Adenylyl Cyclase Amino Acid Sequence: Possible Channel- or Transporter-Like Structure

JOHN KRUPINSKI, FRANÇOISE COUSSEN,* HEATHER A. BAKALYAR, WEI-JEN TANG, PAUL G. FEINSTEIN, KIM ORTH, CLIVE SLAUGHTER, RANDALL R. REED, ALFRED G. GILMAN

Complementary DNA's that encode an adenylyl cyclase were isolated from a bovine brain library. Most of the deduced amino acid sequence of 1134 residues is divisible into two alternating sets of hydrophobic and hydrophilic domains. Each of the two large hydrophobic domains appears to contain six transmembrane spans. Each of the two large hydrophilic domains contains a sequence that is homologous to a single cytoplasmic domain of several guanylyl cyclases; these sequences may represent nucleotide binding sites. An unexpected topographical resemblance between adenylyl cyclase and various plasma membrane channels and transporters was observed. This structural complexity suggests possible, unappreciated functions for this important enzyme.

SINCE THE DISCOVERY OF ADENOSINE 3′,5′-MONOPHOSphate (cyclic AMP) by Rall *et al.* (*1*), intensive research efforts have focused on mechanisms of regulation of the intracellular concentration of this second messenger. Dozens of hormones and neurotransmitters can either stimulate or inhibit the rate of cyclic AMP synthesis by pathways that consist of three distinct types of plasma membrane–associated proteins: hormone receptors, guanine nucleotide-binding regulatory proteins (G proteins), and adenylyl cyclase itself. For example, binding of hormone to a stimulatory receptor at the extracellular surface triggers the activation of the signal transducing G protein, G_s. This species then interacts with adenylyl cyclase at the intracellular face of the plasma membrane to stimulate the formation of cyclic AMP. The primary structures of various G proteins (*2*), as well as those of several relevant receptors (*3*), have been determined. We now describe the primary structure of a hormone-sensitive adenylyl cyclase.

Both genetic (*4*) and biochemical (*5–8*) evidence indicates that there are multiple forms of adenylyl cyclase. The estimated molecular masses of these forms falls in the range of 120 to 150 kD, depending on the source (*7*) or the electrophoretic system used (*6*). The adenylyl cyclase most readily purified from a eukaryote is a glycoprotein with a molecular mass of approximately 120 kD; this enzyme has been purified to apparent homogeneity from bovine brain (*5, 6*). The fundamental breakthrough in the realization of this goal was the discovery that an activator of adenylyl cyclase, the diterpene forskolin, could be covalently coupled to an agarose matrix to permit the selective adsorption and elution of the protein (*9*). Using the eluate from a forskolin column as antigen, Mollner and Pfeuffer (*7*) prepared monoclonal antibodies to bovine brain adenylyl cyclase. One of these antibodies specifically recognized the enzyme in immunoblots of solubilized bovine brain membranes and precipitated adenylyl cyclase activity from detergent solutions that contained the partially purified protein (*7*). The eluate from a forskolin-Sepharose affinity column can be purified further by chromatography on supports that contain immobilized wheat germ agglutinin (*5, 6*) or calmodulin (*8*). A 120-kD component of such preparations binds calmodulin and wheat germ agglutinin in a gel overlay system after sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) (*8*). In addition, hormone-dependent synthesis of cyclic AMP can be demonstrated when this highly purified preparation is reconstituted with homogeneous β-adrenergic receptors and G_s in phospholipid vesicles (*10, 11*). Thus, the 120-kD protein purified from bovine brain is in fact a catalytic moiety of a hormone-regulated adenylyl cyclase system.

Protein sequencing and cDNA cloning. A silver-stained SDS-polyacrylamide gel of approximately 400 ng of adenylyl cyclase, partially purified by affinity chromatography on forskolin-Sepharose (*12*), shows a diffuse, but prominent band just above the 116-kD marker (Fig. 1A). We refer to this protein as 120-kD adenylyl cyclase. We and others have purified this species further by column chromatography with calmodulin agarose (*8, 13*) or wheat germ lectin agarose (*5, 6, 14*).

Attempts to sequence the intact protein were unsuccessful because of a blocked amino terminus; the nature of this block remains unknown. The protein was thus digested with trypsin after transfer to a nitrocellulose filter, and peptides were separated by sequential cycles of high-performance liquid chromatography (HPLC) in acetonitrile gradients, first at *p*H 2.1 and then at *p*H 6.5 (Fig. 1, B and C) (*15, 16*). Fourteen different peptide sequences were ultimately obtained; all of them (underlined in Fig. 3) were subsequently identified in the long open reading frame encoded by the cDNA clones described below.

One of the peptide sequences (peak A in Fig. 1C) was sufficiently hydrophilic to merit an attempt to generate a polyclonal peptide

J. Krupinski, F. Coussen, W.-J. Tang, and A. G. Gilman are in the Department of Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235. H. A. Bakalyar, P. G. Feinstein, and R. R. Reed are in the Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205. K. Orth and C. Slaughter are in the Howard Hughes Medical Institute and Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75235.

*Present address: Laboratoire de Neurobiologie, Ecole Normale Superieure, 46 Rue d'Ulm, 75005 Paris, France.

antibody. The synthetic peptide was coupled to keyhole limpet hemocyanin to serve as antigen (*17*). An immunoblot of forskolin-Sepharose purified adenylyl cyclase (Fig. 1D) shows a diffuse band above the 116-kD marker, indicating that this peptide sequence was in fact obtained from the 120-kD protein; however, the antibody cannot detect the low amounts of this protein in bovine brain membranes.

The peak labeled "II" in Fig. 1B contained three different peptides. This was the one instance in which an unambiguous sequence could be assigned to one of the peptides on the basis of yields from the protein sequencer. We used the sequence of this most prominent peptide to synthesize a single 50-nucleotide antisense probe, following the criteria of Lathe (*18*).

A size-fractionated, randomly primed bovine brain cDNA library in bacteriophage lambda zap was screened; approximately 2 million plaques yielded a series of 36 independent, overlapping, clones (*19*) (Fig. 2). Clone p1-6, which extends from nucleotide 736 to 3191 in the cDNA sequence of Fig. 3, was used to probe bovine brain poly(A)$^+$ RNA at high stringency. The result was surprising. A single message of approximately 11.5 kb was visualized (Fig. 4); it contained an exceptionally long 3′ untranslated region, since another clone that also appeared to hybridize to the same message, p1-8 (Fig. 2), extended for approximately 3.4 kb beyond the end of the open reading frame, but it did not contain a poly(A)$^+$ tail.

The 5′ end of clone pA11 is G-C rich; of the first 200 nucleotides, 91 percent are either G or C (Fig. 3). It is possible that the 5′ end of this clone represents the true 5′ end of the message. However, it is also possible that secondary structure in this G-C rich region prevented the synthesis of longer clones during construction of the cDNA library. Two ATG sequences are nested in the middle of the G-C rich region. The first is in poor agreement with consensus initiation sequences (*20*) and is not in frame with any of the peptide sequences determined by protein sequencing; the associated reading frame encodes only 67 amino acid residues. The second ATG is part of a reasonable consensus initiation sequence and is in frame with all of the tryptic peptide sequences. This second ATG may be the initiator methionine for adenylyl cyclase, even though it is not preceded by upstream termination codons. The molecular mass of the protein deduced by translation of the open reading frame is 124 kD, in excellent agreement with that estimated for the purified protein by SDS-PAGE. Although it is certainly possible that the sequence is incomplete at the 5′ end, we do observe elevated adenylyl cyclase activity in COS-m6 cells after they are transfected with a putative full-length cDNA assembled from three of the clones. This activity is dependent on the amount of plasmid utilized to transfect the cells and the concentration of forskolin in the assay (Fig. 5).

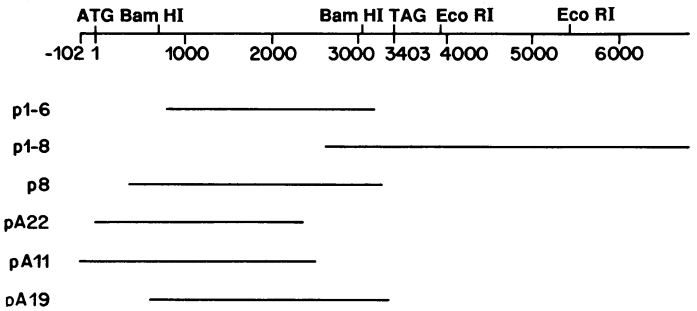

Fig. 2. cDNA clones for bovine brain adenylyl cyclase. Map of several clones arranged relative to the regions of the coding sequence that they specify.

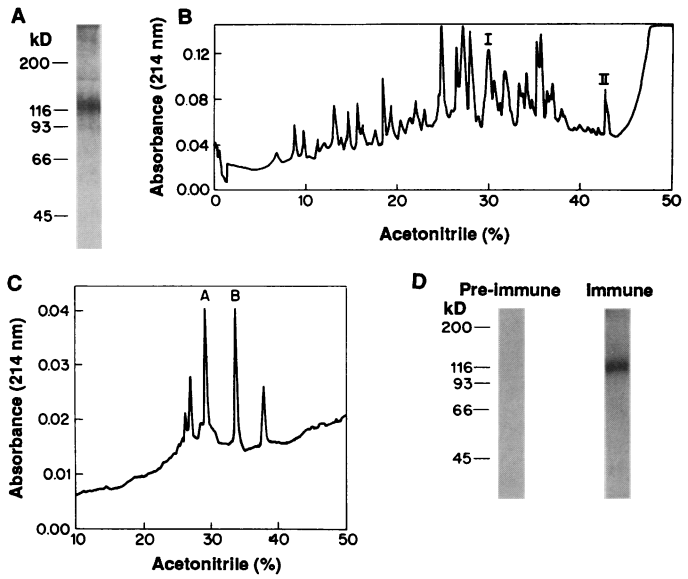

Fig. 1. Purified bovine brain adenylyl cyclase and peptide sequence determination. (**A**) Silver-stained SDS-polyacrylamide gel of approximately 400 ng of adenylyl cyclase partially purified by affinity chromatography on forskolin-Sepharose (*12*). (**B**) Absorbance trace of an HPLC separation of peptides formed from the solid-phase digestion of approximately 300 pmol of 120-kD adenylyl cyclase with 30 pmol of trypsin (*16*). The gradient used for separation of the peptides was 0 to 50 percent (v/v) acetonitrile with 0.1 percent trifluoroacetic acid, *p*H 2.1 (*15*). (**C**) The peak labeled I in (B) was collected, processed, and then chromatographed again in a 10 to 50 percent (v/v) acetonitrile gradient with 0.1 percent ammonium acetate, *p*H 6.5 (*15*). Automated Edman degradation was used to analyze the material in peaks A and B; each gave a distinct peptide sequence, namely, GKGEMLTYFLEGR (*36*) and VGINVGPVVAGVIGAR, respectively. (**D**) A peptide with the sequence CKGEMLTYFLEGR was synthesized, coupled to keyhole limpet hemocyanin, and used as an antigen in rabbits (*17*). Approximately 400 ng of bovine brain adenylyl cyclase purified on a forskolin-Sepharose column was subjected to electrophoresis and transferred to nitrocellulose. A dilution of rabbit antiserum (1 to 100) was used to probe the nitrocellulose.

These data support the possibility that the cDNA's isolated do, in fact, encode an adenylyl cyclase. In previous genetic studies of *Drosophila melanogaster*, a locus (rutabaga) that is thought to encode the structural gene for a calmodulin-dependent form of adenylyl cyclase was identified (*4*). Mutations at this site can be correlated with the loss of calmodulin-stimulated adenylyl cyclase activity in this organism. One of us (*21*) has recently isolated a genomic clone from *Drosophila* by cross-hybridization to the bovine brain cDNA (p1-6); the sequences of these clones are homologous. In situ hybridization of the genomic clone to *Drosophila* polytene chromosomes reveals hybridization within the region 12E1-13A5 on the X chromosome that defines the rutabaga locus. This fact, in combination with the evidence presented above and the sequence homologies to be discussed below, remove all reasonable doubt about the identity of the cDNA's described in this article.

Analysis of the protein sequence. There are two domains within the protein sequence that are similar to one another. A dot matrix comparison (Fig. 6A) reveals a cluster of four line segments along one diagonal, indicating that amino acids 238 to 480 are similar to residues 809 to 1059. When these two regions are aligned (*22, 23*) and analyzed, we see that these two domains of adenylyl cyclase are 28 percent identical and 54 percent similar (*24*).

The sequences of four different guanylyl cyclases are now available (*25–29*). Each of the two homologous domains within adenylyl cyclase is similar to a cytoplasmic domain within each of the guanylyl cyclases (Fig. 6B). Three of the guanylyl cyclases contain domains that are similar to most of the duplicated regions within adenylyl cyclase. These three include two membrane-bound proteins—those from rat brain (*25*) and the sea urchin *Strongylocentrotus*

```
-102          GCCGGCGCCCCAACTCGGCCCGCCGCGCCCCGGCGCCTCGCCGCCCGCCCGCCCGCCCGGCGCCCCGGCCGGCGAGGGGCGCGCCCGCTGCATGGCGCTGGG
1     ATGGCGGGGGCGCCGCGCGGCCGAGGCGGCGGCGGAGGCGGAGGCGGCGCGGGCGAGTCTGGGGGCGCCGAGCGGGCGGCGGGGCCGGGCGGCCGGCGCGGGCTGCGGGCGTGCGATGAG
1     M A G A P R G R G G G G G G G G G A G E S G G A E R A A G P G G R R G L R A C D E
121   GAGTTCGCGTGCCCCGAGCTGGAGGCGCTGTTCCGCGGCTACACGCTGCGGCTGGAGCAGGCGGCGACGCTGAAGGCGCTGGCCGTGCTCAGCCTGCTGGCGGGCGCGCTGGCCCTGGCC
41    E F A C P E L E A L F R G Y T L R L E Q A A T L K A L A V L S L L A G A L A L A
241   GAGCTGCTGGGCGCGCCGGGGCCCGCGCCCGGCCTGGCTAAGGGCTCGCACCCCGTGCACTGCGTGCTCTTCCTGGCGCTGCTCGTGGTCACCAACGTCCGCTCGCTGCAGGTGCCCCAG
81    E L L G A P G P A P G L A K G S H P V H C V L F L A L L V V T N V R S L Q V P Q
361   CTGCAGCAGGTCGGCCAGCTCGCGCTGCTCTTCAGCCTCACCTTCGCTCTGCTGTGCTGTCCCTTCGCGCTCGGCGGCCCCGCCGGTGCCCACGCCGGGGCGGCAGCGGTGCCGGCGACA
121   L Q Q V G Q L A L L F S L T F A L L C C P F A L G G P A G A H A G A A A V P A T
481   GCCGATCAGGGAGTCTGGCAGCTCCTTTTGGTCACCTTCGTGTCCTATGCCCTGCTGCCCGTGCGCAGCCTGCTGGCCATCGGCTTCGGGCTCGTGGTGGCCGCCTCGCACTTGCTGGTC
161   A D Q G V W Q L L L V T F V S Y A L L P V R S L L A I G F G L V V A A S H L L V
601   ACGGCTACGTTGGTCCCCGCCAAGCGCCCACGTCTCTGGAGAACGCTGGGTGCCAACGCTCTGCTCTTCCTCGGTGTGAACGTGTATGGCATCTTCGTGAGGATCCTGGCTGAGCGCGCC
201   T A T L V P A K R P R L W R T L G A N A L L F L G V N V Y G I F V R I L A E R A
721   CAGAGGAAGGCCTTCCTGCAGGCCCGGAACTGCATTGAGGACCGGCTGAGGCTGGAGGATGAGAATGAGAAGCAGGAGCGGCTGCTCATGAGCCTCCTGCCTCGGAATGTTGCCATGGAG
241   Q R K A F L Q A R N C I E D R L R L E D E N E K Q E R L L M S L L P R N V A M E
841   ATGAAGGAGGACTTCCTGAAGCCCCCTGAGAGGATTTTCCACAAGATTTACATCCAGCGGCATGACAACGTGAGCATCCTCTTTGCAGACATCGTGGGCTTCACAGGCTTGGCGTCACAG
281   M K E D F L K P P E R I F H K I Y I Q R H D N V S I L F A D I V G F T G L A S Q
961   TGCACGGCCCAGGAGCTGGTGAAACTCCTCAATGAGCTCTTCGGGAAGTTTGACGAGCTGGCCACAGAGAACCACTGCCGCCGCATCAAGATCCTGGGAGATTGCTACTACTGCGTGTCT
321   C T A Q E L V K L L N E L F G K F D E L A T E N H C R R I K I L G D C Y Y C V S
1081  GGCCTCACTCAGCCCAAGACTGACCACGCCCACTGCTGTGTGGAGATGGGCCTGGACATGATCGACACCATCACGTCCGTGGCTGAGGCCACTGAGGTGGACTTGAACATGCGTGTGGGG
361   G L T Q P K T D H A H C C V E M G L D M I D T I T S V A E A T E V D L N M R V G
1201  CTGCACACCGGCAGGGTCCTCTGCGGGGTCCTGGGCCTGCGTAAGTGGCAGTATGATGTGTGGTCCAACGACGTGACCCTGGCCAACGTCATGGAGGCTGCCGGCCTGCCTGGGAAGGTT
401   L H T G R V L C G V L G L R K W Q Y D V W S N D V T L A N V M E A A G L P G K V
1321  CACATCACAAAGACCACCCTGGCGTGCCTGAATGGTGACTATGAGGTGGAGCCGGGACACGGACACGAGAGGAACAGTTTTCTGAAAACTCATAACATTGAGACCTTTTTTATTGTGCCC
441   H I T K T T L A C L N G D Y E V E P G H G H E R N S F L K T H N I E T F F I V P
1441  TCGCATCGGCGAAAGATATTTCCAGGGCTGATTCTCTCAGACATAAAACCGGCCAAGAGGATGAAGTTCAAGACCGTGTGCTACCTGCTGGTGCAGCTCATGCACTGCCGGAAGATGTTC
481   S H R R K I F P G L I L S D I K P A K R M K F K T V C Y L L V Q L M H C R K M F
1561  AAGGCCGAGATCCCTTTCTCCAACGTCATGACCTGTGAGGATGACGACAAGCGGAGGGCACTGAGAACAGCCTCGGAAAAACTCAGAAACCGCTCGTCTTTCTCTACAAACGTTGTCCAA
521   K A E I P F S N V M T C E D D D K R R A L R T A S E K L R N R S S F S T N V V Q
1681  ACCACCCCCGGCACACGTGTCAACAGGTACATCGGCCGCCTCCTGGAAGCCCGCCAGATGGAGCTGGAGATGGCAGACCTGAACTTCTTCACCCTGAAGTACAAGCAAGCTGAGCGAGAG
561   T T P G T R V N R Y I G R L L E A R Q M E L E M A D L N F F T L K Y K Q A E R E
1801  CGAAAGTACCACCAGCTTCAGGACGAGTATTTCACCAGCGCCGTGGTTCTGGCTCTCATTCTGGCCGCCTTATTCGGCCTTGTCTACCTTCTAATAATCCCACAGAGTGTGGCTGTCCTG
601   R K Y H Q L Q D E Y F T S A V V L A L I L A A L F G L V Y L L I I P Q S V A V L
1921  CTCCTGCTGGTGTTCTGCATCTGCTTCCTGGTGGCCTGTGTCCTGTACCTACACATCACCCGGGTCCAGTGTTTTCCAGGGTGCCTGACCATCCAGATCCGCACCGTCTTGTGCATCTTC
641   L L L V F C I C F L V A C V L Y L H I T R V Q C F P G C L T I Q I R T V L C I F
2041  ATCGTGGTCTTAATCTACTCTGTGGCCCAAGGCTGTGTGGTGGGCTGCCTGCCTTGGTCCTGGAGCTCCAGTCCCAACGGGTCCCTGGTGGTCCTGTCTTCTGGGGGCCGGGACCCAGTG
681   I V V L I Y S V A Q G C V V G C L P W S W S S S P N G S L V V L S S G G R D P V
2161  CTGCCTGTCCCGCCCTGCGAGTCTGCGCCCCATGCCCTGCTGTGCGGCCTCGTGGGCACCCTCCCGCTGGCCATATTCCTGCGGGTCTCCTCCTTGCCAAAAATGATCCTGCTCGCCGTG
721   L P V P P C E S A P H A L L C G L V G T L P L A I F L R V S S L P K M I L L A V
2281  CTCACCACCTCCTACATCCTCGTCCTGGAGCTCAGCGGGTACACGAAGGCCATGGGGGCCGGTGCCATCTCAGGGCGCAGCTTCGAGCCGATCATGGCCATCCTGCTATTCTCGTGCACG
761   L T T S Y I L V L E L S G Y T K A M G A G A I S G R S F E P I M A I L L F S C T
2401  CTGGCCCTGCACGCCCGGCAGGTGGATGTCAAGCTGCGGCTGGACTACCTCTGGGCGGCCCAGGCAGAGGAGGAGCGGGATGACATGGAGAAAGTGAAGCTGGACAACAAGAGGATTCTC
801   L A L H A R Q V D V K L R L D Y L W A A Q A E E E R D D M E K V K L D N K R I L
2521  TTCAACCTCCTGCCAGCCCACGTTGCCCAGCACTTCCTAATGTCCAACCCTCGCAACATGGACCTGTATTACCAGTCATACTCGCAGGTGGGGGTCATGTTTGCATCCATCCCCAACTTC
841   F N L L P A H V A Q H F L M S N P R N M D L Y Y Q S Y S Q V G V M F A S I P N F
2641  AATGACTTCTACATCGAGCTGGATGGCAACAACATGGGGGTGGAATGTCTACGCCTTCTGAATGAGATCATCGCTGACTTTGATGAGCTCATGGACAAAGACTTTTACAAGGACCTAGAG
881   N D F Y I E L D G N N M G V E C L R L L N E I I A D F D E L M D K D F Y K D L E
2761  AAGATCAAGACCATTGGGAGCACGTACATGGCTGCTGTGGGGCTGGCGCCCACTGCTGGGACCAAGGCTAAGAAGTGCATCTCCTCCCACCTCAGCACGTTGGCAGATTTTGCCATCGAG
921   K I K T I G S T Y M A A V G L A P T A G T K A K K C I S S H L S T L A D F A I E
2881  ATGTTTGATGTCCTGGATGAGATCAACTACCAGTCTTATAACGACTTTGTGCTCCGTGTTGGCATCAATGTTGGCCCCGTGGTGGCTGGAGTGATCGGGGCTCGCAGGCCGCAGTATGAC
961   M F D V L D E I N Y Q S Y N D F V L R V G I N V G P V V A G V I G A R R P Q Y D
3001  ATTTGGGGGAACACGGTCAATGTGGCCAGTCGGATGGACAGTACCGGCGTCCAGGGCAGGATCCAGGTCACAGAGGAAGTTCACCGGCTGCTGCGGCGGGGTTCCTACCGCTTCGTGTGC
1001  I W G N T V N V A S R M D S T G V Q G R I Q V T E E V H R L L R R G S Y R F V C
3121  CGAGGCAAAGTCAGTGTCAAGGGCAAGGGTGAGATGCTGACATACTTCCTGGAAGGCAGGACCGATGGAAATGGCTCCCAAACCAGGTCCCTGAACTCAGAGCGGAAAATGTATCCTTTC
1041  R G K V S V K G K G E M L T Y F L E G R T D G N G S Q T R S L N S E R K M Y P F
3241  GGAAGAGCTGGCCTCCAGACCCAGACTGGCTGCGGGTCACCCCCCGGTACCTCCTGCAGCCCGGCCTCCCAGTTGGAGCTGGGCCAGGGGCTCTGCAGGGCTCGGGGCTTGCCCCAGGCCCC
1081  G R A G L Q T R L A A G H P P V P P A A G L P V G A G P G A L Q G S G L A P G P
3361  CCAGGCCAACACCTGCCCCCTGGAGCCTCTGGGAAGGAGGCTTAGTGGAGCCCATGCCAGCCGCTTGGGGCACAGGGCACAAATGCTTGAGGTATTGGGGGTCTTTGGGGCTCCCCCAGG
1121  P G Q H L P P G A S G K E A   1134
3481  GACCAGCCCAGCCAGCAGAGCAGGGCTGGGAGCTGGTGACTGGGCTGGGGAAGGGGCACAGTCCAGGCATGACTTGAAGCAGCTGGGCAGTGAGTCCAGCAGTGAGTCGGGTGAGGGGCG
3601  AACCCTGACCGCACGGACACCGAGGTTTTAGCGGCGACTGTGTTTGCTTTGTCCTCACCTCAGTGGCAGAGGGAGGTTGGCTGGGTCAGCTCCAGGTGTCAGAACGTTCAGGACATTCTC
3721  CAGGGCCATAGGCTACAGTGAGAGCCTCAGAGGTCTGGCTGGCAGAGTGACTGGGGCGCCCACCTTGTGCCCTGAGATGGATCAGCGTCAGCCCAGCAGTGTGGGTCCATGGGGGCAGCT
3841  CCAGCCCTGGGCAGTGGCCTGCAGAGCCCGGAGTGG   3876
```

Fig. 3 (facing page). Complete coding sequence for the 120-kD form of bovine brain adenylyl cyclase. The sequences of the tryptic peptides that were determined by amino acid sequencing are underlined. The residue preceding the peptide is not always R or K because certain amino acids could not be identified unambiguously in initial cycles of sequencing. Clone p1-6 was sequenced largely with the use of the 3′,5′ exonuclease activity of T4 polymerase to generate a series of nested deletion mutants (*56*). The remainder of the sequence was determined by extending specific primers that were synthesized on the basis of the previously determined sequence. All of the sequence was determined on both strands; the modified T7 polymerase (Sequenase, U.S. Biochemicals) was used to perform dideoxy sequencing reactions (*57*). Throughout the entire sequence, at least one strand was determined with a nucleotide analog, either deoxyinosine triphosphate (dITP) or 7-deaza-deoxyguanosine triphosphate (dGTP), in place of dGTP in order to reduce artifacts caused by secondary structure. Near the 5′ end of the sequence, polyacrylamide sequencing gels included 40 percent formamide to reduce compressions. In addition, the sequence in this region was verified with the polymerase from *Thermus aquaticus* so that the chain extension reactions could be performed at 70°C (Thermal Base Sequencing Kit, Stratagene).

purpuratus (*26*)—as well as the 70-kD subunit of the soluble guanylyl cyclase from bovine lung (*27*). In the case of the soluble enzyme, there is no knowledge of the functional characteristics of the resolved subunit of the heterodimer (*27, 28*). In each case, the pairwise alignment of the respective guanylyl cyclase to amino acids 258 to 480 of adenylyl cyclase (or to amino acids 809 to 1059 of adenylyl cyclase) shows similar relationships; the domains are, on average, 30 percent identical and 55 percent similar. The fourth guanylyl cyclase is another membrane-bound protein from a more primitive sea urchin, *Arbacia punctulata* (*29*). In this case, the similarity is limited to the first 57 positions that are common to the other alignments, with a 53 percent similarity to amino acids 258 to 315 of the bovine brain adenylyl cyclase. We presume that the chemistry of catalysis by adenylyl cyclases and guanylyl cyclases is similar; however, the regulatory properties of these enzymes differ (*30*).

When the two similar domains of adenylyl cyclase were first characterized as a profile and then compared with the National Biomedical Research Foundation (NBRF) Protein Identification Resource, the eight highest quality scores observed were with nucleotide-binding proteins (*31*). The second highest score was for adenylyl cyclase from *Saccharomyces cerevisiae*. An alignment of the two similar domains of bovine brain adenylyl cyclase to amino acids 1624 to 1854 of the yeast adenylyl cyclase (Fig. 6B) indicates that only 15 percent of the residues are identical, although 41 percent match if conservative substitutions are counted. However, this region of yeast adenylyl cyclase is within a stretch of 400 amino acid residues that contains the catalytic domain (*32*). The large size (2026 amino acid residues), the lack of obvious transmembrane spans, and the different mechanisms of regulation of yeast adenylyl cyclase relative to mammalian adenylyl cyclases led us to suspect substantial differences between the proteins; however, the fact that they both catalyze the same reaction implies similarities. This may be reflected in the aligned sequences of Fig. 6B, if these regions do indeed specify nucleotide binding domains. It is thus particularly interesting that the mammalian adenylyl cyclase has two of these domains.

A hydropathy plot of the protein sequence deduced from the adenylyl cyclase cDNA's (*33*) shows two intensely hydrophobic regions (Fig. 7). Our analysis of the sequence (*34*) suggests that each of these regions contains six transmembrane spans (numbered 1 to 12 in Fig. 7). In addition, this analysis suggests the possibility of an additional transmembrane span at residues 974 to 994. For the sake of our initial, speculative attempt to model the topology of adenylyl cyclase in the plasma membrane (Fig. 8), we have ignored this 13th potential span for the following reasons. (i) Residues 974 to 994 lie at the heart of the region of homology between residues 238 to 480 and 809 to 1059 of the adenylyl cyclase sequence and between these sequences and the carboxyl-terminal cytoplasmic domain of the guanylyl cyclases, as discussed above. (ii) We have recently isolated cDNA's that appear to encode an additional form of adenylyl cyclase (type II) from both bovine and rat brain libraries (*21*). Hydropathy analysis of the carboxyl-terminal half of the type II adenylyl cyclase suggests six transmembrane spans, grouped just as shown in Fig. 7. In the region in question, the amino acid sequence of the type II adenylyl cyclase is similar to that shown in Fig. 3 (18 of 21 identities

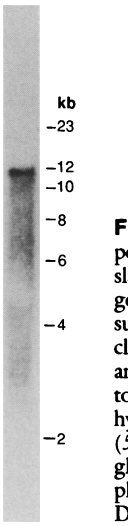

Fig. 4. High stringency RNA blot analysis of bovine brain poly(A)$^+$ RNA. Cerebral cortex was removed from a freshly slaughtered cow, coarsely chopped, and frozen in liquid nitrogen. A portion (20 g) was used for the purification of RNA and subsequent purification of poly(A)$^+$ RNA. The insert from clone p1-6 was prepared by Eco RI digestion, labeled with ^{32}P, and used as the probe. Methods for isolation of RNA, transfer to Nytran filters for RNA blot analysis, labeling of the probe, hybridization, and washing of the filters have been described (*58*). Molecular sizes of the RNA species were determined from glyoxal-treated 1-kb ladder (Bethesda Research Laboratories) placed in an adjacent lane and probed with ^{32}P-labeled marker DNA.

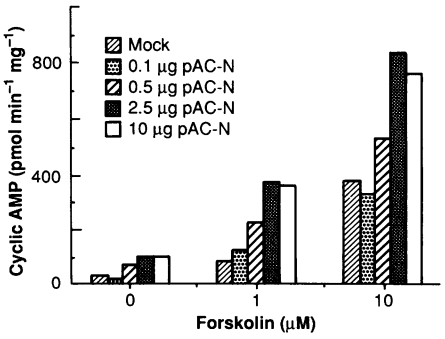

Fig. 5. Transient expression in COS-m6 cells. Cyclic AMP formation was measured as a function of forskolin concentration, with different amounts of plasmid pAC-N in the transfection. For the construction of a cDNA with a full-length coding sequence (pAC-N), two Eco RI–Bam HI fragments were ligated into Eco RI–digested pMK-16; these two fragments were (i) the sequence from nucleotides minus 102 to 701 in pA11 and (ii) the sequence from the Bam HI site at position 3060 to the first Eco RI site in p1-8 (approximately 900 bp). The plasmid containing these two fragments was digested with Bam HI and ligated with the Bam HI fragment of pA19 (nucleotides 702 to 3059) to form a clone with the complete adenylyl cyclase coding sequence (pMK-AC). The DNA containing the entire adenylyl cyclase coding sequence was then released from pMK-AC by Eco RI digestion and cloned into the mammalian expression vector pCMV5 (*59*). The construct pAC-N is oriented so that the cytomegalovirus promoter of pCMV-5 directs expression of the complete open reading frame presented in Fig. 3. COS-m6 cells were seeded on 60-mm plates the day before transfection so that they were approximately 50 to 80 percent confluent for the experiment. The cells were transfected with the indicated amount of plasmid (without carrier) by the DEAE-dextran method (*60*) and treated with glycerol for 4 minutes (*61*). After 48 hours, the cells were scraped from the plate in a solution containing 20 m*M* Hepes (sodium salt) (*p*H 8.0), 2 m*M* $MgCl_2$, and 1 m*M* EDTA and supplemented with the protease inhibitors that are included during the purification of bovine brain membranes (*12*). The cells were collected into a microfuge tube, frozen in liquid nitrogen, and thawed rapidly. This suspension was sedimented (microfuge) for 15 minutes, the supernatant was removed, and the pellet was resuspended in the same solution used to scrape the cells from the plates. The washing was repeated and the final resuspended sample was passed through a 22-gauge needle ten times. Adenylyl cyclase activity was assayed as described (*6*). Cells transfected with plasmid containing the full-length construct in the reverse orientation contained adenylyl cyclase activities equal to those observed with mock-transfected cells. Data shown are from a single experiment that is representative of four such experiments.

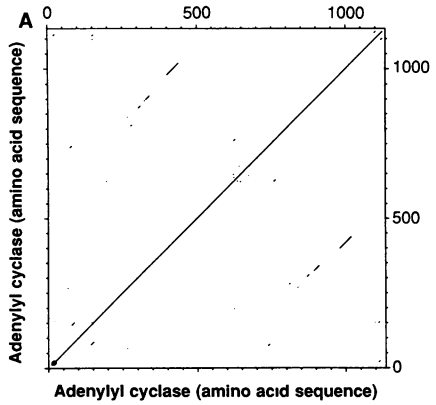

Fig. 6. (**A**) Dot matrix comparison of bovine brain adenylyl cyclase with itself. The Compare and DotPlot programs were used to generate the figure (*23*). A window of 20 with a stringency of 14 was used in the Compare program. In the program the Dayhoff mutational matrix (*24*) is used for assigning a value to every possible substitution of one amino acid for another. The values have been adjusted so that the average pair receives a score of 0 and the standard deviation is 1; an identity receives a score of 1.5 (*62*). (**B**) Sequence alignment of two sequences within bovine brain adenylyl cyclase (AC-A and AC-B), as well as sequences from four different guanylyl cyclases and one sequence from *Saccharomyces cerevisiae* adenylyl cyclase. The sequences shown are membrane-bound guanylyl cyclases from rat brain (RB GC) (*25*), *Strongylocentrotus purpuratus* (SP GC) (*26*), *Arbacia punctulata* (AP GC) (*29*), the soluble guanylyl cyclase subunit from bovine lung (SOL GC) (*27*), and adenylyl cyclase from *Saccharomyces cerevisiae* (SC AC) (*32*). The strong similarity between the bovine adenylyl cyclase and the guanylyl cyclases is readily detected by comparing the entire protein sequences. The Gap program was used to obtain the alignments (*23*). Conservative substitutions were defined as (1) C; (2) S, T, P, A, G; (3) N, D, E, Q; (4) H, R, K; (5) M, I, L, V; (6) F, Y, W. Any match within a group was given a score of 1, and a nonmatch was scored as 0. The gap weight was 5 and the gap-length weight was 0.3. For the consensus sequence listed, if four of the sequences contain the same residue at the given position, the amino acid is listed in the consensus. If four of the sequences contain amino acid residues from the same closely related group, then that group is referred to at that position in the consensus by a number 1 to 6, corresponding to the groups listed above. However, four identities are always given priority over conservative substitutions. The original detection of similarity between bovine brain adenylyl cyclase and the yeast adenylyl cyclase was made by profile analysis (*55*). An alignment of the two similar regions within bovine adenylyl cyclase was used to generate the profile. One alignment with the *Saccharomyces cerevisiae* adenylyl cyclase was again optimized, but with respect to the individual adenylyl cyclase sequences and the above list of conservative substitutions.

B

		1				50
AC-A	258	LEDENEKQER	LLMSLLPRNV	AMEMKEDFLK	PPERIFHKIY	IQRHDNVSIL
AC-B	829	MEKVKLDNKR	ILFNLLPAHV	AQHFLMSNPR	NMD.LYYQSY	SQVGVMFASI
RB GC	807	LE.EKRKAEA	LLYQILPHSV	AEQL.....	...KRGETVQ	AEAFDSVTIY
SP GC	855	LQKEKTKTEQ	LLHRMLPPSI	ASQL.....	...IKGIAVL	PETFEMVSIF
SOL GC	383	LEDEKKKTDT	LLYSVLPPSV	ANELRHKRPV	PA.....KRY	DNVTILFSGI
AP GC	852	LQKEKAKTEQ	LLHRMLPPSI	ASQL.....	...IKGISVL	PETFDMVSIF
SC AC	1624	LYENIQQQNR	...FTLNKNS	LMTRRSTFED	TTLRRLQPEI	SPPTGNLAMV
CONSENSUS		LE-EK-K2E-	LL--5LP2SV	A-3L------	----4--25-	232-35VSI5
		51				100
AC-A	308	FADIVGFTGL	ASQCTAQELV	KLLNELFGKF	DELATENHCR	...RIKILGD
AC-B	878	PNFNDFYIEL	DGNNMGVECL	RLLNEIIADF	DELMDKDFYK	DLEKIKTIGS
RB GC	847	FSDIVGFTAL	SAESTPMQVV	TLLNDLYTCF	DAVIDN...F	DVYKVETIGD
SP GC	896	FSDIVGFTAL	SAASTPIQVV	NLLNDLYTLF	DAIISN...Y	DVYKVETIGD
SOL GC	428	VGFNAFCSKH	ASGEGAMKIV	NLLNDLYTRF	DTLTDSRKNP	FVYKVETVGD
AP GC	893	FSDIVGF...	899			
SC AC	1671	FTDIKSSTFL	WELFP....N	AMRTAIKTHN	DIMRRQLRIY	GGYEVKTEGD
CONSENSUS		F2DIVGFT-L	22--22535V	-LLN3L6T-F	D-5--3----	-5YKV-T5GD
		101				150
AC-A	355	CYYCVSGLTQ	PKTDHAHCCV	.....EMG..	.LDMIDTITS	..VAEATEVD
AC-B	928	TYMAAVGLAP	TAGTKAKKCI	SSHLSTLADF	AIEMFDVL..	DEINYQSYND
RB GC	894	AYMVVSGLPV	RNGQ.....	.LHAREVARM	ALALLDAVRS	FRIRHRPQEQ
SP GC	943	AYMLVSGLPL	RNGDR.....	..HAGQIAST	AHHLLESVKG	FIVPHKPEVF
SOL GC	478	KYMTVSGLPE	PCIHHAR...	.....SICH	LALDMMEIAG	QVQVDG..ES
SC AC	1717	AF..MVAFPT	PTSGLTWCLS	GQLKLLDAQW	PEEITSVQDG	CQVTDRNGNI
CONSENSUS		2YM-VSGLP-	2-2-42----	------5A--	2--553-5-2	--5---2-3-
		151				200
AC-A	395	LNMRVGLHTG	...RVLCGVL	GLRKWQYDVW	SNDVTLANVM	EAAGLPGKVH
AC-B	976	FVLRVGINVG	...PVVAGVI	GARRPQYDIW	GNTVNVASRM	DSTGVQGRIQ
RB GC	937	LRLRIGIHTG	...PVCAGVV	GLKMPRYCLF	GDTVNTASRM	ESNGEALKIH
SP GC	986	LKLRIGIHSG	...SCVAGVV	GLTMPRYCLF	GDTVNTASRM	ESNGLALRIH
SOL GC	517	VQITIGIHTG	...EVVTGVI	GQRMPRYCLF	GNTVNLTSRT	ETTGEKGKIN
SC AC	1765	IYQGLSVRMG	IHWGGCPVPEL	DLVTQRMDYL	GPMVNKAARV	QGVADGGQIA
CONSENSUS		5-5R5GIH2G	---2V52GV5	GL4-PRY-56	G3TVN-ASRM	E2-G-2G4I-
		201			243	
AC-A	442	ITKTTLACLN	GDYEVEPGHG	HERNSFL...	..KTHNIETF	FIV 479
AC-B	1023	VTEEVHRLL.	....RRGSYR	FVCRGK.VSV	KGKGEMLTYF	LEG 1059
RB GC	984	LSSETKAVLE	EFDGFELELR	GDVE.....M	KGKGKVRTYW	LLG 1021
SP GC	1033	VSPWCKQVLD	.....KLGGY	ELEDRGLVPM	NGKGEIHTFW	LLG 1070
SOL GC	564	VSEYTYRCLM	TPENSDPQFH	LEHRGPVSNK	GKKEPMQVWF	LSR 606
SC AC	1815	MSSDFYSEFN	KIMKYHERVV	KGKES.LKEV	YGEEIIGEV.	... 1852
CONSENSUS		5S------L3	----------	---32-5--5	-GK2-5--66	L--

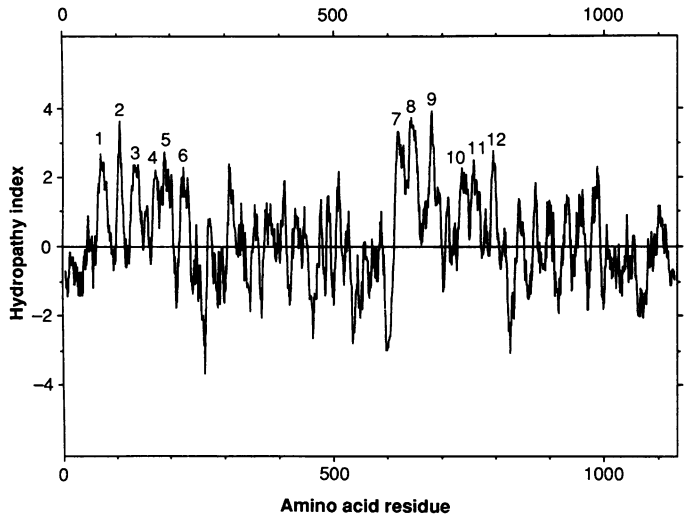

Fig. 7. Hydropathy plot of the amino acid sequence deduced from the adenylyl cyclase cDNA (*33*). Hydrophobic indices were calculated using the Kyte-Doolittle scale and averaging over a window of nine amino acid residues. The 12 highest peaks of hydrophobicity have been numbered in order of their appearance within the sequence. The data were computed with the program PepPlot (*23*).

from residues 974 to 994). Nevertheless, two crucial amino acid substitutions (Val → Arg; Val → Lys) lower the hydropathy score for this sequence enough to make a transmembrane span quite unlikely. The entire carboxyl-terminal 269 amino acid residues of the type II enzyme that follow the six transmembrane spans are probably cytoplasmic; we have modeled the sequence shown in Fig. 3 (designated type I adenylyl cyclase) similarly. Therefore, most of the protein shown in the model (Fig. 8) has been distributed to the cytoplasmic side of the membrane.

Some of the connecting loops between the predicted transmembrane helices (Fig. 8) are short. Slight adjustments in the exact locations of the spans could alleviate these problems without the loss of the stretches of 21 amino acid residues with a sufficiently high average hydrophobic index to be characteristic of transmembrane spans. Although models for the localization of transmembrane spans are based on alpha helices, other conformations are possible, particularly if the protein has many transmembrane spans (*35*). Thus, one layer of amphipathic helices might form a protein insulator that could allow other regions of the polypeptide to traverse the membrane without necessarily coming into contact with the bilayer throughout its entire length.

Several of the proposed transmembrane helices are amphipathic. Charged residues within the membrane have been enclosed in circles (negative) or boxes (positive) in Fig. 8. For example, helix 2 contains two histidines, one lysine, and a threonine that could all be aligned along one face; helix 12 contains two serines and both an arginine and a glutamate that could form a salt bridge in the membrane. Depending on the relative juxtaposition of these helices, these hydrophilic groups could interact with one another in the membrane. Such hypothetical interactions could reflect requirements for protein packing.

An interesting sequence is present near the amino terminus of the molecule. After a stretch of seven glycine residues, the sequence from positions 16 to 21 is GAGESG (*36*)—identical to a highly

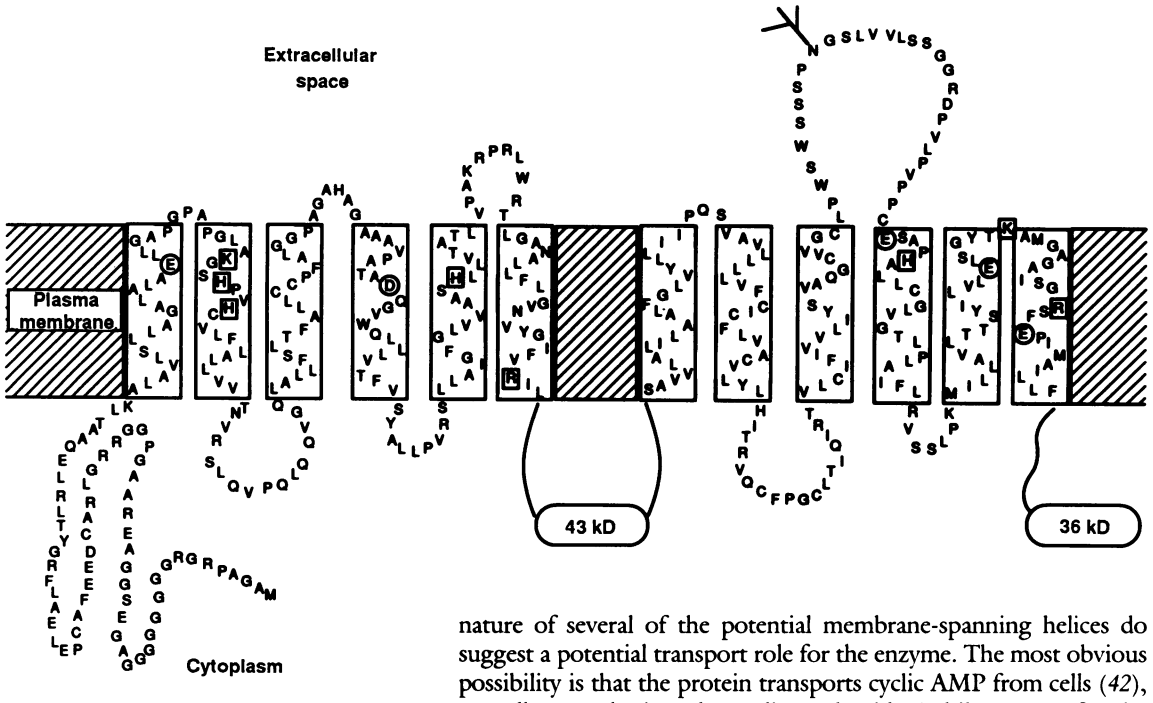

Fig. 8. A schematic, two-dimensional model of adenylyl cyclase in the plasma membrane. Charged residues within the putative membrane spans have been enclosed in circles (negative) or boxes (positive).

conserved region of the guanine nucleotide binding pocket of the G protein family (*2*). This sequence also falls into the category of a consensus nucleotide binding sequence, GXGXXG (*37*). Perhaps two or all of the cytoplasmic domains may need to interact to form a functional catalytic site. Alternatively, the GAGESG sequence or one of the other putative nucleotide binding sequences (or both) could contribute to the so-called P site, where adenosine analogs interact to inhibit adenylyl cyclase activity noncompetitively (*8, 38*). Other sequences of interest include four possible N-linked glycosylation sites. In our model, only one of these would be exposed extracellularly (Asn^{706}). A 120-kD form of adenylyl cyclase is phosphorylated by protein kinase C in frog erythrocytes (*39*); a possible site for modification is Ser^{481}. Similarly, a consensus site for phosphorylation by cyclic AMP–dependent protein kinase is present at Ser^{1035}.

In view of the known function and intracellular site of regulation of adenylyl cyclase, it is surprising that the enzyme contains numerous transmembrane spans. Furthermore, application of the usual criteria for prediction of the location and number of such spans has led us to propose a structure that is reminiscent of those that have been proposed for several transporters and channels. The latter include other G protein–regulated entities, such as dihydropyridine-sensitive Ca^{2+} channels and K^+ channels (*40*). The common motif is one or more sets of six transmembrane spans; large cytoplasmic domains separate multiple sets of spans. Particularly striking is the overall topographical resemblance between adenylyl cyclase and the P glycoprotein—the product of the multidrug resistance gene (*41*). This protein, which is believed to export a variety of chemicals from cells, is thought to consist of a short amino-terminal cytoplasmic tail, a set of six transmembrane spans, a large cytoplasmic domain, a second set of six membrane spans, and a large carboxyl-terminal cytoplasmic domain. Each of the large cytoplasmic domains is believed to contain a binding site for adenosine triphosphate.

Although the topographical similarity between adenylyl cyclase and various channels and transporters is intriguing, it is not complemented by substantial similarity of primary amino acid sequence. Nevertheless, this relationship, the possibility of multiple nucleotide binding sites in adenylyl cyclase, and the amphipathic nature of several of the potential membrane-spanning helices do suggest a potential transport role for the enzyme. The most obvious possibility is that the protein transports cyclic AMP from cells (*42*), as well as synthesizes the cyclic nucleotide. While reasons for the linkage of these two functions in higher eukaryotes are not obvious, the cellular slime mold *Dictyostelium discoideum* synthesizes cyclic AMP and then exports it as an extracellular signal for chemotaxis, aggregation, and differentiation (*43*).

REFERENCES AND NOTES

1. T. W. Rall, E. W. Sutherland, J. Berthet, *J. Biol. Chem.* **224**, 463 (1957).
2. A. G. Gilman, *Annu. Rev. Biochem.* **56**, 615 (1987).
3. H. G. Dohlman, M. G. Caron, R. J. Lefkowitz, *Biochemistry* **26**, 2657 (1987).
4. M. S. Livingstone, P. P. Sziber, W. G. Quinn, *Cell* **37**, 205 (1984).
5. E. Pfeuffer, S. Mollner, T. Pfeuffer, *EMBO J.* **4**, 3675 (1985).
6. M. D. Smigel, *J. Biol. Chem.* **261**, 1976 (1986).
7. S. Mollner and T. Pfeuffer, *Eur. J. Biochem.* **171**, 265 (1988).
8. A. M. Minocherhomjee, S. Selfe, N. J. Flowers, D. R. Storm, *Biochemistry* **26**, 4444 (1987).
9. T. Pfeuffer and H. Metzger, *FEBS Lett.* **146**, 369 (1982).
10. D. C. May, E. M. Ross, A. G. Gilman, M. D. Smigel, *J. Biol. Chem.* **260**, 15829 (1985).
11. D. Feder *et al.*, *EMBO J.* **5**, 1509 (1986).
12. Forskolin-Sepharose was synthesized from 7-succinyl-7-deacetylforskolin as described (*44*). Crude brain membranes were purified from bovine cerebral cortex as described (*45*), except that the homogenization buffers were supplemented with 2 m*M* dithiothreitol; leupeptin and lima bean trypsin inhibitor each at 3.2 mg/liter; aprotinin at 0.5 mg/liter; and L-1-tosylamido-2-phenylethyl chloromethyl ketone, 1-chloro-3-tosylamido-7-amino-2-heptanone, and phenylmethylsulfonyl chloride each at 22 mg/liter. The buffers used for washing were also supplemented with 2 m*M* dithiothreitol, but the mixture of protease inhibitors was present at one-third the concentrations indicated above. Crude bovine brain membranes (1.6 g of total protein) were suspended at 4 mg of protein per milliliter in 50 m*M* Hepes sodium salt (*p*H 8.0), 1 m*M* $MgCl_2$, 1 m*M* EDTA, 2 m*M* dithiothreitol, and the mixture of protease inhibitors used for homogenization. Lubrol PX was added with stirring to a final concentrations of 0.6 percent. This mixture was homogenized manually, insoluble material was removed by centrifugation (25 minutes at 50,000*g*), and NaCl was added to the supernatant to a final concentration of 0.5*M*. This extract was incubated with 30 ml of forskolin-Sepharose overnight at 4°C, with gentle rocking. The resin was poured into a column and the column was washed sequentially with 500 ml of buffer A containing 0.5*M* NaCl, 500 ml of buffer A with 2*M* NaCl, and 60 ml of buffer A with 0.5*M* NaCl and 2 percent dimethyl sulfoxide. [Buffer A is 10 m*M* Hepes sodium salt (*p*H 8.0), 1 m*M* EDTA, 1 m*M* $MgCl_2$, 2 m*M* dithiothreitol, 0.05 percent Lubrol PX, and the protease inhibitors used for the homogenization of crude membranes at one-half the indicated concentrations.] Adenylyl cyclase was eluted with 30 ml of 0.2 m*M* forskolin in buffer A containing 0.5*M* NaCl; the flow was stopped for 15 minutes, and an additional 30 ml of the elution buffer was added. This cycle of washing and incubation was repeated four times. Adenylyl cyclase activity was assayed as described (*6*). The specific activity of the enzyme eluted from the hydroxylapatite column was approximately 2 to 4 μmol/(mg·min) in the presence of 100 μ*M* forskolin and 20 m*M* $MgCl_2$. The yield of activity from this procedure averaged 0.5 to 1 μmol/min. For the SDS-polyacrylamide gel (Fig. 1A), adenylyl cyclase was precipitated from 1 ml of forskolin-Sepharose eluate by addition of trichloroacetic

acid (to 15 percent) and centrifugation for 15 minutes at 14,000*g* at 4°C. The pellet was washed with acetone at −20°C, resuspended in gel electrophoresis sample buffer, boiled for 5 minutes, and subjected to electrophoresis (SDS-PAGE) on a 7.5 percent gel; the gels were then stained with silver. The markers are myosin (200 kD), β-galactosidase (116 kD), phosphorylase B (92.5 kD), bovine serum albumin (66.2 kD), and ovalbumin (45 kD).
13. G. B. Rosenberg and D. R. Storm, *J. Biol. Chem.* **262**, 7623 (1987).
14. These preparations usually contain an additional 150-kD protein band (faintly visible in Fig. 1A). cDNA clones that correspond to this band were isolated, nucleotide sequences corresponding to approximately 70 percent of the protein were determined, and antisera to synthetic peptides were generated. There is no obvious homology between the deduced amino acid sequence and that of any other known protein. Quantitative immunoblotting indicates that this 150-kD species constitutes 1 percent of total brain membrane protein. A small fraction of this 150-kD species contaminates forskolin-Sepharose eluates and is not enriched by this column. We presume that it is not related to adenylyl cyclase, nor does it appear to be the same 150-kD protein noted by Mollner and Pfeuffer (*7*).
15. Bovine brain adenylyl cyclase (1 to 2 nmol; partially purified on forskolin-Sepharose) was concentrated and separated from remaining contaminants by SDS-PAGE (*12*), except that we used a preparative gel and did not stain it. The protein in the gel was transferred to nitrocellulose (*46*), this filter was stained with amido black (*47*), and the 120-kD band was excised (yield, approximately 300 pmol). The equipment and basic protocol for tryptic digestion and the first HPLC separation were as described (*48*); peaks were collected in plastic tubes instead of onto filter disks. Lubrol PX (10 percent) was added to the samples to a final concentration of 0.1 percent, the volumes were reduced (a stream of nitrogen gas) to approximately 20 μl, and the samples were adjusted to approximately 100 μl with 0.1 percent aqueous ammonium acetate, *p*H 6.5. These samples were chromatographed as above for 60 minutes at a flow rate of 50 μl/min on an acetonitrile gradient (10 to 50 percent, v/v) with 0.1 percent ammonium acetate, *p*H 6.5 (*49*). Peaks were collected onto 1-cm Whatman GF/C disks, alkylated, and reduced (*50*), and then sequenced (*48*).
16. R. H. Aebersold, J. Leavitt, R. A. Saavedra, L. E. Hood, S. B. H. Kent, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6970 (1987).
17. N. Green *et al.*, *Cell* **28**, 477 (1982).
18. R. Lathe, *J. Mol. Biol.* **183**, 1 (1985). The portion of the peptide sequence utilized to construct a probe was NEIIADFDELMDKDFYK (*36*). The nucleotide sequence for the complementary strand probe was 5′-TT GTA GAA GTC CTT GTC CAT CAG CTC ATC AAA GTC AGC AAT GAT CTC ATT-3′. This sequence proved to be correct in 46 of 50 positions.
19. Unless otherwise indicated, routine manipulations were performed as described (*51*). Initially, an oligonucleotide probe (*18*) was used to screen 1.5 × 10^5 phage plaques from the bovine brain library generated with inserts greater than 4 kb (5 × 10^4 independent recombinants). Plaque lifts were performed with nitrocellulose filters, and filters were washed at 50°C in 3*M* tetramethylammonium chloride (*52*). Four independent clones were isolated. The open reading frame predicted by these clones specifies 12 of the tryptic peptides. An oligonucleotide probe was synthesized complementary to the sequence corresponding to nucleotides 831 to 868 (Fig. 3, the 5′ end of p1-6). One clone isolated with this probe extends 347 nucleotides in the 5′ direction. Its coding sequence specifies an additional tryptic peptide. With the Eco RI restriction site in the polylinker, a 312-bp Eco RI–Bam HI restriction fragment was isolated. This fragment was ^{32}P-labeled by random priming (*53*) and was used to screen the libraries containing the 1- to 2-kb and 2- to 4-kb size fractions of the randomly primed, bovine brain cDNA (*54*). We isolated 29 independent clones. Although none extended more 5′ than pA11, an independent clone extended to within 15 nucleotides of the 5′ end of pA11. The coding sequence of pA11 specifies the last of the tryptic peptides. Clones were defined as overlapping by two criteria. (i) Digestion with selected restriction enzymes gave fragments consistent with other sequenced clones. (ii) Approximately 200 nucleotides of exact overlapping sequence was determined. Typically, half of this corresponded to sequence from one end of the clone, while the other half was determined with an internal primer.
20. M. Kozak, *Nucleic Acids Res.* **15**, 8125 (1987).
21. R. R. Reed, unpublished observations.
22. T. F. Smith and M. S. Waterman, *Adv. Appl. Math.* **2**, 482 (1981).
23. J. Devereux, P. Haeberli, O. Smithies, *Nucleic Acids Res.* **12**, 387 (1984).
24. M. O. Dayhoff, W. C. Barker, L. T. Hunt, *Methods Enzymol.* **91**, 524 (1983).
25. M. Chinkers *et al.*, *Nature* **338**, 78 (1989).
26. D. S. Thorpe and D. L. Garbers, *J. Biol. Chem.* **264**, 6465 (1989).
27. D. Koesling *et al.*, *FEBS Lett.* **239**, 29 (1988).
28. M. Nakane, S. Saheki, T. Kuno, K. Ishii, F. Murad, *Biochem. Biophys. Res. Commun.* **157**, 1139 (1988).
29. S. Singh *et al.*, *Nature* **334**, 708 (1988).
30. Singh *et al.* (*29*) noted an amino acid similarity between a region of the guanylyl cyclase from *A. punctulata* and the catalytic domain of the protein kinase family. This similarity includes 26 of 33 amino acids that are highly conserved among the protein kinases. They suggest that these variations may reflect differences in nucleotide binding specificity, or that guanylyl cyclase may have previously unobserved phosphotransferase activity. The region of amino acid similarity between the guanylyl cyclases and the protein kinases is not a part of the region that is similar to bovine brain adenylyl cyclase.
31. The repeated regions of bovine brain adenylyl cyclase were aligned and then characterized as a profile (*23*). Profile analysis includes all of the information in the alignment, so that in the profile a position in the sequence is represented by 21 entries in a matrix (*55*). The first 20 represent a score for finding each of the 20 possible amino acids at that given position within the sequence. These entries are evaluated by means of the Dayhoff mutational distance matrix (*24*). The 21st entry in the matrix represents a gap penalty for that given position. For example, if some of the sequences in the alignment included a gap at a given position, then the gap penalty associated with that postion would be lower than that associated with a position in which none of the sequences required a gap (*55*). The mean score for the comparison of the profile with the protein data banks was 52 ± 25 (mean ± SD, n = 9138). The eight highest quality scores observed were associated with proteins that bind nucleotides. They were type II restriction enzyme Bsu RI (95.86); adenylyl cyclase from *Saccharomyce cerevisiae* (94.14); GMP (guanosine 5′-phosphate) synthase from *Escherichia coli* (93.99); fatty acyl–coenzyme A oxidase from *Candida tropicalis* (93.57); the Na-,K-ATPase (adenosine triphosphate) from rat (93.18), and pig (93.07), and a second form from rat (92.95), and the exodeoxyribonuclease V from *E. coli* (92.92).
32. T. Kataoka, D. Broek, M. Wigler, *Cell* **43**, 493 (1985).
33. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
34. D. Eisenberg, E. Schwarz, M. Komaromy, R. Wall, *ibid.* **179**, 125 (1984).
35. H. F. Lodish, *Trends Biochem. Sci.* **13**, 332 (1988).
36. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
37. R. K. Wierenga and W. G. J. Hol, *Nature* **302**, 842 (1983).
38. C. Londos, J. Wolff, D. M. F. Cooper, in *Physiological and Regulatory Functions of Adenosine and Adenine Nucleotides*, H. P. Baer and G. I. Drummond, Eds. (Raven, New York, 1979), p. 271.
39. T. Yoshimasa *et al.*, *Nature* **327**, 67 (1987).
40. W. A. Catterall, *Science* **242**, 50 (1988).
41. M. M. Gottesman and I. Pastan, *J. Biol. Chem.* **263**, 12163 (1988).
42. P. R. Davoren and E. W. Sutherland, *ibid.* **238**, 3009 (1963).
43. G. Gerisch, *Annu. Rev. Biochem.* **56**, 853 (1987).
44. E. Pfeuffer, R.-M. Drehev, H. Metzger, T. Pfeuffer, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3086 (1985).
45. P. C. Sternweis and J. D. Robishaw, *J. Biol. Chem.* **259**, 13806 (1984).
46. H. Towbin, T. Staehelin, J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350 (1979).
47. W. Schaffner and C. Weissmann, *Anal. Biochem.* **56**, 502 (1973).
48. G. Gil *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8963 (1988).
49. M. Noda *et al.*, *Nature* **312**, 121 (1984).
50. P. C. Andrews and J. E. Dixon, *Anal. Biochem.* **161**, 1 (1987).
51. T. Maniatis, E. Fritsch, J. Sambrook, *Molelcular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
52. W. I. Wood, J. Gitschier, L. A. Lasky, R. M. Lawn, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1585 (1985).
53. A. P. Feinberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983).
54. W. M. Strauss, in *Current Protocols in Molecular Biology*, F. M. Ausubel *et al.*, Eds. (Wiley-Interscience, New York, 1987), p. 6.3.1.
55. M. Gribskov, A. D. McLachlan, D. Eisenberg, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4355 (1987).
56. R. M. K. Dale, B. A. McClure, J. P. Houchins, *Plasmid* **13**, 31 (1985).
57. S. Tabor and C. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4767 (1987).
58. D. T. Jones and R. R. Reed, *J. Biol. Chem.* **262**, 14241 (1987).
59. S. Andersson, D. N. Davis, H. Dahlback, H. Jornvall, D. W. Russell, *ibid.* **264**, 8222 (1989).
60. G. Adams and J. Rose, *Cell* **41**, 1007 (1985).
61. M. A. Lopata *et al.*, *Nucleic Acids Res.* **12**, 5707 (1984).
62. M. Gribskov and R. R. Burgess, *ibid.* **14**, 6745 (1986).
63. We thank D. Garbers (Vanderbilt University) for providing guanylyl cyclase sequences before their publication and D. Garbers, E. M. Ross, and J. R. Falck for helpful discussion; R. Dixon (Merck Sharp & Dohme) for the cDNA library utilized in this work; and K. Vickroy, C. Moomaw, and G. Dorsey for technical assistance. Supported by NIH grants GM34497 and CA16519, American Cancer Society grant BC555K, and the Raymond and Ellen Willie Chair of Molecular Neuropharmacology; NIH postdoctoral fellowship GM12230 (J.K.); Fogarty International Fellowship Award TW03913 (F.C.); NIH training grant 3T32 GM07445 (H.A.B); and grants from the Perot Family Foundation and the Lucille P. Markey Charitable Trust. The nucleotide sequence data reported in this article have been submitted to GenBank and assigned accession number M25579.

14 March 1989; accepted 24 May 1989