A Chemoattractant Receptor Controls Development in *Dictyostelium discoideum*

PETER S. KLEIN, TZELI J. SUN, CHARLES L. SAXE III, ALAN R. KIMMEL
RONALD L. JOHNSON, PETER N. DEVREOTES*

During the early stages of its developmental program, *Dictyostelium discoideum* expresses cell surface cyclic adenosine monophosphate (cyclic AMP) receptors. It has been suggested that these receptors coordinate the aggregation of individual cells into a multicellular organism and regulate the expression of a large number of developmentally regulated genes. The complementary DNA (cDNA) for the cyclic AMP receptor has now been cloned from λgt-11 libraries by screening with specific antiserum. The 2-kilobase messenger RNA (mRNA) that encodes the receptor is undetectable in growing cells, rises to a maximum at 3 to 4 hours of development, and then declines. In vitro transcribed complementary RNA, when hybridized to cellular mRNA, specifically arrests in vitro translation of the receptor polypeptide. When the cDNA is expressed in *Dictyostelium* cells, the undifferentiated cells specifically bind cyclic AMP. Cell lines transformed with a vector that expresses complementary mRNA (antisense) do not express the cyclic AMP receptor protein. These cells fail to enter the aggregation stage of development during starvation, whereas control and wild-type cells aggregate and complete the developmental program within 24 hours. The phenotype of the antisense transformants suggests that the cyclic AMP receptor is essential for development. The deduced amino acid sequence of the receptor reveals a high percentage of hydrophobic residues grouped in seven domains, similar to the rhodopsins and other receptors believed to interact with G proteins. It shares amino acid sequence identity and is immunologically cross-reactive with bovine rhodopsin. A model is proposed in which the cyclic AMP receptor crosses the bilayer seven times with a serine-rich cytoplasmic carboxyl terminus, the proposed site of ligand-induced receptor phosphorylation.

UPON STARVATION, *Dictyostelium* AMOEBAE SYNCHRONOUSly aggregate into a multicellular organism and cells in specific positions differentiate into stalk or spore cells (*1, 2*). The aggregation of individual cells into the multicellular structure is coordinated by a developmentally regulated cyclic AMP signaling system. At aggregation centers, cyclic AMP is periodically synthesized and secreted at 6-minute intervals. The oscillations in cyclic AMP initiate chemical waves that are propagated through the cell monolayer. The leading edge of each of these waves provides a gradient that directs the migration of chemotactically responsive cells into the aggregation center (*3*). The oscillations are required for early gene expression as well as proper morphogenesis. Mutants with defective oscillators fail to differentiate, but can be rescued by periodic application of exogenous cyclic AMP. A constant level of extracellular cyclic AMP does not rescue these mutants and suppresses early gene expression in wild-type cells. Late gene expression requires continuous extracellular cyclic AMP. Pharmacologic studies indicate that the effects of cyclic AMP on both early and late gene expression are mediated via the chemoattractant receptor (*2, 4*).

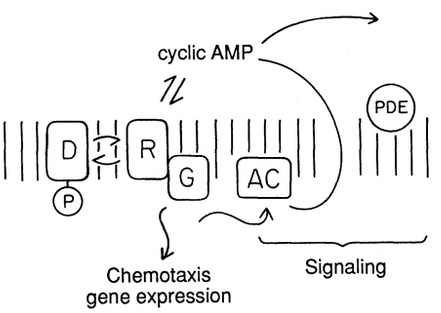


Fig. 1. Diagram of *Dictyostelium* cyclic AMP signaling system. Interactions among receptor (R), modified receptor (D), guanine nucleotide binding protein or proteins (G), adenylate cyclase (AC), and phosphodiesterase (PDE).

The cyclic AMP oscillator includes the surface receptor, adenylate cyclase, and cyclic AMP phosphodiesterase (Fig. 1). Binding to the surface receptor triggers activation of adenylate cyclase and newly synthesized cyclic AMP is secreted, completing a positive feedback loop that amplifies the initial stimulus and relays the chemotactic signal to other cells. Within a few minutes, the response ceases as the cells adapt to persistent stimulation. Adaptation is reversible; when the stimulus is removed the cells resensitize, and within a few minutes the next cycle is triggered.

The cyclic AMP receptor, identified by photoaffinity labeling (*5*), oscillates between two interconvertible forms designated R (40 kilodaltons) and D (43 kilodaltons) in parallel with the oscillations in cyclic AMP synthesis (*6*). Stimulation with cyclic AMP converts the R form to the D form. The kinetics and dose-dependence of this conversion are similar to those of adaptation (*7*). Adaptation and the shift in electrophoretic mobility are closely correlated with a fivefold increase in phosphorylation of the receptor, suggesting that phosphorylation of the receptor plays a central role in adaptation (*8*). Receptor phosphorylation has also been correlated with deactivation of rhodopsin and the β-adrenergic receptors (*9*).

Recent evidence, including in vitro stimulation of adenylate

P. S. Klein, T. J. Sun, R. L. Johnson, and P. N. Devreotes are in the Department of Biological Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205; and C. L. Saxe III and A. R. Kimmel are in the Laboratory of Cellular and Developmental Biology, NIDDK (6/B1-12), National Institutes of Health, Bethesda, MD 20892.

*To whom correspondence should be addressed.

cyclase with guanosine triphosphate (GTP), cyclic AMP–stimulated binding of GTP to membranes, and modulation of cyclic AMP binding affinity by GTP and guanosine diphosphate (GDP), suggests that the cyclic AMP receptor is coupled to a GTP-dependent regulatory protein analogous to mammalian G proteins and transducin (*1, 10, 11*). The cDNA's for two G protein α subunits and a β subunit have been cloned from *Dictyostelium* and show considerable identity to mammalian G protein subunits (*12, 13*).

Thus, the chemotactic signaling system in *Dictyostelium* appears to be a progenitor of vertebrate signaling systems. The opsins, the adrenergic receptors, and other receptors coupled to G proteins contain multiple hydrophobic regions, and it has been suggested that they traverse the membrane seven times (*14*). We present the amino acid sequence of the cyclic AMP receptor and propose that it and other eukaryotic chemoattractant receptors are structurally related to the growing class of G protein–linked receptors. We also provide evidence that the receptor plays a central role in the developmental program of *Dictyostelium*.

Isolation of receptor cDNA. We have purified the cyclic AMP receptor and raised, in rabbits, a monospecific antiserum that reacts with both the R and D forms (*15, 16*) (Fig. 2). We screened two λgtd11 cDNA libraries with the antiserum and obtained seven cDNA clones. These clones produced β-galactosidase fusion proteins (118 kD to 150 kD) that reacted strongly with antiserum in immunoblots (Fig. 2). A library prepared from late aggregation stage NC-4 cells yielded two clones in an initial screen of 60,000. The second library, prepared from 3-hour stage AX-3 cells, yielded five clones in an initial screen of 160,000. The AX-3 library was again screened with cDNA probes to obtain 16 additional clones, some of which included the entire coding sequence.

That the cDNA clones encode the cyclic AMP receptor was verified by multiple criteria. Antibodies affinity purified from several of the fusion proteins specifically stained the R and D forms of the receptor in immunoblots of crude membranes. Antibodies affinity purified from two small fusion proteins (A and C) representing nonoverlapping amino acid sequences did not cross-react with each other in immunoblots, but did react with the receptor and a large fusion protein (B) that encompassed both of the shorter sequences, showing that the B cDNA encodes multiple receptor epitopes (Fig. 2). Furthermore, each of these fusion proteins (A, B, and C) generated antiserum (in rabbits) that specifically recognizes the cyclic AMP receptor in cell extracts.

In vitro translation of mRNA from early aggregation stage cells and subsequent immunoprecipitation of the expressed polypeptide has established characteristics of the mRNA encoding the cyclic AMP receptor. (i) The receptor mRNA encodes a nascent 37-kD polypeptide (*16*); (ii) the receptor mRNA is undetectable in growing cells, is maximally expressed 3 to 4 hours after starvation, and then declines (*16*); (iii) the receptor mRNA is approximately 2 kb, as determined by in vitro translation of size-fractionated poly(A)$^+$ RNA. Each of these characteristics was confirmed with the cDNA clones. Thus, RNA blot analysis with a cDNA probe identifies a 2-kb band absent in vegetative cells (0 hours), rising to a maximum at 4 hours of development, and then declining (Fig. 3A). Furthermore, RNA transcribed in vitro from a cDNA encompassing the entire coding region (cDNA 6B) translates into a 37-kD polypeptide, and complementary RNA generated from cDNA 6B arrests translation of the cyclic AMP receptor mRNA when hybridized to cellular RNA (Fig. 3B).

To test whether the cDNA encoded a cyclic AMP binding

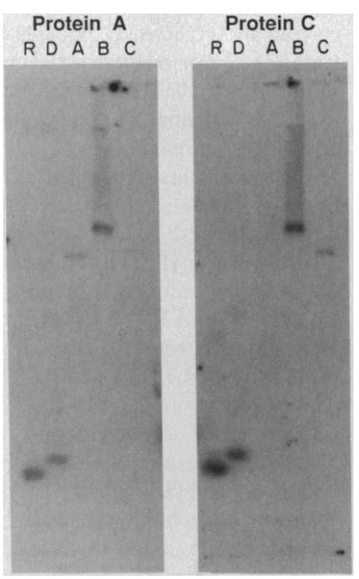


Fig. 2. Isolation of cyclic AMP receptor cDNA and affinity purification from fusion proteins. Immunoblot analysis with antibodies affinity-purified from fusion protein A (left) or fusion protein C (right). The five lanes in each panel are: Membranes from aggregation stage cells before (lane R) or after stimulation with cyclic AMP (lane D) and β-galactosidase fusion proteins encoded by cDNA's for A (lane A), for B (lane B), or for C (lane C). Two λgt11 cDNA libraries, representing early and late aggregation stage cells, were screened with antiserum to the cyclic AMP receptor (*42*). Positive phage clones were purified and inserts were subcloned. A 5′ fragment of one of the larger cDNA's was used as a probe to again screen the AX-3 library. A 1.3-kb insert (6B), which was used as a probe in several experiments described below, was cloned (3′–5′) into Bluescript KS$^-$ and was designated p6B. β-Galactosidase fusion proteins were prepared by infection of CAG 456 cells with plaque purified phage (*43*). Lysates containing approximately 100 μg of each fusion protein, were separated on preparative SDS-PAGE gels, transferred to nitrocellulose filters, and incubated with the antiserum to the receptor. The stained band was excised, and the antibodies were eluted in 0.1*M* glycine, *p*H 3.0, as described (*44*).

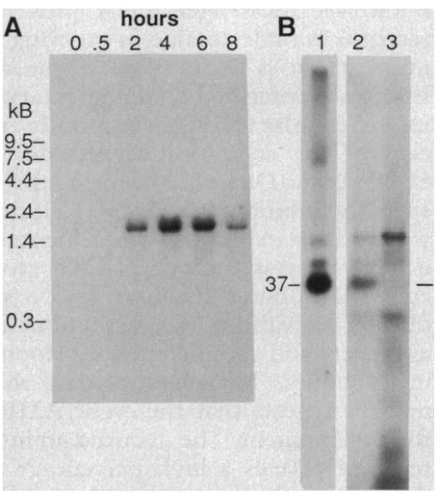


Fig. 3. In vitro synthesis of cyclic AMP receptors and analysis of receptor mRNA. (**A**) Developmental regulation. RNA isolated from AX-3 cells at the indicated hours during development was subjected to RNA blot analysis with a ^{32}P-labeled probe generated from p6B. (**B**) In vitro translation and hybrid arrest: RNA transcribed in vitro from p6B was translated in reticulocyte lysates in the presence of [^{35}S]methionine (lane 1). Cellular RNA isolated from cells at the 3-hour stage was similarly translated either in the absence (lane 2) or presence (lane 3) of in vitro transcribed cRNA. The reaction products were immunoprecipitated with the antiserum receptor, subjected to electrophoresis on SDS-PAGE and fluorography for 2 hours (lane 1) or 15 hours (lanes 2 and 3). The position of the nascent receptor is indicated with a bar at 37 kD. Development was initiated by transfer of AX-3 cells from nutrient-rich (HL-5) medium to development buffer (*16*). RNA was isolated after lysis of cells in 6*M* guanidine HCl with subsequent precipitation by ethanol (*45*). Total RNA (10 μg) from each time point was subjected to electrophoresis in formaldehyde gels (*46*) and transferred to GeneScreen Plus. The Eco RI fragment of p6B was purified and labeled (*47*) to 10^9 cpm/μg. Hybridization conditions were as described (GeneScreen protocol). A parallel filter was hybridized with a cDNA probe for a constitutive mRNA, which showed equivalent intensity bands for each time point. RNA was transcribed in vitro with either T3 or T7 RNA polymerase (*48*) (Stratagene protocol was used). The template p6B was linearized at either the 3′ (for coding strand) or 5′ end (for noncoding strand). In vitro translations were carried out with nuclease-treated reticulocyte lysate (Promega) in a final volume of 50 μl containing 50 μCi of [^{35}S]methionine (1100 Ci/mmol, NEN) for 2 hours. The translation mixture was immunoprecipitated with receptor antiserum, and analyzed by electrophoresis and fluorography (*16*). In vitro transcribed RNA was present at 0.02 μg/μl and cellular RNA was present at 1 μg/μl. For hybrid arrest of translation, 50 μg of cellular RNA was incubated with 1 μg of in vitro transcribed RNA corresponding to the noncoding strand (T7 promoter) in 14 μl of water at 50°C for 15 minutes before in vitro translation.

Table 1. Expression of cDNA 6B in growing cells. pB6 (p6B inverted) was linearized with RV, and Bgl II linkers (CAGATCTG) were added. The insert was excised with Bam HI–Bgl II and cloned (5′ to 3′) into the Bgl II site of plasmid pBS18. After transformation of AX-3 cells with construct, pBS18B6, single colonies were screened by immunoblotting for those producing maximal amounts of the 40-kD receptor polypeptide. All experiments shown were carried out with a single cell line. The transformation was repeated two additional times with similar results. Shown is the specific binding of ^{3}H-labeled cyclic AMP expressed as 10^3 sites per cell. Ammonium sulfate binding assays (*34*) were carried out on cells at the 0-hour stage. Several replicate measurements were performed in each of four independent experiments.

experiment	Wild-type	pBS18	pBS18B6
I	15	18	475, 555, 275
II	15	N.D.	115
III	12	22	437, 810
IV	14	20	790

protein, we constructed a cell line that expressed the gene product of cDNA 6B in undifferentiated cells. Wild-type cells at this stage display few cyclic AMP binding sites and express very low levels of receptor mRNA and 40-kD protein (Fig. 3A). Complementary DNA 6B was cloned into a vector that uses the actin 15 promoter to direct RNA synthesis (*17*). The actin 15 promoter is maximally active in "0-hour" stage cells (*17*). At the 0-hour stage, the cells transformed with the construct containing cDNA 6B expressed a 40-kD protein as detected with receptor antiserum, while those transformed with the vector alone did not. Table 1 shows binding of ^{3}H-labeled cyclic AMP to the wild-type and transformed cell lines. The wild-type and the control transformed cell lines display only 1.5×10^4 sites per cell at this stage. Those transformed with the construct containing cDNA 6B display 1×10^5 to 8×10^5 sites per cell (*18*).

The cyclic AMP receptor controls development. To explore the role of the cyclic AMP receptor in morphogenesis and development, we constructed a cell line that failed to express receptors. Cells were transformed with a construct designed to express RNA complementary to the entire coding region of the cyclic AMP receptor mRNA. The cDNA 6B was cloned into a plasmid vector such that the noncoding strand would be synthesized as a fusion to the neomycin phosphotransferase mRNA (*19*). The noncoding strand of 6B arrests in vitro translation of endogenous receptor mRNA when hybridized to cellular mRNA (Fig. 3B). A control set of cells was transformed with the vector lacking receptor sequences. During starvation of the cells, the antisense transformants remained as a cell monolayer, without aggregating; the control cells aggregated and formed streams in parallel with wild-type cells (Fig. 4A). The cyclic AMP receptor was undetectable in the antisense transformants, as determined by immunoblot analysis (Fig. 4B). Furthermore, in the antisense transformants the endogenous 2-kb mRNA for the receptor was undetectable by RNA blot analysis and antisense transcripts were detectable (*20*). These data show that the receptor is required for aggregation and for cells to progress through the developmental program.

Sequence analysis. The nucleotide sequence of cDNA encompassing the entire coding region is shown in Fig. 5. The largest open reading frame encodes a polypeptide of 392 amino acids with a calculated molecular size of 44,243 daltons. The NH_2-terminal two-thirds of the coding region is enriched in hydrophobic residues in clusters of 21 to 25 amino acids, while the last 130 amino acids are predominantly hydrophilic and are particularly enriched in serines and threonines. A single consensus sequence for N-linked glycosylation is present near the NH_2-terminus (*21*). The higher mobility of the nascent polypeptide (37 kD) relative to the mature R form (40 kD) suggests a cotranslational modification step such as glycosylation (*16*). However, the electrophoretic mobility of the receptor in polyacrylamide gels is not altered by treatment of the partially purified receptor with endoglycosidases or by inhibitors of glycosylation, including tunicamycin [even at doses in excess of that needed to inhibit glycosylation of other *Dictyostelium* glycoproteins (*22*)]. Both the nascent and mature polypeptides migrate more rapidly than would be expected from the deduced sequence (44,243 daltons), as observed for other highly hydrophobic membrane proteins (*23*).

The hydropathy profile of the cyclic AMP receptor determined by the method of Kyte and Doolittle (*24*) reveals six strongly hydrophobic regions (M1 to M6) and a seventh domain (M7), which is less hydrophobic (Fig. 6). The pattern is remarkably similar to the profile of bovine rhodopsin, the β-adrenergic receptor, and several other recently identified receptors, although the lengths of the loops and the COOH-terminal regions vary widely (*14, 25–28*). For rhodopsin, assignment of the hydrophobic regions as membrane-spanning domains is supported by extensive biochemical data, including accessibility to proteases, peptide antibodies, and lipophilic modifying reagents (*29*). The cyclic AMP receptor and bovine rhodopsin share 22 percent amino acid identity over 270 residues in the NH_2-terminal regions of the two receptors (Fig. 7), and the percentage of matched residues increases to 32 percent with conservative replacements. The identities are distributed in the hydrophobic and hydrophilic domains, and the optimal alignment matches each of the putative transmembrane domains and the hydrophilic

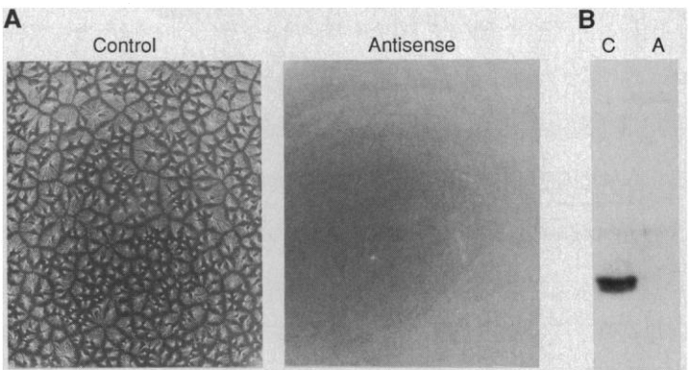


Fig. 4. Antisense transformation mimics receptor mutation. (**A**) Morphology of transformed cells. AX-3 cells were transformed with the control vector (*19*) or with the antisense construct described below. Under starvation conditions, the control transformants behaved as wild-type cells, aggregating with streams of cells entering the aggregation centers. Each of four independent clones of antisense transformants remained as a uniform cell monolayer. The phenotype of one of these clones and the control of the 7.5-hour stage of development are shown. After 36 hours, the antisense transformants began to show weak signs of morphogenesis. (**B**) Immunoblot analysis. Cells starved in suspension were removed at the 6-hour stage of starvation and subjected to immunoblot analysis with the antiserum to the cyclic AMP receptor as described in the legend to Fig. 2. The receptor was expressed in the control transformants (**C**) at the same level as wild-type cells but was undetectable in the antisense transformants (A). A Bam HI–Sal I digest of p6B was subcloned into pA6NPTII digested with Bam HI and Sal I; the Bam HI site in the polylinker of p6B is 13 bp beyond the 3′ end of the cDNA insert and the Sal I site is 26 bp beyond the 5′ end. Thus, the resulting construct includes the entire 6B cDNA (with 39 additional bases from the Bluescript polylinker) fused, in the 3′ to 5′ orientation, to the neomycin phosphotransferase gene under control of the actin 6 promoter. AX-3 cells were transformed (*19*) and selected in G418 at 20 μg/ml. Clonal cells (each from a separate transformed dish) were isolated from low density (20 to 40 colony-forming units per 10-cm dish) bacterial plates when the diameter of the plaque was about 1 mm and maintained in HL5 with G418 at 20 μg/ml. Cells were allowed to differentiate at 5×10^7 cells per milliliter in suspension or at 5×10^7 cells per 10-cm dish.

joining regions in order, suggesting conserved structural features that extend beyond the hydrophobic character of the two proteins. For example, the hydrophilic regions joining the first and second and the fifth and sixth hydrophobic domains share several identities and both are especially enriched in positively charged side chains. Although the identities do not persist in the hydrophilic COOH-terminal regions of the two receptors, both have an abundance of serines and threonines.

There are also identities with other receptors in this family, such as the α- and β-adrenergic, muscarinic acetylcholine, and substance K receptors (*14, 27, 28*). Certain groups of residues distributed throughout the sequences appear to be highly conserved in many of these receptors (Fig. 7). Domain M6 of many of these receptors contains a cluster of identities within the sequence FXXCWXPO (*30*), where X represents an aliphatic residue and O represents phenylalanine or tyrosine. The identities between the cyclic AMP receptor and bovine rhodopsin are particularly numerous here: five out of nine residues are identical, and eight out of nine match with conservative replacements. There appears to be no sequence homology to other known cyclic AMP–binding proteins such as the catabolite activator protein (CAP) of *Escherichia coli* or the regulatory subunit of cyclic AMP–dependent protein kinase from *Dictyostelium* or from mammals (*31, 32*).

The structural relation of the cyclic AMP receptor to bovine rhodopsin is supported by immunological data. Antiserum to bovine rhodopsin recognizes the cyclic AMP receptor in immunoblots of crude membranes from aggregation stage *Dictyostelium*, in which the receptor is 0.1 percent of the total protein (Fig. 7, inset). While the intensity of staining is low compared to the signal for rhodopsin, the major band is clearly evident at 40 kD (R form) and shifts to 43 kD (D form) on stimulation with cyclic AMP (Fig. 2). The cross-reactive epitope or epitopes could be one of several regions of concentrated sequence identity (Fig. 7), for example, residues 95 to 112 for which six of eight residues are identical, or residues 227 to 235, discussed above.

Model of the cyclic AMP receptor. On the basis of the hydropathy analysis and the functional and structural analogy to bovine rhodopsin, we propose a working model for the structure of the cyclic AMP receptor (Fig. 8). In this model, there are seven transmembrane domains, the NH_2-terminus is extracellular, and the serine-rich COOH-terminal tail is intracellular and is the site of ligand-induced phosphorylation.

In many of the receptors proposed to cross the membrane seven times, including the cyclic AMP receptor, the seventh domains are considerably less hydrophobic than would be predicted for a typical transmembrane domain (*24*). For the rhodopsins, this is partly due to the lysine, which covalently binds retinal. In the cyclic AMP receptor, the strongly polar residues of the seventh domain are aligned on one face of the proposed α helix (average hydropathy value = −0.9), the opposing face predominantly hydrophobic with an average hydropathy value of +2.0 (*33*). These polar residues, in a manner analogous to the lysine of rhodopsin, could be involved in

```
1          TCT CTC TCT CTC TCT AAT ATC AAA TCA AAA CGA TAA ATC GAA   42

43  TAA ATT TAA GAT TTT TCA CAC ATA TAT ATA TAT AAA TAA AAA AAT AAA   90

                                    *
91  ATG GGT CTT TTA GAT GGA AAT CCA GCC AAT GAA ACA TCA TTG GTT TTA  138
    MET Gly Leu Leu Asp Gly Asn Pro Ala Asn Glu Thr Ser Leu Val Leu

139 TTA TTA TTT GCC GAT TTT TCT TCA ATG TTG GGT TGT ATG GCA GTG TTG  186
    Leu Leu Phe Ala Asp Phe Ser Ser MET Leu Gly Cys MET Ala Val Leu

187 ATT GGT TTT TGG AGA TTA AAA CTT TTG CGT AAT CAT GTT ACA AAA GTA  234
    Ile Gly Phe Trp Arg Leu Lys Leu Leu Arg Asn His Val Thr Lys Val

235 ATT GCA TGT TTT TGT GCT ACC TCC TTT TGT AAA GAT TTT CCT TCA ACA  282
    Ile Ala Cys Phe Cys Ala Thr Ser Phe Cys Lys Asp Phe Pro Ser Thr

283 ATT TTA ACT TTA ACA AAT ACA GCT GTC AAT GGT GGT TTC CCT TGT TAT  330
    Ile Leu Thr Leu Thr Asn Thr Ala Val Asn Gly Gly Phe Pro Cys Tyr

331 TTA TAT GCA ATT GTA ATT ACA TAT GGT AGT TTT GCA TGT TGG TTG TGG  378
    Leu Tyr Ala Ile Val Ile Thr Tyr Gly Ser Phe Ala Cys Trp Leu Trp

379 ACT TTA TGT CTT GCA ATT AGT ATT TAT ATG TTA ATT GTA AAA AGA GAA  426
    Thr Leu Cys Leu Ala Ile Ser Ile Tyr MET Leu Ile Val Lys Arg Glu

427 CCA GAA CCA GAA AGA TTT GAA AAA TAT TAT TAT TTA TTA TGT TGG GGT  474
    Pro Glu Pro Glu Arg Phe Glu Lys Tyr Tyr Tyr Leu Leu Cys Trp Gly

475 TTA CCA TTA ATT TCA ACA ATT GTT ATG TTG GCA AAA AAT ACA GTA CAA  522
    Leu Pro Leu Ile Ser Thr Ile Val MET Leu Ala Lys Asn Thr Val Gln

523 TTT GTT GGT AAT TGG TGT TGG ATT GGT GTA AGT TTC ACT GGA TAT AGA  570
    Phe Val Gly Asn Trp Cys Trp Ile Gly Val Ser Phe Thr Gly Tyr Arg

571 TTT GGG TTA TTC TAT GGT CCA TTC TTA TTT ATT TGG GCA ATT TCA GCA  618
    Phe Gly Leu Phe Tyr Gly Pro Phe Leu Phe Ile Trp Ala Ile Ser Ala

619 GTA TTG GTT GGT TTA ACA TCA CGT TAT ACC TAT GTT GTT ATT CAT AAT  666
    Val Leu Val Gly Leu Thr Ser Arg Tyr Thr Tyr Val Val Ile His Asn

667 GGT GTT AGT GAT AAT AAA GAG AAA CAT TTA ACA TAC CAA TTC AAA TTG  714
    Gly Val Ser Asp Asn Lys Glu Lys His Leu Thr Tyr Gln Phe Lys Leu

715 ATC AAC TAT ATC ATA GTT TTC CTT GTT TGT TGG GTA TTT GCA GTT GTT  762
    Ile Asn Tyr Ile Ile Val Phe Leu Val Cys Trp Val Phe Ala Val Val

763 AAT CGT ATT GTA AAT GGT TTA AAT ATG TTT CCA CCA GCA CTC AAT ATT  810
    Asn Arg Ile Val Asn Gly Leu Asn MET Phe Pro Pro Ala Leu Asn Ile

811 CTC CAC ACC TAT TTG AGT GTA TCC CAT GGT TTT TGG GCA TCT GTC ACA  858
    Leu His Thr Tyr Leu Ser Val Ser His Gly Phe Trp Ala Ser Val Thr

859 TTT ATC TAT AAT AAT CCA TTA ATG TGG CGT TAT TTT GGT GCA AAG ATT  906
    Phe Ile Tyr Asn Asn Pro Leu MET Trp Arg Tyr Phe Gly Ala Lys Ile

907 TTA ACA GTT TTC ACA TTC TTT GGT TAC TTT ACT GAT GTA CAA AAG AAA  954
    Leu Thr Val Phe Thr Phe Phe Gly Tyr Phe Thr Asp Val Gln Lys Lys

955 TTG GAA AAG AAT AAA AAC AAT AAC AAT CCA TCA CCA TAT AGT TCA AGT 1002
    Leu Glu Lys Asn Lys Asn Asn Asn Asn Pro Ser Pro Tyr Ser Ser Ser

1003 CGT GGT ACT TCT GGT AAA ACT ATG GGT GGT CAT CCA ACT GGT GAT GAT 1050
     Arg Gly Thr Ser Gly Lys Thr MET Gly Gly His Pro Thr Gly Asp Asp

1051 GTG CAA TGT AGT TCC GAT ATG GAG CAA TGT TCT TTG GAA AGA CAT CCA 1098
     Val Gln Cys Ser Ser Asp MET Glu Gln Cys Ser Leu Glu Arg His Pro

1099 AAT ATG GTA AAT AAT CAA CAA AAT CTA AAT AAT AAT TAT GGT TTA CAA 1146
     Asn MET Val Asn Asn Gln Gln Asn Leu Asn Asn Asn Tyr Gly Leu Gln

1147 CAA AAT TAT AAT GAT GAA GGT TCA TCT TCA TCT TCT TTA TCA TCA TCT 1194
     Gln Asn Tyr Asn Asp Glu Gly Ser Ser Ser Ser Ser Leu Ser Ser Ser

1195 GAT GAA GAA AAA CAG ACT GTT GAA ATG CAA AAT ATT CAA ATT TCC ACT 1242
     Asp Glu Glu Lys Gln Thr Val Glu MET Gln Asn Ile Gln Ile Ser Thr

1243 TCA ACA AAT GGT CAA GGA AAT AAT TGA TTA TAA AAA AAA AAA GAA AAA 1290
     Ser Thr Asn Gly Gln Gly Asn Asn

1291 AAA AAA AAA AAA TTA AAA AAC C                                  1312
```

Fig. 5. Nucleotide and deduced amino acid sequence of the cyclic AMP receptor. The sequence of cDNA 5A, which encompasses the entire coding region, was determined from both strands. In addition, nine independent, overlapping cDNA's that comprised the entire coding region were partially sequenced. The 1.3-kb sequence shown was identical in the multiple, independent cDNA's. In the deduced amino acid sequence, the hydrophobic domains are underlined and numbered M1 to M7. The asterisk (*) indicates the single consensus site for N-linked glycosylation, although the receptor has not been shown to be glycosylated. Exonuclease deletions of various lengths were prepared (*49*) from cDNA 5A, which had been subcloned into Bluescript as described in the legend to Fig. 2. DNA sequence was determined by means of the dideoxy-nucleotide chain termination method (*50*), with modified T7 DNA polymerase (Sequenase, United States Biochemical Corp.). Several cDNA's were obtained from the two libraries that were identical in the 1.3-kb region shown above but diverged in either the 5′ or 3′ ends. These divergent sequences did not hybridize to the receptor mRNA in blots or to genomic restriction fragments that included the coding region. The origin of these terminal sequences is unclear.

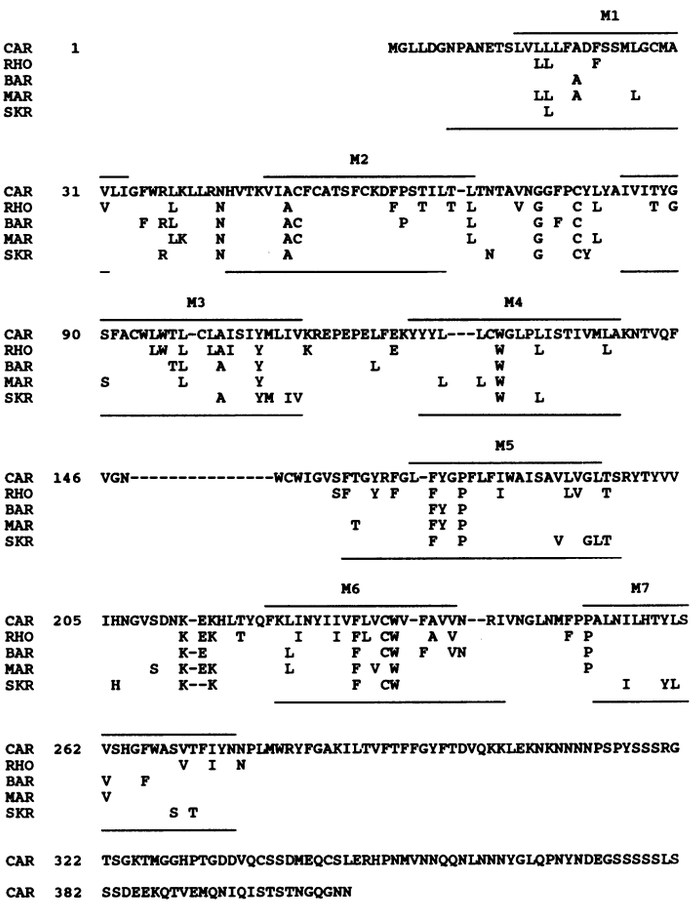
the binding of cyclic AMP. On the basis of binding studies with cyclic AMP analogs, Van Haastert and Kien have proposed that activation of the cyclic AMP receptor involves the formation of a covalent bond between a nucleophilic residue of the receptor and the phosphate of the cyclic AMP molecule (34); the polar face of this proposed helix could offer one of several nucleophiles as well as a positive charge (histidine) to balance the resulting negative charge on the phosphoryl oxygen.

The cyclic AMP receptor also has two adjacent, charged residues (Lys^{59} and Asp^{60}) in the second hydrophobic domain and an aspartic acid in the first domain. Pairs of oppositely charged residues occur in domain M3 of all of the opsin sequences, and there is a highly conserved aspartic acid within the second transmembrane domain of many of the opsins, the β-adrenergic, muscarinic, and substance K receptors (*14, 25–28, 35*). These charged residues occur in contexts of predominantly hydrophobic residues, significantly lowering the average hydropathy value of these domains (*24*). In the opsins, intramembrane charged residues have been proposed to interact with the retinal moiety (*26, 35*) and in the β-adrenergic receptor, they have been proposed to play a role in ligand binding (*36*). The charged residues within the hydrophobic domains of the cyclic AMP receptor could similarly interact with cyclic AMP in the binding site. Thus, the seven transmembrane domains could be arranged like the staves of barrel, forming a central core cradling the ligand binding site, as has been proposed for rhodopsin and the β-adrenergic receptor (*26, 35*), with charged amino acid side chains facing into the binding site.

The last 132 amino acids of the cyclic AMP receptor have 18 serine and 6 threonine residues. In our model, this cytoplasmic tail is the site of stimulus dependent phosphorylation, which occurs predominantly on serine (*16*). A particularly interesting stretch of serines lies from residue 360 to 368, with eight serines interrupted by one leucine, and bounded on each side by two to three acidic residues. Since the cyclic AMP receptor appears to be multiply phosphorylated in response to stimulus, these serines may represent the specific site of stimulus dependent phosphorylation.

Novel roles for G protein–linked receptors. Although many of the genes required for prokaryotic chemotaxis have been identified (*37*), the biochemical mechanisms of eukaryotic chemotaxis are less well understood. On the basis of the primary sequence presented, we propose a model for the cyclic AMP receptor, which structurally relates a chemoattractant receptor to other G protein–linked receptors. Since chemotaxis in *Dictyostelium* is a primitive sensory process involved in cell-cell interactions, it is not unreasonable to suppose that the cyclic AMP receptor and those involved in sensory and endocrine systems of higher organisms evolved from a common ancestor. Furthermore, structural similarities are likely to be found among other eukaryotic chemoattractant receptors when their primary structures are determined.

The cyclic AMP chemoattractant receptor is believed to play a central role in the differentiation of *Dictyostelium*, regulating the expression of genes in both early and late developmental stages. Recent studies with mutants defective in signal transduction have



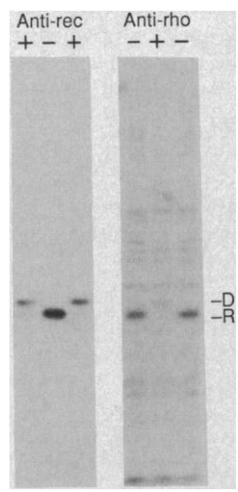


Fig. 7. Homology to bovine rhodopsin and other G protein–linked receptors. The amino acid sequences of the cyclic AMP receptor (CAR) and bovine rhodopsin (RHO) were aligned with the DFASTP program (*51*). A 15-residue gap was introduced in the sequence of the cyclic AMP receptor (between residues 148 and 149) before running the alignment program. Statistical significance was assessed with the use of the RDF program with 100 random shuffles of either sequence and gave *Z* values of 14.2 and 14.6 (*51*). The solid line above the cyclic AMP receptor sequence indicates the proposed transmembrane domains, labeled M1 to M7 (see Figs. 5 and 6). The hamster β-adrenergic receptor (BAR), the porcine cerebral muscarinic acetylcholine receptor (MAR), and the substance K receptor (SKR) were also aligned with bovine rhodopsin and then compared with the cyclic AMP receptor sequence by inspection. Identities with the cyclic AMP receptor are indicated below the sequence. The solid line beneath these sequences indicates the transmembrane domains of bovine rhodopsin as proposed by Hargrave (*26, 30*). (Right) Immunoblot with antiserum to receptor (anti-rec) or antiserum to rhodopsin (anti-rho). Membranes from aggregation stage AX-3 cells were prepared from unstimulated (R) and cyclic AMP stimulated (D) cells and analyzed by immunoblot (*16*) with anti-rec at 1:500 dilution (left) or with antiserum to bovine rhodopsin at 1:150 dilution (*52*). The cyclic AMP receptor represents 0.1 percent of the total protein in these membranes. The immunoblot was visualized with ^{125}I-labeled protein A, and film was exposed for 5 hours (left) or 150 hours (right).

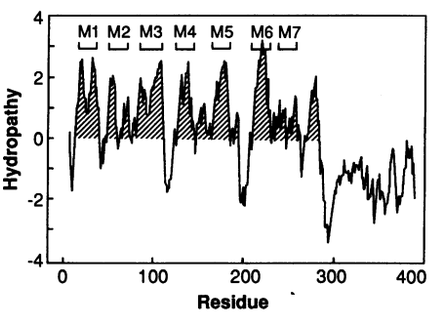


Fig. 6. Hydropathy analysis of deduced amino acid sequence. The deduced amino acid sequence was analyzed with the Kyte and Doolittle hydropathy algorithm; windows of 9 and 19 residues were used (*24*). The 9-residue window is plotted. The assignment of the transmembrane domains is based on the average hydropathy from a 19-residue window, as well as analogies with bovine rhodopsin, as discussed in the text; the domains are indicated by bars labeled M1–7 and are also delineated in Fig. 5.

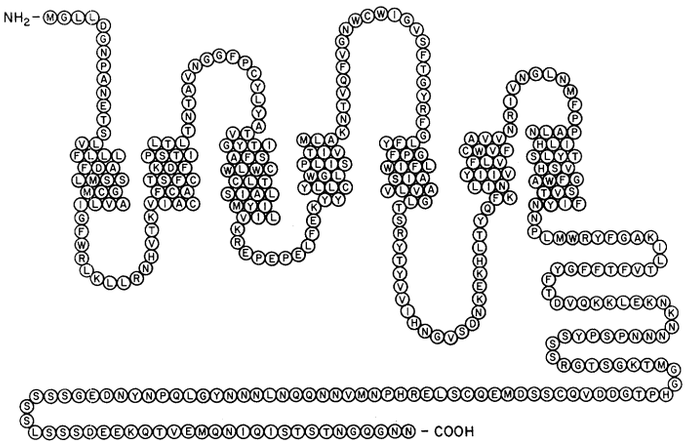


Fig. 8. Model of the cyclic AMP receptor. The seven hydrophobic domains are arranged as α helices in the lipid bilayer. The proposed extracellular domains are shown above the α helices and the intracellular domains, including the long COOH-terminus, are shown below these transmembrane helices.

suggested that the effect of cyclic AMP on early gene expression is mediated through a G protein–coupled receptor (*38*). The control of gene expression in a developing organism is a novel role for G protein–coupled receptors. This could be a general theme in development: embryonic inducers doubling as chemoattractants and acting through G protein–coupled receptors. In this way, short-term stimuli could be integrated into long-term changes in the organism. Similar mechanisms may also play a role in neural plasticity. In *Aplysia*, a mechanism has been proposed for long-term information storage whereby transient stimuli, correlated with G protein–mediated activation of adenylate cyclase, effect long-term changes in synaptic junctions that are dependent on gene expression (*39*).

The recent advances in the transformation of *Dictyostelium* have expanded the possibilities of genetic manipulation of these relatively simple eukaryotic cells (*19, 40, 41*). In our experiments, transformation of *Dictyostelium* with an antisense construct of the cyclic AMP receptor cDNA has yielded cells that are severely impaired in the aggregation process. These data now raise the question of whether the disruption of receptor expression blocks development by inhibiting the expression of cyclic AMP–regulated genes required for early development or whether the failure to develop is because the cells are unable to respond to a chemotactic stimulus. Disruption of the cyclic AMP receptor, by antisense transformation and by homologous recombination, as well as replacement with receptor genes carrying defined mutations, can now be used to further elucidate the role of the receptor in both chemotaxis and development.

REFERENCES AND NOTES

1. P. M. W. Janssens and P. J. M. van Haastert, *Microbiol. Rev.* **51**, 396 (1987).
2. G. Gerisch, *Annu. Rev. Biochem.* **56**, 853 (1987).
3. K. J. Tomchik and P. N. Devreotes, *Science* **212**, 443 (1981).
4. P. Schaap and R. van Driel, *Exp. Cell Res.* **159**, 388 (1985); A. Kimmel, *Dev. Biol.* **122**, 163 (1987); M. Oyama and D. D. Blumberg, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4819 (1986); B. Haribabu and R. P. Dottin, *Mol. Cell Biol.* **6**, 2402 (1986); S. Mann and R. A. Firtel, *ibid.* **7**, 458 (1987).
5. M. Juliani and C. Klein, *J. Biol. Chem.* **256**, 613 (1981); A. Theibert, P. Klein, P. N. Devreotes, *ibid.* **259**, 12318 (1984).
6. P. Klein, A. Theibert, D. Fontana, P. N. Devreotes, *ibid.* **260**, 1757 (1985).
7. P. N. Devreotes and J. Sherring, *ibid.*, p. 6378.
8. P. Klein, D. Fontana, B. Knox, A. Theibert, P. N. Devreotes, *Cold Spring Harbor Symp. Quant. Biol.* **50**, 787 (1985); C. Klein, J. Lubs-Harkness, S. Simons, *J. Cell Biol.* **100**, 714 (1985); R. Vaughan and P. N. Devreotes, *J. Biol. Chem.*, in press.
9. D. R. Sibley, J. L. Benovic, M. G. Caron, R. J. Lefkowitz, *Cell* **48**, 913 (1987).
10. A. Theibert and P. N. Devreotes, *J. Biol. Chem.* **261**, 15121 (1986).
11. P. J. M. van Haastert, *Biochem. Biophys. Res. Commun.* **124**, 597 (1984).
12. M. Pupillo, G. Pitt, A. Kumagai, R. A. Firtel, P. N. Devreotes, in preparation.
13. P. Lilly and P. N. Devreotes, in preparation.
14. H. G. Dohlman, M. G. Caron, R. J. Lefkowitz, *Biochemistry* **26**, 2657 (1987).
15. P. Klein, B. Knox, J. Borleis, P. N. Devreotes, *J. Biol. Chem.* **262**, 352 (1987).
16. P. Klein, R. Vaughan, J. Borleis, P. N. Devreotes, *ibid.*, p. 358.
17. The pBS18 vector was a gift of A. Kumagai and R. Firtel.
18. R. Johnson and P. Devreotes, in preparation.
19. The pA6NPTII vector was a gift of D. Knecht. For a complete description of the vector see D. A. Knecht and W. F. Loomis, *Science* **236**, 1081 (1987).
20. T. J. Sun and P. Devreotes, in preparation.
21. S. C. Hubbard and R. J. Ivatt, *Annu. Rev. Biochem.* **50**, 555 (1981).
22. G. Bertholdt, J. Stadler, S. Bozzaro, B. Fichtner, G. Gerisch, *Cell Differ.* **16**, 187 (1985).
23. M. Mueckler *et al.*, *Science* **229**, 941 (1985).
24. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
25. N. Nakayama, A. Miyajima, K. Arai, *EMBO J.* **4**, 2643 (1985); A. C. Burkholder and L. H. Hartwell, *Nucleic Acids Res.* **13**, 8463 (1985); D. Hagen, G. McCaffrey, G. F. Sprague, Jr., *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1418 (1986).
26. P. A. Hargrave, in *The Retina*, R. Adler and D. Farber, Eds. (Academic Press, New York, 1986).
27. Y. Masu *et al.*, *Nature* **329**, 836 (1987).
28. B. K. Kobilka *et al.*, *Science* **238**, 650 (1987).
29. P. A. Hargrave, *Prog. Retinal Res.* **1**, 1 (1982).
30. Single letter 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.
31. H. Aiba, S. Fujimoto, N. Ozaki, *Nucleic Acids Res.* **10**, 1345 (1982); P. Cossart and B. Gicquel-Sanzey, *ibid.*, p. 1363.
32. R. Mutzel, M. L. Lacombe, M. N. Simon, S. de Gunzburg, M. Veron, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 6 (1987); K. Takio, S. B. Smith, E. G. Krebs, K. A. Walsh, K. Titani, *ibid.* **79**, 2544 (1982); K. Titani *et al.*, *Biochemistry* **23**, 4193 (1984).
33. A very strong amphipathic stretch of sequence also extends from G_{244} (within the loop between proposed sixth and seventh domains) to G_{265} (within the proposed seventh domain). Although we did not propose this region as a membrane spanning α helix because of the adjacent prolines, it may be of considerable interest.
34. P. J. M. van Haastert and E. Kien, *J. Biol. Chem.* **258**, 9636 (1983).
35. J. Nathans, D. Thomas, D. S. Hogness, *Science* **232**, 193 (1986).
36. C. D. Strader *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4384 (1987); R. A. F. Dixon *et al.*, *Nature* **326**, 73 (1987).
37. A. Boyd, K. Kendall, M. I. Simon, *ibid.* **301**, 623 (1983); A. F. Russo and D. E. Koshland, Jr., *Science* **220**, 1016 (1983); M. Springer, M. Goy, J. Adler, *Nature* **280**, 279 (1979).
38. F. Kesbeke, E. Snaar-Jagalska, P. J. M. van Hasstert, *J. Cell Biol.*, in press.
39. E. R. Kandel and J. H. Schwartz, *Science* **218**, 433 (1982).
40. C. Reymond, R. Gomer, M. Mehdy, R. A. Firtel, *Cell* **39**, 141 (1984).
41. A. De Lozanne and J. A. Spudich, *Science* **236**, 1086 (1987).
42. The NC-4 late aggregation library was a gift of C. H. Siu. For the early aggregation stage library, poly(A)$^+$ RNA was isolated from AX-3 cells after 3 to 4 hours of development, and the cDNA library was then prepared by Clonetech Laboratories (*43*).
43. M. Snyder, S. Elledge, D. Sweetser, R. Young, R. W. Davis, *Methods Enzymol.* **154**, 107 (1987); T. V. Huynh, R. A. Young, R. Davis, in *DNA Cloning Techniques: A Practical Approach*, D. Glover, Ed. (IRL Press, Oxford, 1984).
44. C. Weinberger, *Science* **228**, 740 (1985).
45. International Biotechnologies, Inc. Catalog vol II, 1986–87, p. 69. Guan: see IBI catalog for reference.
46. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
47. A. Fineberg and B. Vogelstein, *Anal. Biochem.* **132**, 6 (1983).
48. M. J. Chamberlin, in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, New York, 1982), vol. 15B, p. 61.
49. T. M. Roberts, R. Kacich, M. Ptashne, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 760 (1979).
50. F. Sanger, S. Nicklen, A. R. Coulsen, *ibid.*, **74**, 5463 (1977).
51. D. J. Lipman and W. R. Pearson, *Science* **227**, 1435 (1985).
52. Gift of B. Knox and G. Khorana.

53. We thank J. Nathans and D. Shortle for critical reading of the manuscript, R. Doolittle and P. Hargrave for helpful discussions. C. H. Siu for providing the NC-4 cDNA library, D. Knecht and R. Firtel for providing transformation vectors, and J. Borleis for technical assistance. Supported by NIH grant GM 34933 to P.N.D.

18 May 1988; accepted 11 August 1988