that the amino acid composition in positions 1 through 4 is the same for α_{s1} and α_{s2}, the 14 amino acid shorter α_{s1} would have a predicted M_r of 44294. At this moment we do not know how this polypeptide would migrate on SDS-PAGE. It could have extremely anomalous behavior or be modified post-translationally to migrate with apparent M_r of 42000, as that of human erythrocytes. On the other hand, it might be the second member of the 50–52 kDa [9] – or 47–50 kDa [11,17] – doublet observed previously in ADP-ribosylated samples. The answer to this question will have to await the isolation of the full-length cDNA of this subunit and its expression in cells.

The location of the 14 amino acids that are missing in α_{s1} is of interest. α subunits of G_o, G_i and G_t (transducin) have 361, 355 and 350 amino acids. The predicted α_{s2} sequence, as well as that of the previously cloned α_s subunits, has 394 amino acids. This makes it the longest of the known signal transducing α-subunits. All α-subunits, including α_s, have long stretches of homology [20,21] and the 40–60 'extra' amino acids present in the 394 amino acid form of α_s appear as discrete sets of 'inserts'. When compared to bovine α_i, these inserts are only 6, having lengths of 7, 3, 12,

2, 13 and 2 amino acids and spanning positions 7–14, 69–71, 80–91, 303–304, 327–339 and 357–358, respectively. As illustrated in figs 2 and 3C, the α_{s1} cDNA differs from the α_{s2} cDNA in that the dinucleotide CA in α_{s1} is replaced by a 44-nucleotide long sequence (nucleotides 213–255 of α_{s2}). This sequence codes not only for 14 additional amino acids (positions 72–85), but also changes the identity of each of the flanking amino acids from Asp^{71} and Ser^{72} in α_{s1} to Glu^{71} and Gly^{86} in α_{s2} (shown in shaded boxes in fig.3C). Thus α_{s1} differs from α_{s2} in that the missing-plus-altered amino acids span two of the six 'insertion' regions that arise when comparing the 394 amino acid α_{s2} to the 355 amino acid α_i. Fig.3C presents a comparison of the amino acid sequences of the α-subunits of various G proteins in the region of difference between human liver α_{s1} and α_{s2}.

This report presents the first experimental evidence that α_s subunits exhibit heterogeneity on the basis of amino acid sequence difference. This difference should constitute, at least in part, the molecular basis for the heterogeneity observed in SDS-PAGE analyses. It is tempting to suggest that there also may be a functional correlate to the structural difference observed between α-subunits of G_s.

ACKNOWLEDGEMENTS

This work was supported in part by grants AM-19318, HD-09581, AM-27685 and HD-09495 from the National Institutes of Health, and by a grant from the Welch Foundation, Houston, Texas. We thank T. Chandra, who collaborated with V.K. to construct the human liver λgt11 cDNA library using the cloning system developed by Young and Davis [22]. R.M. is an Honors Fellow of the National Research Council of Argentina (CONICET).

REFERENCES

[1] Rodbell, M. (1980) Nature 284, 11–22.
[2] Sternweis, P.C., Northup, J.K., Smigel, M.D. and Gilman, A.G. (1981) J. Biol. Chem. 256, 11517–11526.
[3] Codina, J., Hildebrandt, J.D., Sekura, R.D., Birnbaumer, M., Bryan, J., Manclark, C.R., Iyengar, R. and Birnbaumer, L. (1984) J. Biol. Chem. 259, 5871–5886.
[4] 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.
[5] Hildebrandt, J.D., Codina, J., Risinger, R. and Birnbaumer, L. (1984) J. Biol. Chem. 259, 2039–2042.
[6] Gilman, A.G. (1984) Cell 36, 577–579.
[7] Gill, D.M. and Meren, R. (1978) Proc. Natl. Acad. Sci. USA 75, 3050–3054.
[8] Cassel, D. and Pfeuffer, T. (1978) Proc. Natl. Acad. Sci. USA 75, 2669–2673.
[9] Johnson, G.L., Kaslow, H.R. and Bourne, H.R. (1978) J. Biol. Chem. 253, 7120–7123.
[10] Hudson, T.H. and Johnson, G.L. (1980) J. Biol. Chem. 255, 7480–7486.
[11] Cooper, D.M.F., Jagus, R., Somers, R.L. and Rodbell, M. (1981) Biochem. Biophys. Res. Commun. 101, 1179–1185.
[12] Graves, C.B., Klaven, N.B. and McDonald, J.M. (1983) Biochemistry 22, 6291–6296.
[13] Rebois, R.V., Beckner, S.K., Brady, R.O. and Fishman, P.H. (1983) Proc. Natl. Acad. Sci. USA 80, 1275–1279.
[14] Watkins, P.A., Moss, J. and Vaughan, M. (1981) J. Biol. Chem. 256, 4895–4899.
[15] Kassis, S. and Fishman, P.H. (1982) J. Biol. Chem. 257, 5312–5318.
[16] Rich, K.A., Codina, J., Floyd, G., Sekura, R., Hildebrandt, J.D. and Iyengar, R. (1984 J. Biol. Chem. 259, 7893–7901.
[17] Malbon, C.C. (1982) Biochim. Biophys. Acta 714, 429–434.
[18] Harris, B.A., Robishaw, J.D., Mumby, S.M. and Gilman, A.G. (1985) Science 229, 1274–1277.
[19] Robishaw, J.D., Russell, D.W., Harris, B.A., Smigel, M.D. and Gilman, A.G. (1986) Proc. Natl. Acad. Sci. USA 83, 1251–1255.
[20] Nukada, T., Tanabe, T., Takahashi, H., Noda, M., Hirose, T., Inayama, S. and Numa, S. (1986) FEBS Lett. 195, 220–224.
[21] Itoh, H., Kozasa, T., Nagata, S., Nakamura, S., Katada, T., Ui, M., Iwai, S., Ohtsuka, E., Kawasaki, H., Suzuki, K. and Kaziro, Y. (1986) Proc. Natl. Acad. Sci. USA 83, 3776–3780.
[22] Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194–1198.
[23] Young, R.A. and Davis, R.W. (1983) Science 222, 778–782.
[24] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[25] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.

[26] Kwok, S.C.M., Ledley, F.D., DiLella, A.G., Robson, K.J.H. and Woo, S.L.C. (1985) Biochemistry 24, 556–561.
[27] Woo, S.L.C., Dugaiczyk, A., Tsai, M.-J., Lai, E.C., Catterall, J.F. and O'Malley, B. (1978) Proc. Natl. Acad. Sci. USA 75, 3688–3692.
[28] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[29] Sanger, F., Nicklen, S. and Coulson, A.B. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[30] Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) Proc. Natl. Acad. Sci. USA 80, 3963–3965.
[31] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
[32] Rave, N., Crkvenjakov, R. and Boedtker, H. (1979) Nucleic Acids Res. 6, 3559–3567.
[33] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[34] Kaslow, H.R., Johnson, G.L., Brothers, V.M. and Bourne, H.R. (1980) J. Biol. Chem. 255, 3736–3741.
[35] Kaslow, H.R., Cox, D., Groppi, V.E. and Bourne, H.R. (1981) Mol. Pharmacol. 19, 406–410.