Inhibitory and stimulatory G proteins of adenylate cyclase: cDNA and amino acid sequences of the α chains

(hormone action/homology/transducin/ras/elongation factor Tu)

Kathleen A. Sullivan^{*}, Yu-Chien Liao[†], Azadeh Alborzi^{*}, Barry Beiderman^{*}, Fu-Hsiung Chang^{*}, Susan B. Masters^{*}, Arthur D. Levinson[†], and Henry R. Bourne^{*‡}

*Departments of Pharmacology and Medicine and the Cardiovascular Research Institute, University of California, San Francisco, CA 94143; and †Department of Molecular Biology, Genentech Inc., South San Francisco, CA 94080.

Communicated by Harold E. Varmus, May 13, 1986

ABSTRACT The G protein family of signal transducers includes five heterotrimers, which are most clearly distinguished by their different α chains. The family includes G_s and G_i, the stimulatory and inhibitory GTP-binding regulators of adenylate cyclase; Go, a protein of unknown function abundant in brain; and transducin 1 and transducin 2, proteins involved in retinal phototransduction. Using a bovine α_{t1} cDNA as a hybridization probe, we have isolated mouse cDNAs that encode α chains of two G proteins. One encodes a polypeptide of 377 amino acids (M_r 43,856), identified as α_s because it specifically fails to hybridize with any transcript in an α_s deficient S49 mouse lymphoma mutant, cyc⁻; the other encodes a polypeptide of 355 amino acids (M_r 40,482), presumed to be α_i . These α chains and those of the retinal transducins exhibit impressive sequence homology. Of the four, α_{t1} and α_{t2} are most alike (81% identical amino acid residues), whereas the presumptive α_i is more similar than α_s to α_{t1} (63% vs. 38% identical residues). Sequence homologies with p21ras and elongation factor Tu identify regions of the α chains that form the site for GTP binding and hydrolysis. Further comparison of the α -chain sequences suggests additional regions that may contribute to interactions with $\beta\gamma$ subunits and the receptor and effector components of different signal transduction systems.

Hormones stimulate cAMP synthesis in mammalian cells by means of G_s, the stimulatory GTP-binding regulatory protein of adenylate cyclase. In addition to G_s , the G protein family of transmembrane signaling molecules includes G_i, the inhibitory regulatory protein of adenylate cyclase (reviewed in ref. 1); Go, a G protein of unknown function that is abundant in brain (2, 3); and two retinal proteins that transduce photon signals in rods (transducin 1) and cones (transducin 2) (refs. 4, 5; J. B. Hurley, personal communication). Recent evidence also points to involvement of G proteins in regulation of phosphotidylinositol metabolism (see ref. 6) and ion channels (7-9). All of the G proteins are heterotrimers, with virtually identical β chains, similar γ chains, and distinctive α chains. The α chains, with $M_{\rm r}$ s ranging from 39,000 to 45,000, bind and hydrolyze GTP and serve as substrates for ADP-ribosylation by the exotoxins of Vibrio cholerae (in G_s and transducin 1) and Bordetella pertussis (in Gi, Go, and transducin 1). In the GTP-bound form, the α chains of G_s (α _s) and transducin 1 (α_{t1}) activate their respective effector enzymes, adenylate cyclase and cGMP phosphodiesterase (for review, see refs. 1, 4).

Using a cDNA encoding bovine α_{t1} (10) as a hybridization probe, we have isolated and sequenced murine cDNAs that encode α_s and a second α chain, presumed to be α_i . Homologies and differences among the deduced amino acid sequences of the G protein and transducin α chains point to specific regions that may interact with guanine nucleotides, receptors, effector enzymes, and the G protein $\beta\gamma$ complex.

MATERIALS AND METHODS

Detection of cDNA Inserts. David Goeddel (Genentech) provided a λ gt10 cDNA library prepared by using RNA from the murine macrophage cell line PU-5 (11). We screened this library for plaques that hybridized (12) to the bovine α_{t1} cDNA (10) at low stringency [in 35% formamide, 0.75 M NaCl, 75 mM sodium citrate, 5× concentrated Denhardt's solution (0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone), 0.5% NaDodSO₄, 0.5 mg of denatured salmon sperm DNA per ml, and 5 × 10⁵ cpm of probe per ml for 48 hr at 42°C] (13). Filters were washed three times for 5 min at 25°C in 0.3 M NaCl/30 mM sodium citrate/0.1% NaDodSO₄ and twice for 30 min at 50°C in 0.15 M NaCl/15 mM sodium citrate/0.5% NaDodSO₄.

RNA Blot Hybridization. Poly(A)⁺ RNA (14) was denatured, fractionated on 1% agarose/2.2 M formaldehyde gels (15), and transferred to Zetaprobe (Bio-Rad). Blots were hybridized at low stringency (described above) or at high stringency (in 50% formamide, 0.75 M NaCl, 75 mM sodium citrate, $5 \times$ concentrated Denhardt's solution, 0.5% NaDod-SO₄, and 0.5 mg of denatured salmon sperm DNA per ml at 42°C for 36 hr). After hybridization, filters were washed as described for low stringency (above), except that washes at 50°C were performed in 15 mM NaCl/1.5 mM sodium citrate/0.1% NaDodSO₄.

DNA Sequencing. The entire insert and restriction fragments of clones containing the longest cDNA inserts for presumptive α_s and α_i were subcloned in phage M13mp18 and M13mp19. The DNA sequence was determined by sequencing both strands by means of the dideoxy chain-termination method (16, 17) using the M13 universal primer and synthetic oligonucleotides derived from α_s and α_i sequences. Two independently derived presumptive α_i cDNAs were sequenced; these differed by a single nucleotide (substitution of cytosine for adenosine in position 960).

RESULTS

Isolation of α_s and α_i cDNAs. The structural homology between α_{t1} and α_o (18) and functional similarities between α_{t1} and other members of the G protein family suggested that the bovine α_{t1} cDNA (10) could serve as a hybridization probe for nucleotide sequences encoding α_s and α_i . We therefore probed poly(A)⁺ RNA transcripts from wild-type and cyc⁻

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: G_s and G_i , stimulatory and inhibitory GTP-binding regulators of adenylate cyclase; G_o , a G protein of unknown function abundant in brain; α_s , α_i , α_o , α_{t1} , and α_{t2} , α chains of G_s , G_i , G_o , transducin 1, and transducin 2, respectively; EF-Tu, elongation factor Tu.

[‡]To whom reprint requests should be addressed.

S49 mouse lymphoma cells (Fig. 1) with radiolabeled α_{t1} cDNA. At low stringency, the α_{t1} probe hybridized to two transcripts in wild-type extracts. The larger transcript (2600 bases) was detected in wild-type and cyc⁻ extracts, whereas the smaller (2100 bases) was detected only in wild-type extracts. Harris *et al.* (19) recently reported that S49 cells bearing the *cyc⁻* mutation lack mRNA detectable by hybridization to a putative bovine α_s cDNA. This finding is in keeping with biochemical (20) and somatic genetic (21) evidence that S49 cyc⁻ cells lack functional α_s activity. We surmised that the transcript lacking in cyc⁻ corresponded to α_s , whereas the larger transcript corresponded to α_i , which is present in wild type and cyc⁻ (22, 23).

The ability of α_{t1} cDNA to detect presumed α_s and α_i messages provided a rationale for using this cDNA under low stringency conditions to screen a mouse macrophage cDNA library for α_s and α_i recombinants. From a screen of $\approx 3.5 \times 10^5$ phage, 12 positive clones were isolated. The cDNA inserts of each were subcloned into pUC13 and used to probe RNA transfer blots of S49 wild-type and cyc⁻ poly(A)⁺ RNA. cDNAs from 5 of the 12 phage hybridized specifically at high stringency to each other and to the 2100-base message that is present in wild type but is not detectable in cyc⁻ (see example in lanes 3 and 4 of Fig. 1); by this criterion, these cDNA inserts were presumed to encode α_s .

The remaining seven inserts hybridized to the 2600nucleotide message present in wild-type and cyc⁻ extracts (Fig. 1); these cDNAs are presumed to correspond to α_i .

The longest of the five original α_s cDNA inserts (corresponding to bases 383–1493 in Fig. 2 *Left*) contained the poly(A)⁺ tract corresponding to the 3' terminus of the mRNA but was too short to contain the entire α_s coding region. Accordingly, we used a restriction fragment representing the 5' end of the original insert as a probe to rescreen the macrophage library for phage carrying overlapping inserts with additional 5' sequence. Five new α_s cDNA inserts were isolated by this method; the longest of these (corresponding to bases 1–780 in Fig. 2 *Left*) was sequenced.

cDNA Sequences. The overlapping α_s inserts specified an open reading frame, extending from nucleotide 1 to the stop codon at position 1132, which would encode a polypeptide of 377 amino acids, with a M_r of 43,856 (Fig. 2 *Left*), similar to the apparent molecular weight of α_s (45,000) on NaDodSO₄/



FIG. 1. RNA blot analysis. Poly(A)⁺ mRNA (5 μ g) isolated from wild-type (lanes 1, 3, and 5) and cyc⁻ (lanes 2, 4, and 6) cells was denatured, separated on 1% agarose/formaldehyde gels, transferred to Zetaprobe, and hybridized under low (lanes 1 and 2) or high (lanes 3-6) stringency conditions. Insert DNA was labeled with [α -3²P]dATP by random priming with calf thymus oligonucleotides (13). cDNA probes used were α_{t1} cDNA (lanes 1 and 2); α_s cDNA (lanes 3 and 4); α_i cDNA (lanes 5 and 6). The positions of 28S and 18S rRNA are indicated by arrows.

polyacrylamide gels (24). The available 5' upstream sequence is too short to establish unequivocally that the methionine predicted at amino acid position 1 (Fig. 2 *Left*) initiates the translated α_s chain; the corresponding DNA sequences conform to the consensus sequence for initiation of translation from eukaryotic messages (25). The α_s cDNA sequence from nucleotides 1 to 55 is unusually rich in guanosine and cytosine and predicts an unusual amino acid sequence at the amino terminus of α_s , including three consecutive glycine residues.§

The presumptive α_i cDNA insert encodes a single open reading frame that includes the indicated coding region extending from nucleotide 1 to the stop codon at nucleotide 1056 (Fig. 2 *Right*). This region would encode a protein of 355 amino acids with a calculated M_r of 40,482, similar to the reported M_r 41,000 size of α_i (27, 28). The methionine residue at amino acid position 1 is presumed to be the site for initiation of translation, based on homology with the deduced amino acid sequences of α_{t1} and α_{t2} (Fig. 3); the 97 bases upstream of this position do not encode a methionine in the same reading frame.

DISCUSSION

Of the two cDNAs reported here, the first is identified as α_s by virtue of its ability to detect a transcript that is present in wild-type S49 cells but absent in the α_s -deficient cyc⁻ mutant and by virtue of its homology to α_{t1} . The predicted size of the α_s polypeptide, M_r 43,856, corresponds to only one of the two α_s chains found in many mammalian cells, with apparent M_r s of 45,000 and 52,000. Somatic genetic analysis in S49 lymphoma cells (21) showed that both polypeptides are encoded by the same gene, and cyc⁻ cells lack both chains as well as the transcript(s) detected by α_s cDNA (Fig. 1 and ref. 19). Although the α_s cDNA provides no hint as to the origin of the M_r 52,000 polypeptide, we presume that the two polypeptides could be produced either by covalent modification or by alternative splicing of an α_s transcript.

Strong but circumstantial evidence suggests that the second cDNA corresponds to α_i . (i) It corresponds to the only G protein α transcript detected by low stringency hybridization to the α_{t1} cDNA in extracts of S49 cyc⁻ cells; these cells contain only one known G protein α chain, that of G_i (22, 23). (ii) It encodes a polypeptide chain whose carboxyl terminus contains a pertussis toxin ADP-ribosylation site almost identical to that of α_{t1} (Fig. 3), in which Cys-347 (corresponding to Cys-352 in the presumed α_i sequence) is the residue covalently modified by the toxin (30). (iii) This cDNA is unlikely to correspond to α_0 , because the smaller (M_r 39,000) α_0 chain has not been detected in S49 cells and because the predicted α_i amino acid sequence differs significantly from homologous sequences of proteolytic fragments of bovine α_o (18); residues 22-34 and 210-234 of the presumed α_i chain differ from the corresponding peptide sequences of α_0 by 42% and 18%, respectively.

Fig. 3 compares the amino acid sequences of murine α_s and presumptive α_i to those of bovine α_{t1} (10, 31, 32), and α_{t2} (5)

[§]The sequence of a cDNA encoding the bovine α_s chain has been reported recently (26). The bovine and mouse α_s chains contain identical residues at 335 positions but differ markedly in two regions. The bovine sequence contains an "extra" region of 15 residues inserted after residue 69 of the mouse sequence. Subsequent experiments revealed a second bovine α_s cDNA lacking the extra region and showed that the shorter and longer cDNAs corresponded, respectively, to the M_r 45,000 and M_r 52,000 α_s polypeptides (50); the two polypeptides are thus presumed to result from alternative splicing of a single bovine α_s transcript. The second major difference is at the extreme amino terminus of α_s , where only 2 of 15 amino acids are identical. Further experiments will be required to determine whether the unusual amino acid sequence predicted by the mouse α_s cDNA in this region represents true interspecies variation or an undetected artefact of the cloning procedure.

-16	GTTGCCGAGGCAGCCC	-97	CTCAGGCCCGCAGCCGAGGCGGTGGGAGCGGAGTGGG
(1)	MetålaålaårgGlyålaålaGlyLeuårgGlyGlyGlyGluLysålaGlnårgGluåla ATGGCGGCGGGGGGGCGGCGGGCGGGGGGGGGGGGGGGG	-60	TCGGGCGGGGCGAGCCGGGCGTGGGGCCTGTGGGGGCCAGGCCGGGCCGGGGGGGG
(21) 61	AsnLysLysIleGluLysGlnLeuGlnLysAspLysGlnValTyrArgAlaThrHisArg AACAAAAAGATCGAGAAGCAGCTGCACGAGGACAAGCAGGTCTACCGGGCCACGCACCGC	(1)	MetGlyCysThrValSerAlaGluAspLysAlaAlaAlaGluArgSerLysMetIleAsp ATGGGCTGCACCGTGAGGCGCGGAGGACAAGGCGGCGGCGGGGCGCGCTCTAAAATGATCGAC
(41) 121	LeuLeuLeuGlyAlaGlyGluGseGlyLysSerThrlieValLysGlnHetArgIle CTGCTGCTGCTGGGTGCTGGAGAGTCTGGCAAAAGCACCATTGTGAAGCAGCATGAGGATC Basby	(21) 61	LysAsnLeuArgGluAspGlyGluLysAlaAlaArgGluValLysLeuLeuLeuLeuGly AAGAACCTGCGGGAGGACGGCGAGAAGGCGGCGCGGGAGGTGAAGTTGCTTCTGTTAGGT
(61) 181	LeuHisValAsnGlyPheAsnGlyAspGluLysAlaThrLysValGlnAspIleLysAsn CTGCATGTTAATGGGTTTAACGGAGATGAGAAGGCCACTAAAGTGCAGGACATCAAAAAC	(41) 121	AlaGlyGluSerGlyLysSerThrIleValLysGlnHetLysIleIleHisGluAspGly GCTGGAGAATCAGGGAAGAGCACCATCGTCAAGCAGATGAAGATCATCCATGAAGATGGC
(81) 241	AspleulysGluAlaIleGluThrIleValAlaAlaMetSerAspleuValProProVal AACCTGAAGGAGGCCATTGAAACCATTGTGGCCGCCATGAGCAACCTGGTGCCCCCTGTG	(61) 181	TyrSerGluGluGluCysArgGlnTyrArgAlaValValTyrSerAsnThrIleGlnSer TACTCAGAAGAGGAGTGCCGGCAGTACCGTGCCGTGGTCTACAGCAACACCATCCAGTCT
101) 301	GluleualaasaProGluasaGlaPbearg¥alaspTyrIleleuSer¥alMetasaLeu GAGCTGGCCAACCCTGAGAACCAGTTCAGAGTGGACTACATTCTGAGCGTGATGAACCTG	(81) 241	IleLeuAlaIleValLysArgMetGlyAsnLeuGlnIeAspPheAlaAspProGlnArg ATACTGGCCATCGTGAAGCGCATGGGCAACCTGCAGATCGACTTTGCTGACCCCCAGCGT
361	ProAsnPheaspPheProProGluPheTyrGluHisAlaLysAlaLeuTrpGluAspGlu CCGAACTTTGACTTCCCACCTGAATTCTATGAGCATGCCAAGGCTCTGTGGGAGGATGAG EcoRI	(101) 301	PstI AlaAspAspAlaArgGlnLeuPheAlaLeuSerCysAlaAlaGluGluGlnGlyMetLeu GCGGATGATGCCAGGCAGCTGTTGGCCCTGTGCTGCAGAGGAACAAGGGATGCTT
(141) 421	GlyValArgAlaCysTyrGluArgSerAssGluTyrGlsLeuIleAspCysAlaGlsTyr GGAGTGCGTGCCTGCTACGAGCGCTCCAATGAGTACCAGCTGATTGACTGTGCCCAGTAC	(121)	ProGluAspLeuSerGlyVallieArgArgLeuTrpAlaAspHisGlyValGlnAlaCys crraagaccrosccostortatccootgcoctCtoGcctaCcAcGcTGTGCAAGCCTGC
(161) 481	PheLeuAspLysIleAspValIleLysGlnAleAspTyrValProSerAspGlnAspLeu TTCCTGGACAAGATTGATGTGATCAAGCAGGCCGACTACGTGCCAAGTGACCAGGACCTG	(141)	PheGlyArgSerArgGluTyrGluLeuAsnAspSerAlaAlaTyrTyrLeuAsnAspLeu
(181) 541	LeuArgCysArgValLeuThrSerGlyIlePheGluThrLysPheGlnValAspLysVal CTTCGCTGCCGTGTCCTGACCTCTGGAATCTTTGAGACCAAGTTCCAGGTGGACAAAGTC	(161)	GluargIleAlaGinSerAspTyrIleProThrGinGinAspValLeuArgThrArgVal
(201) 601	AsnPheHisMetPheAspValGlyGlyGlnArgAspGluArgArgLysTrpIleGlnCys AACTTCCACATGTTCGATGTGGGGCGGCCAGCGCGATGAACGCCGCAAGGGATCCAGTGC BasHI	481 (181)	GAGGGCATTGCACAGAGTGAGTACATCCCTACACAGCAGGAGGGGGGGG
(221) 661	PheAsnAspValThrAlaIleIlePheValValAlaSerSerSerTyrAsnMetValIle TTCAATGATGTGACTGCCATCATCTTCGTGGTGGCCAGCAGCAGCTACAACATGGTCATT	541	ANGACCACGGGCATCGTGGAAACACACACTTCACCTTCAAGGACTTACACTTCAAGATGTTT
(241) 721	ArgGluAspAsnGlnThrásnArgLeuGlnGluAlsLeuAsnLeuPheLysSerIleTrp CGGGAGGACAACCAGACTAACCGCCTGCAGGAGGCTCTGAACCTCTTCAAGAGCATCTGG Det	601	GATGTGGGTGGTCAGCGGTCTGAGCGCAAGAAGTGGATCCACTGCTTTGAGGGCGTCACG BamHI
(261) 781	AsnAsnArgTrpLeuArgThrIleSerValIleLeuPheLeuAsnLysGlnAspLeuLeu AACAACAGATGGCTGCGCACCATCTCTGTGATTCTCTTCCCTAACAAGCAAG	(221) 661	AlaileilePheCysValAlaLeuSerAlaTyrAspLeuValLeuAlaGluAspGluGlu GCCATCATCTTCTGTGTCGCCTTGAGCGCCTATGACTTGGTGCTGGCTG
(281) 841	ÅlæGluLysValleuÅlæGlyLysSerLysIleGluÁspTyrPheProGluPheÅlæÅrg GCTGÅGÅÅAGTCCTCGCTGGCAAATCGÅAGÅTTGAGGACTACTTTCCAGÅGTTCGCTCGC	(241) 721	HetAsnArgHetHisGluSerHetLysLeuPheAspSerIleCysAsnAsnLysTrpPhe ATGAACCGCATGCATGAGAGCATGAAGCTGTTTGACAGCATCTGCAACAACAAGTGGTTC SphI
(301) 901	TyrThrThrProGluAspAlaThrProGluProGlyGluAspProArgValThrArgAla TACACCACTCCTGAGGATGCGACTCCCGAGCCCGGGAGGACCCACCGGGGGCC	(261) 781	ThrAspThrSerIleIleLeuPheLeuAsnLysLysAspLeuPheGluGluLysIleThr ACAGACACCTCCATCATCCTCTCTCCTCAACAAGAAGGACCTGTTTGAAGAAAAGATCACA
(321) 961	LysTyrPheIle&rg&spGluPheLeu&rgIleSerThrAlaSerGlyAspGlyArgHis AAgtacttCattcgggatgagtttctg&gaatcagcactgctagtgg&gatgggccc&& Beri	(281) 841	GinSerSerLeuThrIleCysPheProGluTyrThrGlyAlaAsnLysTyrAspGluAla CAGAGCTCCCTGACCATCTGTTTCCCTGAGTACACGGGGGCCAACAAGTACGACGAGGCA Sact
(341) 1021	TyrCysTyrProHisPheThrCysAla¥alAspThrGluAsnIleArgArg¥alPheAsn TACTGCTACCTCACTTTACCTGCGCCGTGGACACTGAGAACATCCGCCGTGTCTTCAAC	(301) 901	AlaSerTyrIleGlnSerLysPheGluAspLeuAsnLysArgLysAspThrLysGluIle GCCAGCTACATCCAGAGCAAGTTTGAGGATCTAAATAAGCGCAAAGACACCACGAGGGAG Bell1
(361) 1081	AspCysArgAspIleIleGlnArgMetHisLeuArgGlnTyrGluLeuLeu GACTGCCGTGACATCATCCAGCGCATGCATCTCCGCCAATACGAGCTGCTCTAAGAAGGG	(321) 961	TyrThrHisPheThrCysAlaThrAspThrLysAsnValGlnPheValPheAspAlaVal TACACGCACTTCACATGCGCCACCGACACCAAGAACGTGCACTTGTGTTCCATGCCGTC
1141	AACACCCAAATTTAATTCAGCCTTAÅGCACAATTAATTAAGAGTGAAACGTAATTGTACA	(341)	ThrAspValIleIleĹysAsnAsnLeuLysAspCysGlyLeuPhe
1201	AGCAGTTGGTCACCCACCATAGGGCATGATCAACACCGCAACCTTTCCTTTTTCCCCCAG	1021	ACTGACGTCATCATCAAGAACAACCTGAAGGACTGTGGCCTCTTCTGAGGGGCAGTGGAC
1261	TGATTCTGAAAAACCCCTCTTCCCTTCACACCTTGCTTAGATTTCCAAATTTAGTAAGCT BindIII	1081	CTGGCAGGATTGCGGAGCCTCTGCCTACCCACTCTCTGCCGCTCGGAGGCCCCAAAGGAA
1321	TAAGGGCCCTGCAGAAGAAAAAAGAAAAAAAAGTTCCCTCTCACTTTCAGTA <u>AAAAAAA</u> ATT Psti	1141	AAAGCACAAGAAGCGTGAGAGATACGCCATTCTGGAGACAAAGCCACCTGCTATTCTCGT
1381	AAAAGCAGCAACAGA <u>AATAAA</u> GA <u>AATAAA</u> TGAAATTCAAAATGAAATAAAATATTGTGTTG	1201	AGCTTTANAGAAAAAAAAAAAAAAAAA
1441	TGCAGCATTAAAAAATC <u>AATAAA</u> AATTAAAAATGAGCAAAAAAAAAAAAAAAAA		

FIG. 2. Nucleotide and predicted amino acid sequence of mouse α_s (*Left*) and α_i (*Right*). Numbers represent the positions of nucleotides or amino acids. Specific restriction endonuclease cleavage sites are indicated below the nucleotide sequence. Poly(A) signals corresponding to the consensus sequence AAUAAA are underlined.

as well as two additional GTP-binding proteins, bacterial EF-Tu (33) and p21^{c-Ha-ras} (34). The striking overall homology among the four G protein α chains indicates their evolution as a family from a common precursor. Of the four α chains, α_{t1} and α_{t2} are most closely related (81% identical residues); α_i is more closely related to α_{t1} (63% identical residues) than is α_s (38%), which is the most distinctive of the four (Fig. 3). Detailed comparison of the sequences also provides insight into the structural basis for functional differences and similarities among different G protein α chains.

The almost identical carboxyl-terminal regions of α_i , α_{t1} , and α_{t2} contain the conserved cysteine residues (Cys-347 in α_{t1}) that is ADP-ribosylated by pertussis toxin (30). α_s , unlike α_i and α_t , is not a substrate for pertussis toxin. As expected, the carboxyl terminus of α_s differs strikingly from that of the other three α chains (Fig. 3); a tyrosine residue in α_s replaces the cysteine modified by pertussis toxin in α_t . In contrast, cholera toxin ADP-ribosylates α_s and α_t , a modification that reduces the rate of GTP hydrolysis by both polypeptides (35, 36). As we would have predicted, the sequence surrounding the amino acid ADP-ribosylated by cholera toxin in α_{t1} , Arg-174 (37), closely resembles that surrounding the presumptive cholera toxin ADP-ribosvlation site of α_s , Arg-182 (Fig. 3). Surprisingly, because G_i is not usually considered a cholera toxin substrate, the corresponding region of α_i also closely resembles the cholera toxin modification site of α_s and contains a conserved arginine residue at position 176. This observation may explain recent reports (38, 39) that cholera toxin can, under certain conditions, ADP-ribosylate α_i and alter its function.

As portions of the presumptive site for binding and hydrolysis of GTP in α_s , α_i , and α_t , we postulate four regions whose sequences are nearly identical in the three polypeptides. These regions are designated A, C, E, and G in Fig. 3, in accord with four regions shown (40) to exhibit the greatest homology among two other groups of GTP binding proteins, the ras polypeptides and bacterial elongation factors. Homologies among the α chains, EF-Tu, and the ras proteins are strongest in regions A and G and relatively weak in regions C and E. The three-dimensional crystal structure of the GDP-binding domain of EF-Tu (41) has revealed that each of these four regions is situated at a turn between a β strand and an α helix. Specific residues in or near these turns interact with the bound GDP molecule or with an associated Mg²⁺ ion. The same residues are conserved in the four G protein α chains

In EF-Tu, region A coils close to the α and β phosphates of GDP (41). Gly-12 in the A region of p21^{c-Ha-ras}, which corresponds to Gly-47 in α_s (Fig. 3), is critical for GTP hydrolysis; its replacement by other amino acids produces a p21 molecule that hydrolyzes GTP more slowly than wild type and that has an increased capacity for oncogenic transformation (42).

In the C region of EF-Tu, Asp-80 binds to the Mg²⁺ associated with the β phosphate of the bound GDP molecule



FIG. 3. Comparison of amino acid sequences of mouse α_s and α_i with bovine α_{t1} and α_{t2} and the presumptive GTP-binding regions of elongation factor Tu (EF-Tu) and c-Ha-*ras1*. Gaps, indicated by hyphens, were introduced to obtain maximum homology. The A, C, E, and G regions corresponding to the presumptive guanine nucleotide-binding site are indicated above the sequences. (c) and (p) indicate positions of the arginine and cysteine residues ADP-ribosylated by cholera and pertussis toxins, respectively. Boxes surround residues that are identical in any two of the four G protein α chains and residues in EF-Tu and c-Ha-*ras* that are identical or conserved (29) with respect to amino acids at corresponding positions in any one of the four α chains. Single-letter abbreviations for amino acid residues are as follows: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine. The following Dayhoff conservative categories (29) were used: C; S, T, P, A, G; N, D, E, Q; H, R, K; M, I, L, V; F, Y, W.

(41). This region (surrounding Asp-206 in α_s) is highly conserved among all of the G protein α chains (Fig. 3). It is situated downstream from the arginine (Arg-182 of α_s) that is ADPribosylated by cholera toxin, a modification that decreases GTPase activity; moreover, binding of hydrolysis-resistant GTP analogs specifically prevents tryptic cleavage at a closely adjacent residue of α_{t1} (Arg-203, corresponding to Arg-214 of α_s and Arg-209 of α_i) (18). The E region (Fig. 3) is a stretch of hydrophobic residues that are well conserved among the α chains; similarly, hydrophobic residues in the corresponding region of EF-Tu form part of a hydrophobic pocket for the guanine ring of GDP. Region G in EF-Tu comprises a β -to- α turn located close to the guanine moiety of GDP (41); in EF-Tu, Asn-135 is located directly over the plane of the guanine ring and Asp-138 interacts with the amino substituent on the guanine; identical residues are found at the corresponding positions (275 and 278 of α_s) of all four α chains. The G region shows extremely high homology within the family of G protein α chains and among these chains as well as the ras proteins, EF-Tu, and elongation factor G.

Amino acid sequences in certain regions outside the presumptive GTP-binding site are also conserved, whereas other regions vary greatly. Thus, amino acids corresponding to residues 133-204 and 344-377 in α_s are 43% and 42% identical, respectively, in α_s , α_i , and α_{t1} ; most of the nonidentical residues are conservative substitutions. These conserved regions, in which polypeptide folding is probably also conserved, may participate in binding to the $\beta\gamma$ complex, which is common to all G protein-mediated transduction systems. In contrast, two other regions of the α chains (corresponding to residues 61–132 and 287–343 in α_s) are quite different, showing only 11% and 4% identity, respectively, among α_s , α_i , and α_{t1} . Although the poorly conserved sequences may be functionally unimportant, they may instead form sites for specific interaction with the structurally distinct receptor and effector elements of different signal transduction systems.

In the α chains a long stretch of sequence is interposed between GTP-binding regions A and C. In EF-Tu, the much shorter stretch of interposed amino acids (49 residues vs. 139–145 residues in the α chains) appears to participate in binding aminoacyl tRNA (33); the affinity of this interaction is much higher when GTP, rather than GDP, is bound to EF-Tu. By analogy, binding functions of the corresponding interposed peptide sequences of the α chains may also be subject to a GTP-triggered conformational switch. The first portion of this sequence in the α chains, corresponding to residues 61–132 of α_s , is highly variable (11% identical amino acids in α_s , α_i , and α_{t1}). This region may form part of the α chain domains that activate adenylate cyclase, cGMP phosphodiesterase, and other effector molecules in a GTPdependent fashion. The second half of the interposed sequence in the α chains, corresponding to residues 133–204 in α_s , is highly conserved. It would not be surprising if this conserved region turns out to form part of the α chains' interface with the $\beta\gamma$ complex, because this binding interaction is also subject to a GTP-regulated conformational switch (1). Biochemical evidence (43, 44) indicates that the aminoterminal 18 residues of α_{t1} are also required for binding to $\beta\gamma$.

Finally, three kinds of evidence suggest that the carboxylterminal 25 residues of the α chains form at least part of their receptor-binding domain. (i) Pertussis toxin-catalyzed attachment of ADP-ribose to a conserved cysteine four residues from the carboxyl terminus prevents transducin 1 and G_i from interacting with photorhodopsin (45) and inhibitory receptors (46), respectively; this effect might result from steric hindrance by the ADP-ribose moiety. (ii) The carboxylterminal 25 residues of α_{t1} exhibit impressive homology to a sequence (47) in a second retinal protein, arrestin, which also binds specifically to photorhodopsin and prevents binding and activation of transducin (4). (iii) The abilities of the G proteins to be stimulated by different receptor elements correlate closely with patterns of sequence homology in this 25-residue carboxyl-terminal region of the α chains. Thus, photorhodopsin briskly stimulates GTP hydrolysis by transducin 1 and G_i (48), whose α chains exhibit 88% identical residues in the carboxyl-terminal region (Fig. 3); photorhodopsin is a very poor stimulator of GTP hydrolysis by α_s (49), which is identical to either α_i or α_{t1} at only 28–32% of the carboxyl-terminal 25 residues. Conversely, a different detector element, the β -adrenoceptor, stimulates GTPase activity of G_s much more effectively than it does that of either G_i or transducin 1 (49).

Further experiments will be required to test and refine these tentative assignments of α -chain regions to specific functions in GTP-dependent transmembrane signaling.

We thank Daniel Chin for helpful advice and A. G. Gilman for telling us, before publication, that α_s message is absent in cyc⁻ cells. This work was supported in part by a grant from the March of Dimes and by National Institutes of Health Grants GM27800, GM28310, and HL07192.

- 1. Gilman, A. G. (1984) Cell 36, 577-579.
- Sternweis, P. C. & Robishaw, J. D. (1984) J. Biol. Chem. 259, 13806–13813.
- Neer, E. J., Lok, J. M. & Wolf, L. G. (1984) J. Biol. Chem. 259, 14222-14229.
- 4. Stryer, L. (1986) Annu. Rev. Neurosci. 9, 87-119.
- Lochrie, M. A., Hurley, J. B. & Simon, M. I. (1985) Science 228, 96–99.
- Berridge, M. J. & Irvine, R. F. (1984) Nature (London) 312, 315-321.
- Pfaffinger, P. J., Martin, J. M., Hunter, D. D., Nathanson, N. M. & Hille, B. (1985) Nature (London) 317, 536-538.
- Breitwieser, G. E. & Szabo, G. (1985) Nature (London) 317, 538-540.
- Holz, G. G., Rane, S. G. & Dunlap, K. (1986) Nature (London) 319, 670-672.
- Medynski, D. C., Sullivan, K., Smith, D., Van Dop, C., Chang, F. H., Fung, B. K. K., Seeburg, P. H. & Bourne, H. R. (1985) Proc. Natl. Acad. Sci. USA 82, 4311-4315.
- Pennica, D., Hayflick, J. S., Bringman, T. S., Palladino, M. A. & Goeddel, D. V. (1985) Proc. Natl. Acad. Sci. USA 82, 6060-6064.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- 13. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
- Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
- 15. Thomas, P. S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.

- 16. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
- 17. Messing, J., Crea, R. & Seeburg, P. (1981) Nucleic Acids Res. 9, 309-321.
- Hurley, J. B., Simon, M. I., Teplow, D. B., Robishaw, J. D. & Gilman, A. G. (1984) Science 226, 860-862.
- Harris, B. A., Robishaw, J. D., Mumby, S. M. & Gilman, A. G. (1985) Science 229, 1274–1277.
- 20. Ross, E. M. & Gilman, A. G. (1977) J. Biol. Chem. 252, 6966-6969.
- Bourne, H. R., Beiderman, B., Steinberg, F. & Brothers, V. M. (1982) Mol. Pharmacol. 22, 204-210.
- Hildebrandt, J. D., Sekura, R. D., Codina, J., Iyengar, R., Manclark, C. R. & Birnbaumer, L. (1983) Nature (London) 302, 706-709.
- Aktories, K., Schultz, G. & Jakobs, K. H. (1983) FEBS Lett. 158, 169–173.
- Northup, J. K., Sternweis, P. C., Smigel, M. D., Schleifer, L. S., Ross, E. M. & Gilman, A. G. (1980) Proc. Natl. Acad. Sci. USA 77, 6516-6520.
- 25. Kozak, M. (1984) Nucleic Acids Res. 12, 857-873.
- Robishaw, J. D., Russell, D. W., Harris, B. A., Smigel, M. D. & Gilman, A. G. (1986) Proc. Natl. Acad. Sci. USA 83, 1251–1255.
- 27. Bokoch, G. M., Katada, T., Northup, J. K., Hewlett, E. L. & Gilman, A. G. (1983) J. Biol. Chem. 258, 2072–2075.
- Codina, J., Hildebrandt, J., Iyengar, R., Birnbaumer, L., Sekura, R. D. & Manclark, C. R. (1983) Proc. Natl. Acad. Sci. USA 80, 4276-4280.
- 29. Dayhoff, M. O., ed. (1978) Atlas of Protein Sequence and Structure (National Biomedical Research Foundation, Washington, DC), pp. 345-352.
- 30. West, R. E., Moss, J., Vaughan, M., Liu, T. & Liu, T. Y. (1985) J. Biol. Chem. 260, 14428-14430.
- Tanabe, T., Nukada, T., Nishikawa, Y., Sugimoto, K., Suzuki, H., Takahashi, H., Noda, M., Haga, T., Ichiyama, A., Kangawa, K., Minamino, N., Matsuo, H. & Numa, S. (1985) Nature (London) 315, 242-245.
- 32. Yatsunami, K. & Khorana, G. (1985) Proc. Natl. Acad. Sci. USA 82, 4316-4320.
- Laursen, R. A., L'Italien, J. J., Nagarkatti, S. & Miller, D. L. (1981) J. Biol. Chem. 256, 8102–8109.
- Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. (1983) Nature (London) 302, 33-37.
- 35. Cassel, D. & Selinger, Z. (1977) Proc. Natl. Acad. Sci. USA 74, 3307-3311.
- Abood, M. E., Hurley, J. B., Pappone, M.-C., Bourne, H. R. & Stryer, L. (1982) J. Biol. Chem. 257, 10540-10543.
- Van Dop, C., Tsubokawa, M., Bourne, H. R. & Ramachandran, J. (1984) J. Biol. Chem. 259, 696–698.
- Owens, J. R., Frame, L. T., Ui, M. & Cooper, D. M. F. (1985) J. Biol. Chem. 260, 15946–15952.
- Aksamit, R. R., Backlund, P. S., Jr., & Cantoni, G. L. (1985) Proc. Natl. Acad. Sci. USA 82, 7475-7479.
- 40. Halliday, K. (1984) J. Cyclic Nucleotide Res. 9, 435-448.
- 41. Jurnak, F. (1985) Science 230, 32-36.
- McGrath, J. P., Capon, D. J., Goeddel, D. V. & Levinson, A. D. (1984) Nature (London) 310, 644–649.
- 43. Fung, B. K.-K. (1983) J. Biol. Chem. 258, 10495-10502.
- 44. Fung, B. K.-K. & Nash, C. R. (1983) J. Biol. Chem. 258, 10503-10510.
- Van Dop, C., Yamanaka, G., Steinberg, F., Sekura, R. D., Manclark, C. R., Stryer, L. & Bourne, H. R. (1984) J. Biol. Chem: 259, 23-26.
- 46. Ui, M., Katada, T., Murayama, T., Kurose, H., Yajima, M., Tamura, M., Nakamura, T. & Nogimori, K. (1984) Adv. Cyclic Nucleotide Res. 17, 145–151.
- 47. Wistow, G. J., Katial, A., Craft, C. & Shinohara, T. (1986) FEBS Lett. 296, 23-28.
- Kanaho, Y., Tsai, S. C., Adamik, R., Hewlett, E. L., Moss, J. & Vaughan, M. (1984) J. Biol. Chem. 259, 7378-7381.
- Cerione, R. A., Staniszewski, C., Benovic, J. L., Lefkowitz, R. J., Caron, M. G., Gierschik, P., Somers, R., Spiegel, A. M., Codina, J. & Birnbaumer, L. (1985) J. Biol. Chem. 260, 1493-1500.
- 50. Robishaw, J. D., Smigel, M. D. & Gilman, A. G. (1986) J. Biol. Chem., in press.