*Proc. Natl. Acad. Sci. USA*
Vol. 84, pp. 6–10, January 1987
Biochemistry

# Cloning and cDNA sequence of the regulatory subunit of cAMP-dependent protein kinase from *Dictyostelium discoideum*

**(protein evolution/λgt11 expression library)**

RUPERT MUTZEL, MARIE-LISE LACOMBE, MARIE-NOËLLE SIMON, JEAN DE GUNZBURG*, AND MICHEL VERON

Unité de Biochimie Cellulaire, Département de Biochimie et Génétique Moléculaire, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

*Communicated by Edmond H. Fischer, August 22, 1986*

**ABSTRACT** **cDNA clones encoding the regulatory subunit of the cAMP-dependent protein kinase (ATP:protein phosphotransferase, EC 2.7.1.37) from *Dictyostelium discoideum* were isolated by immunoscreening of a cDNA library constructed in the expression vector λgt11. High-affinity cAMP-binding activity was detected in extracts from bacteria lysogenized with these clones. Nucleotide sequence analysis of three overlapping clones allowed the determination of a 1195-base-pair cDNA sequence coding for the entire regulatory subunit and containing nontranslated 5′ and 3′ sequences. The open reading frame codes for a protein of 327 amino acids, with molecular weight 36,794. The regulatory subunit from *Dictyostelium* shares a high degree of homology with its mammalian counterparts, but is lacking the $NH_2$-terminal domain required for the association of regulatory subunits into dimers in other eukaryotes. On the basis of the comparison of the regulatory subunits from *Dictyostelium*, yeast, and bovine tissues, a model for the evolution of these proteins is proposed.**

Considerable information has accumulated on the biochemistry of cyclic nucleotide-dependent protein kinases, which are found in all eukaryotic cells (reviewed in refs. 1 and 2). Best known are the mammalian cAMP-dependent protein kinases, which are composed of regulatory (R) and catalytic (C) subunits. Two major types of R subunits have been found ($R_I$ and $R_{II}$). The holoenzyme is a tetramer ($R_2C_2$) which, in the presence of cAMP, dissociates into active monomeric C subunits and $R_2$ dimers. Each R subunit carries two high-affinity cAMP binding sites. The determination of the amino acid sequence of bovine $R_I$ and $R_{II}$ subunits (3, 4) has provided strong structural arguments for the organization of each subunit into distinct domains responsible for dimerization of R, interaction with C, and cAMP binding (5).

cAMP-dependent protein kinases have also been found in lower eukaryotes, including *Neurospora* and yeast (6–8). Of particular interest is the enzyme from *Dictyostelium discoideum*, since cAMP is known to play a crucial role in the expression of developmentally regulated genes in this primitive eukaryote (9, 10). In contrast to their mammalian counterparts, R subunits from *Dictyostelium* are isolated as monomers carrying only one high-affinity binding site for cAMP (11–13). Reconstitution experiments using purified R and C subunits have led to the proposal that the *Dictyostelium* holoenzyme is a dimer (RC) composed of only one R and one C subunit (13). This structure is (so far) unique, since the cAMP-dependent protein kinase from yeast, although carrying only one cAMP binding site per R subunit, was shown to be a tetramer (6).

In order to analyze its structure in more detail and to study its role in the regulation of differentiation, we have isolated cDNA clones for the R subunit from *Dictyostelium*.

## MATERIALS AND METHODS

**Materials, Phages, and Bacterial Strains.** Nitrocellulose filters were from Schleicher & Schuell. 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal), isopropyl β-D-thiogalactopyranoside (IPTG), and restriction enzymes were from Boehringer Mannheim. The replicative form (RF) of phage M13 mp19 was purchased from Pharmacia. Sequencing reactions were performed using the M13 sequencing kit from Amersham. The "Cyclone" kit used for the M13 deletions was from International Biotechnologies (New Haven, CT). $^{125}$I-labeled protein A was generously provided by N. Guiso (Institut Pasteur, Paris). Bovine serum albumin (grade V) was from Sigma. [$^{3}$H]cAMP (41 Ci/mmol; 1 Ci = 37 GBq), [α-[$^{35}$S]thio]dATP (>400 Ci/mmol), and [α-$^{32}$P]GTP (>3000 Ci/mmol) were from Amersham.

λ phage and M13 DNA were purified by mini methods as described (14–16). λgt11 phages were lysogenized into *Escherichia coli* Y1089 (17) as described by Huynh *et al.* (18).

**Screening of the λgt11 Library.** The λgt11 library used was constructed by M.-L.L., G. J. Podgorski, J. Franke, and R. H. Kessin (39). In brief, cDNA was synthesized by the method of Gubler and Hoffman (19) from poly(A)$^+$ RNA isolated from *D. discoideum* Ax3 starved for 3 hr in the presence of 1 mM cAMP. The cDNA was ligated into the *Eco*RI site of λgt11, packaged *in vitro*, and used to infect *E. coli* Y1088 (17).

Rabbit antibodies against the purified R subunit from *Dictyostelium* (13) were preabsorbed with extracts from *E. coli* BTA282(λAp3) (20) in order to eliminate crossreaction with phage- and bacteria-encoded proteins. For this, the total immune serum was diluted 1:10 in Tris-buffered saline (TBS: 10 mM Tris·HCl, pH 7.3/0.15 M NaCl) and incubated 1 hr at 4°C with an equal volume of DNase I-treated bacterial extract (50 mg of protein per ml). Cell debris was removed by centrifugation, and the procedure was repeated twice. The absorbed immune serum (diluted 1:250) detected as little as 0.5 ng of purified R subunit spotted together with 20 μg of *E. coli* protein onto a nitrocellulose filter, whereas no reaction occurred with *E. coli* protein alone.

The cDNA library was screened according to ref. 17, using strain Y1090 and 3 × 10$^4$ plaque-forming units per 100-mm plate. A positive control consisting of 1 ng of R subunit was also spotted on each filter. After blotting as described (17), the filters were successively rinsed twice, incubated with 2% bovine serum albumin (1 hr at 37°C) and then with preabsorbed anti-R immune serum (1 hr at 37°C followed by 16 hr at 4°C). Each filter was treated separately in a Petri dish containing 20 ml of the appropriate medium: after five 5-min washes with TBS containing 5% powdered skimmed milk

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: R subunit, regulatory subunit; C subunit, catalytic subunit; IPTG, isopropyl β-D-thiogalactopyranoside; kb, kilobase(s).

*Present address: Whitehead Institute for Biochemical Research, Cambridge, MA 02142.

(Regilait, Lyon, France) (21), the filters were incubated with 1 $\mu$Ci of $^{125}$I-labeled protein A at room temperature for 1 hr. They were then washed five times with 5% milk/TBS and three times with TBS alone, dried, and exposed for 16 hr to Kodak X-Omat S films for autoradiography.

**Determination of cAMP-Binding Activity in Crude Bacterial Extracts.** Crude extracts from Y1089 cells lysogenized with various λgt11 recombinants were prepared before and after *lac* induction by IPTG as described (18). cAMP-binding activity was measured in 20 mM potassium phosphate (pH 7.5) as in ref. 13, by incubating 10–25 $\mu$g of protein with 0.2 $\mu$M [$^3$H]cAMP for 1 hr at 4°C. Binding data were corrected for background radioactivity (without extract) and for nonspecific binding (0.1 mM nonradioactive cAMP present).

**Subcloning and Sequencing of cDNAs.** cDNA inserts from the λgt11 clones of interest were excised with *Eco*RI, ligated into M13 mp19 RF DNA previously cut with *Eco*RI, and transformed into *E. coli* strain JM103 according to Messing (16). For selection of complementary cDNA strands, single-stranded DNAs purified from individual recombinant M13 phages were hybridized two by two and analyzed by agarose gel electrophoresis (16). DNA sequence analysis was performed by the method of Sanger *et al.* (22) as modified by Biggin *et al.* (23), using either the full-length inserts or processive deletions of these DNAs obtained according to Dale *et al.* (24). The nucleotide sequence of a short 3′ sequence was determined according to Maxam and Gilbert (25)

## RESULTS AND DISCUSSION

**Isolation of cDNA Clones for the R Subunit.** Screening of 3.6 × 10$^5$ plaque-forming units from the λgt11 cDNA library with specific immune serum against the R subunit from *Dictyostelium* allowed the detection of 38 clones. In order to check that these clones indeed coded for the R subunit, high-affinity cAMP-binding activity was assayed in lysates of bacteria lysogenized with the clones containing the longest inserts. Binding assays were performed with 0.2 $\mu$M cAMP. Under these conditions cAMP binding to bacterial catabolite gene activator protein (if any) would not have been detected (26).

Table 1 shows that extracts from bacteria lysogenized with clones 11.2, 1.1, 2.2, and 2.1, harboring inserts 1.08–1.15 kilobases (kb) long, contained high cAMP-binding activity after induction by IPTG. In the absence of induction, cAMP binding was very low, similar to the control values. These results are exactly those expected for the expression of a LacZ/R-subunit cDNA-encoded fusion protein. Surprisingly, clones 7.4, 2.3, and 10.1, with inserts >1.2 kb, exhibited significant cAMP-binding activity (at least 10 times background) even in the absence of the *lac* inducer. Instead of increasing binding activity, addition of IPTG slightly reduced it. Notwithstanding this unusual regulation, which will be discussed below, the cAMP-binding activity in these clones clearly indicated that the corresponding inserts also coded for the R subunit. Therefore, no further characterization of the clones shown in Table 1 (e.g., by hybrid-selected translation) was considered necessary.

Table 1. cAMP-binding activity in crude extracts from *E. coli* Y1089 harboring various λgt11 lysogens

| λgt11 clone | cDNA insert, kb | cAMP bound, nmol/mg of protein | |
|---|---|---|---|
| | | − IPTG | + IPTG |
| Control | No insert | 0.07 | 0.02 |
| 11.2 | 1.08 | 0.1 | 13.3 |
| 1.1 | 1.1 | 0.25 | 19.9 |
| 2.2 | 1.12 | 0.13 | 14.2 |
| 2.1 | 1.15 | 0.14 | 14.7 |
| 7.4 | 1.22 | 2.76 | 1.1 |
| 2.3 | 1.25 | 1.27 | 0.9 |
| 10.1 | 1.4 | 1.42 | 0.9 |

Cells growing exponentially in two parallel cultures at 30°C were shifted to 42°C for 20 min to induce phage production. Then, *lac* expression was induced by 10 mM IPTG in one sample, and both cultures were further incubated for 60 min at 37°C. Cell pellets were resuspended in 20 mM potassium phosphate (pH 7.5), and binding of 0.2 $\mu$M [$^3$H]cAMP was measured in duplicate in the crude extracts previously lysed by freeze–thaw. Difference between binding activities of duplicates was <20%. cDNA insert sizes were determined after *Eco*RI digestion and agarose gel electrophoresis.

**Nucleotide Sequence of the Complete cDNA of the R Subunit.** Nucleotide sequence analysis was performed using three clones (Fig. 1). Clone 2.1 was selected for its high inducible cAMP-binding activity, whereas clones 10.1 and 2.3 were chosen for the large size of their cDNA inserts (see Table 1). Fig. 2 shows the complete nucleotide sequence of the cDNA of the R subunit, which contains an open reading frame of 981 nucleotides. Both 5′ and 3′ non-translated sequences are extremely A+T-rich, as previously observed in *Dictyostelium* (27). The 5′ end contains two stop codons (TAA) in-frame with the first ATG at position 139. The nucleotide sequence surrounding this ATG agrees with the consensus sequence (ANNATGR) for initiation of translation often found in eukaryotes (28). Moreover, the ATG is preceded by an A, as has been found for most *Dictyostelium* genes (27). We thus conclude that translation of the R-subunit mRNA starts at this ATG codon.

The isolation by immunoscreening of a λgt11 cDNA library of clones containing a 5′ nontranslated sequence is unexpected. Indeed, this sequence contains nonsense codons, which will stop translation, in all three reading frames. To explain the synthesis of R subunit and the lack of IPTG induction of the cAMP-binding activity in clones 2.3 and 10.1, we hypothesize that the very A+T-rich 5′ sequence contains signals that can serve for transcription termination as well as reinitiation of both transcription and translation.

The coding region predicts a protein of 327 amino acids, with molecular weight 36,794. The codon usage is analogous to that observed with other *Dictyostelium* genes (29), with the exception of two codons (CTG and GCG) that had not been found previously.

**Comparison of the Primary Structure of *Dictyostelium* and Bovine R Subunits.** Since *Dictyostelium* diverged from the mainstream of eukaryotic descent at the earliest branch point yet characterized by molecular phylogeny (30), the compar-

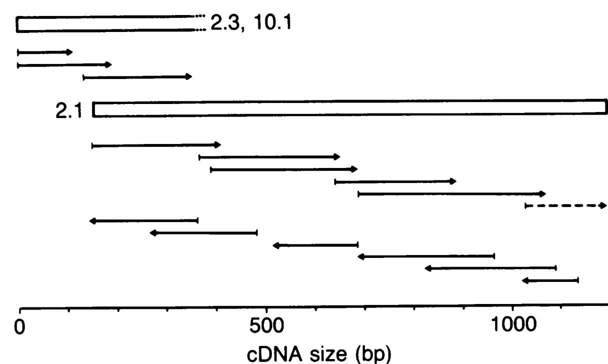


FIG. 1. Sequence analysis of cDNA inserts. Solid arrows show the direction and length [in base pairs (bp)] of individual sequence determinations by the method of Sanger *et al.* (22). The sequence shown by the dashed arrow was determined according to Maxam and Gilbert (25). Sequencing of clone 2.1 was done in both directions except for nucleotides 482–523, which were analyzed only on one strand but using three independent deletions. Similarly, the 5′ end was determined only on one strand but separately in clones 2.3 and 10.1.

| NUCLEOTIDE NUMBER | | AMINO ACID NUMBER |
| --- | --- | --- |
| 1-60 | TAA AAA ATT AAA AAA TTA GGA CAA AAA ACT AAA GTT TTA AAT TAT AGT AAT AGT AAT AAT | |
| 61-120 | AAT AAT AAA ACA ATA AAA ATA GTA TAA AGA AAA ATT TTA TTA TTT ATT ATT TTT TTT GAG | |
| | MET THR ASN ASN ILE SER HIS ASN GLN LYS ALA THR GLU LYS | 1-14 |
| 121-180 | AAA ACA AAA AAA AAA AAA ATG ACA AAT AAT ATA TCA CAT AAC CAA AAA GCA ACA GAA AAA | |
| | VAL GLU ALA GLN ASN ASN ASN ASN ILE THR ARG LYS ARG ARG GLY ALA ILE SER SER GLU | 15-34 |
| 181-240 | GTA GAA GCA CAA AAT AAT AAT AAT ATT ACA CGA AAA AGA AGA GGT GCA ATT AGT AGT GAA | |
| | PRO LEU GLY ASP LYS PRO ALA THR PRO LEU PRO ASN ILE PRO LYS THR VAL GLU THR GLN | 35-54 |
| 241-300 | CCA CTG GGA GAT AAA CCA GCA ACA CCA TTA CCT AAT ATT CCA AAA ACA GTA GAG ACA CAA | |
| | GLN ARG LEU GLU GLN ALA LEU SER ASN ASN ILE MET PHE SER HIS LEU GLU GLU GLU GLU | 55-74 |
| 301-360 | CAA CGT TTA GAA CAA GCA TTA TCA AAT AAT ATT ATG TTT AGT CAT TTA GAA GAG GAG GAA | |
| | ARG ASN VAL VAL PHE LEU ALA MET VAL GLU VAL LEU TYR LYS ALA GLY ASP ILE ILE ILE | 75-94 |
| 361-420 | AGA AAC GTT GTA TTT TTA GCA ATG GTT GAA GTA CTC TAT AAA GCG GGT GAT ATC ATC ATA | |
| | LYS GLN GLY ASP GLU GLY ASP LEU PHE TYR VAL ILE ASP SER GLY ILE CYS ASP ILE TYR | 95-114 |
| 421-480 | AAA CAA GGT GAT GAA GGT GAT CTA TTT TAT GTT ATT GAT TCT GGT ATT TGT GAT ATT TAT | |
| | VAL CYS GLN ASN GLY GLY SER PRO THR LEU VAL MET GLU VAL PHE GLU GLY GLY SER PHE | 115-134 |
| 481-540 | GTT TGT CAA AAT GGT GGT TCC CCA ACT TTA GTA ATG GAA GTA TTT GAA GGT GGT AGT TTT | |
| | GLY GLU LEU ALA LEU ILE TYR GLY SER PRO ARG ALA ALA THR VAL ILE ALA ARG THR ASP | 135-154 |
| 541-600 | GGT GAA TTA GCT TTA ATT TAT GGT AGT CCA AGA GCT GCA ACT GTT ATT GCA AGA ACT GAT | |
| | VAL ARG LEU TRP ALA LEU ASN GLY ALA THR TYR ARG ARG ILE LEU MET ASP GLN THR ILE | 155-174 |
| 601-660 | GTT AGA TTA TGG GCA TTA AAT GGA GCA ACT TAT AGA CGT ATA TTA ATG GAT CAA ACA ATT | |
| | LYS LYS ARG LYS LEU TYR GLU GLU PHE LEU GLU LYS VAL SER ILE LEU ARG HIS ILE ASP | 175-194 |
| 661-720 | AAA AAG AGA AAA TTA TAT GAA GAA TTT TTA GAA AAA GTA TCA ATT TTA CGT CAT ATT GAT | |
| | LYS TYR GLU ARG VAL SER LEU ALA ASP ALA LEU GLU PRO VAL ASN PHE GLN ASP GLY GLU | 195-214 |
| 721-780 | AAA TAT GAA AGA GTA TCA TTA GCA GAT GCA TTG GAA CCT GTT AAT TTT CAA GAT GGT GAG | |
| | VAL ILE VAL ARG GLN GLY ASP PRO GLY ASP ARG PHE TYR ILE ILE VAL GLU GLY LYS VAL | 215-234 |
| 781-840 | GTT ATT GTG CGT CAA GGT GAT CCA GGT GAT AGA TTT TAC ATT ATC GTT GAA GGT AAA GTT | |
| | VAL VAL THR GLN GLU THR VAL PRO GLY ASP HIS SER THR SER HIS VAL VAL SER GLU LEU | 235-254 |
| 841-900 | GTT GTC ACT CAA GAA ACA GTT CCT GGT GAT CAT TCC ACT AGT CAT GTA GTC TCT GAA TTA | |
| | HIS PRO SER ASP TYR PHE GLY GLU ILE ALA LEU LEU THR ASP ARG PRO ARG ALA ALA THR | 255-274 |
| 901-960 | CAT CCT TCT GAT TAC TTT GGT GAA ATT GCA TTA CTT ACT GAT AGA CCA AGA GCT GCA ACT | |
| | VAL THR SER ILE GLY TYR THR LYS CYS VAL GLU LEU ASP ARG GLN ARG PHE ASN ARG LEU | 275-294 |
| 961-1020 | GTA ACT TCT ATT GGT TAT ACA AAA TGT GTA GAA TTG GAT AGA CAA AGA TTT AAT CGT CTT | |
| | CYS GLY PRO ILE ASP GLN MET LEU ARG ARG ASN MET GLU THR TYR ASN GLN PHE LEU ASN | 295-314 |
| 1021-1080 | TGT GGT CCT ATT GAT CAA ATG CTT CGT CGT AAT ATG GAA ACT TAT AAT CAA TTT TTA AAT | |
| | ARG PRO PRO SER SER PRO ASN LEU THR SER GLN LYS SER | 315-327 |
| 1081-1140 | AGA CCA CCT TCT TCA CCA AAT TTA ACC TCT CAA AAA TCT TAA TTT CTT TTT TTT TTT TTT | |
| 1141-1195 | TAT AAT AAC AAC AAA CAC CAA GGT AAT AC(A)25 | |

FIG. 2. Nucleotide sequence of the cDNA for the R subunit. Stop codons are underlined. Amino acid sequence corresponding to the open reading frame is shown above the nucleotide sequence.

ison of the amino acid sequence of the R subunit from *Dictyostelium* ($R_D$) with that of its mammalian counterparts ($R_I$ and $R_{II}$) gives information on the evolution of these proteins. Fig. 3 shows that an optimal alignment of these three sequences can be obtained by introducing a few gaps in only two regions of the sequence.

Most striking is the absence in $R_D$ of the 70 residues corresponding to the $NH_2$ terminus of $R_I$ or $R_{II}$. It has been proposed (31) that the domain responsible for the association of bovine R subunit into dimers is located at the $NH_2$ extremity of $R_I$ and $R_{II}$; this domain has been localized in $R_{II}$ to the 45 $NH_2$-terminal residues (32, 33). The absence of the corresponding sequence in $R_D$ provides structural evidence for its monomeric nature and consequently for the organization of *Dictyostelium* cAMP-dependent protein kinase as an RC dimer.

Starting from the $NH_2$ terminus in $R_D$, the first highly conserved sequence (residues 27–34) corresponds to the so-called "hinge region" previously shown (34) to be involved in the interaction of the C subunit both with $R_I$ and $R_{II}$ and with the heat-stable inhibitor protein. Homology in this region is in agreement with the observation that hybrid holoenzymes can be reconstituted from $R_D$ and bovine C subunit (13). Considering that the residues phosphorylated in $R_I$ and $R_{II}$ are, respectively, Ser-99 and Ser-95 (4, 5), it is likely that *in vitro* phosphorylation of $R_D$ (12) occurs at Ser-32. The maximal homology extends throughout the sequence from Phe-67 to the COOH terminus of $R_D$. In $R_I$ and $R_{II}$, the analogous sequence (starting at Phe-136) corresponds to the two cAMP-binding domains (4, 5, 35). When defined by their residue number in the $R_D$ sequence, these domains extend approximately from Phe-67 to Leu-190 (domain A)

| TYPE OF REGULATORY SUBUNIT | | AMINO ACID NUMBER |
|---|---|---|
| R.D | - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - | 0-0 |
| R.I | A S G T T A S E E E R S L R E C E L Y V Q K H N I Q A L L K D S I V Q L C T | 1-38 |
| R.II | - S H I Q I P P G L T E L L Q G Y T V E V L R Q R P P D L V D F A V D Y F T | 1-37 |
| R.D | - - - - - - - - - - - - - - - - - - - - - - - - - - - - - M T N N I S H N Q | 1-9 |
| R.I | A R P E R P M A F L R E Y F E K L E K E E A K Q I Q N L Q K A G S R A D S R | 39-76 |
| R.II | R L R E A R S R A S T P P A A P P S G S Q D F D P G A G L V A D A V A D S - | 38-74 |
| R.D | K A T E K V E A Q N N N N I T R K R R G A I S S E P L - - - - G D K P A T P | 10-43 |
| R.I | E D E I S P P P P N P V V K G R R R R G A I S A E V Y T - - E E D A A S Y V | 77-112 |
| R.II | E S E D E E D L D V P I P G R F D R R V S V C A E T Y N P D E E E E D T D P | 75-112 |
| R.D | L P N I P K T V E T Q Q R L E Q A L S N N I M F S H L E E E E R N V V F L A | 44-81 |
| R.I | R K V I P K D Y K T M A A L A K A I E K N V L F S H L D D N E R S D I F D A | 113-150 |
| R.II | R V I H P K T D Q Q R C R L Q E A C K D I L L F K N L D P E Q L S Q V L D A | 113-150 |
| R.D | M V E V L Y K A G D I I I K Q G D E G D L F Y V I D S G I C D I Y V C Q N G | 82-119 |
| R.I | M F P V S F I A G E T V I Q Q G D E G D N F Y V I D Q G E M D V Y V N N E W | 151-188 |
| R.II | M F E R T V K V D E H V I D Q G D D G D N F Y V I E R G T Y D I L V T K D - | 151-187 |
| R.D | G S P T L V M E V F E G G S F G E L A L I Y G S P R A A T V I A R T D V R L | 120-157 |
| R.I | A T S - - - - - V G E G G S F G E L A L I Y G T P R A A T V K A K T N V K L | 189-221 |
| R.II | N Q T R S V G Q Y D N H G S F G E L A L M Y N T P R A A T I V A T S E G S L | 188-225 |
| R.D | W A L N G A T Y R R I L M D Q T I K K R K L Y E E F L E K V S I L R H I D K | 158-195 |
| R.I | W G I D R D S Y R R I L M G S T L R K R K M Y E E F L S K V S I L E S L D K | 222-259 |
| R.II | W G L D R V T F R R I I V K N N A K K R K M F E S F I E S V P L L K S L E V | 226-263 |
| R.D | Y E R V S L A D A L E P V N F Q D G E V I V R Q G D P G D R F Y I I V E G K | 196-233 |
| R.I | W E R L T V A D A L E P V Q F E D G Q K I V V Q G E P G D E F F I I L E G S | 260-297 |
| R.II | S E R M K I V D V I G E K V Y K D G E R I I T Q G E K A D S F Y I I E S G E | 264-301 |
| R.D | V V V T Q E - - - - T V P G D H S T S H V V S E L H P S D Y F G E I A L L T | 234-267 |
| R.I | A A V L Q R - - - - - - R S E N E E F V E V G R L G P S D Y F G E I A L L M | 298-329 |
| R.II | V S I L I K S K T K V N K D G E N Q E V E I A R C H K G Q Y F G E L A L V T | 302-339 |
| R.D | D R P R A A T V T S I G Y T K C V E L D R Q R F N R L C G P I D Q M L R R N | 268-305 |
| R.I | N R P R A A T V V A R G P L K C V K L D R P R F E R V L G P C S D I L K R N | 330-367 |
| R.II | N K P R A A S A Y A V G D V K C L V M D V Q A F E R L L G P C M D I M K R N | 340-377 |
| R.D | M E T Y N Q F L N R P P S S P N L T S Q K S | 306-327 |
| R.I | I Q Q Y N S F V S L S V | 368-379 |
| R.II | I S H Y E E Q L V K M F G S S M D L I D P G Q | 378-400 |

FIG. 3. Comparison of the primary structures of R subunits from *Dictyostelium* and bovine heart. Amino acid sequences of $R_I$ and $R_{II}$ are from refs. 3 and 4. Identities of $R_D$ with one or both of the mammalian proteins are boxed. Standard one-letter abbreviations are used.

and from Arg-191 to the COOH terminus (domain B). The presence of two cAMP-binding domains in $R_D$ is unexpected, since binding experiments have shown only one high-affinity cAMP-binding site (11–13). This site is likely to correspond to the $NH_2$-proximal domain (domain A), since fast dissociation kinetics are observed both for $R_D$ (11, 13) and for cAMP binding to domain A in bovine R subunit (36). On the basis of the sequence similarities between the two domains, we hypothesize that domain B in $R_D$ could also bind cAMP, but with a much lower affinity than domain A. In fact, binding of the nucleotide with a $K_d$ in the micromolar range could have failed to be detected under the conditions used for measurement of cAMP binding to purified $R_D$ (12, 13).

On the basis of a comparison of the R subunits from *Dictyostelium*, yeast, and mammals, it is now possible to propose a more detailed scheme for the evolution of these proteins (Fig. 4). The presence of two cAMP-binding domains in $R_D$ indicates that duplication of a gene for a putative ancestral cAMP-binding protein occurred earlier than the acquisition of the dimerization domain D. Therefore, the presence of two cAMP-binding domains in the yeast R subunit is predicted by its similarity to bovine $R_{II}$ (6), although for the yeast protein only one high-affinity cAMP binding site has been demonstrated and no amino acid sequence data are yet available. Since the affinity of cAMP for catabolite gene activator protein is in the micromolar

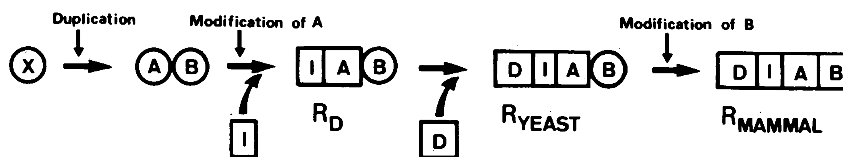


FIG. 4. Hypothetical pathway for the evolution of the regulatory subunits of cAMP-dependent protein kinases. Proteins are designed as sequences of functionally and structurally homologous domains. Arrows show the acquisition of new functions. Evolution of R subunits is assumed to start from an ancestral low-affinity cyclic nucleotide-binding protein (X) and to proceed by a series of genetic events involving gene duplication, gene fusion, and modification by point mutation. Note that the presence of two cAMP-binding domains in R subunit of yeast is predicted by the model (see text) but not yet established on the basis of amino acid sequence data. A and B are cAMP-binding domains carrying low-affinity (circles) or high-affinity (squares) cAMP-binding sites; I is the domain of interaction between R and C subunits; D (boxed) is the domain responsible for dimerization of R subunits.

range (26), we hypothesize that the original duplicated cAMP-binding protein also had a "low" affinity for cAMP. Evolution leading to the primitive eukaryotic cAMP-dependent protein kinase would have involved both acquisition of an I domain required for interaction with the catalytic subunit and modification of domain A to a high-affinity cAMP-binding site with $K_D$ in the nanomolar range. Finally, a second high-affinity binding site in domain B would have appeared later to give mammalian cAMP-dependent protein kinases additional regulatory flexibility through cooperativity between the two cAMP-binding sites.

This model takes into account most of the biochemical properties of the cAMP-dependent protein kinases. However, it leaves open two important questions. What is the advantage conferred by the dimerization domain, and what is the role of cAMP-binding domain B in primitive eukaryotes? We have no answer at present, since the *Dictyostelium* kinase seems as sensitive to *in vitro* activation by cAMP as its more sophisticated homologues (37). The possibility that R subunits could serve another function in addition to the regulation of the catalytic activity has been considered (38). It is tempting to speculate that this (these) putative other function(s) for R subunits is contributed by cAMP-binding domain B and that its importance in the regulation of eukaryotic cell functions is responsible for the conservation of the corresponding sequence from slime molds to mammals.

We thank Drs. G. J. Podgorski, J. Franke, and R. H. Kessin for their expert and friendly advice throughout this work. We are grateful to Dr. Georges N. Cohen for his constant support and helpful discussions. We thank Drs. Claude Parsot, Jean Houmard, and Mario Zakin for help in the sequence determination; Dominique Part for technical assistance; and Lucile Girardot for typing the manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique (U.A. 1129 and ATP 955.113). R.M. is a recipient of a European Molecular Biology Organization fellowship.

1. Flockhart, D. A. & Corbin, J. D. (1982) *Crit. Rev. Biochem.* **12,** 133–186.
2. Lohmann, S. M. & Walter, U. (1984) *Adv. Cyclic Nucleotide Res.* **18,** 63–117.
3. Titani, K., Sasagawa, T., Ericsson, L. H., Kumar, S., Smith, S. B., Krebs, E. G. & Walsh, K. A. (1984) *Biochemistry* **23,** 4193–4199.
4. Takio, K., Smith, S. B., Krebs, E. G., Walsh, K. A. & Titani, K. (1984) *Biochemistry* **23,** 4200–4206.
5. Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A. & Titani, K. (1984) *Biochemistry* **23,** 4207–4218.
6. Hixson, C. S. & Krebs, E. G. (1980) *J. Biol. Chem.* **255,** 2137–2145.
7. Lopes Gomes, S., Juliani, M. H., Maia, J. C. & Rangel-Aldao, R. (1983) *J. Biol. Chem.* **258,** 6972–6978.
8. Trevillyan, J. M. & Pall, M. L. (1982) *J. Biol. Chem.* **257,** 3978–3986.
9. Williams, J. G., Tsang, A. S. & Mahbubani, H. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 7171–7175.
10. Medhy, M. C., Ratner, D. & Firtel, R. A. (1983) *Cell* **32,** 763–771.
11. Arents, J. C. & Van Driel, R. (1982) *FEBS Lett.* **137,** 201–204.
12. Majerfeld, I. H., Leichtling, B. H., Meligeni, J. A., Spitz, E. & Rickenberg, H. V. (1984) *J. Biol. Chem.* **259,** 654–661.
13. de Gunzburg, J., Part, D., Guiso, N. & Véron, M. (1984) *Biochemistry* **23,** 3805–3812.
14. Davis, R. W., Botstein, D. & Roth, J. R. (1980) *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
16. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
17. Young, R. A. & Davis, R. W. (1983) *Science* **222,** 778–782.
18. Huynh, T. V., Young, R. A. & Davis, R. W. (1984) in *DNA Cloning: A Practical Approach*, ed. Glover, D. (IRL, Oxford), pp. 49–78.
19. Gubler, U. & Hoffman, B. J. (1983) *Gene* **25,** 263–269.
20. Hall, R., Hyde, J. E., Goman, M., Simmons, D. L., Hope, I. A., Mackay, M., Scaife, J., Merkli, B., Richle, R. & Stocker, J. (1984) *Nature (London)* **311,** 379–382.
21. Part, D., de Gunzburg, J. & Véron, M. (1985) *Cell Diff.* **17,** 221–227.
22. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
23. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.
24. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13,** 31–40.
25. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
26. Takahashi, M., Blazy, B. & Baudras, A. (1980) *Biochemistry* **19,** 5124–5130.
27. Kimmel, A. R. & Firtel, R. A. (1983) *Nucleic Acids Res.* **11,** 541–552.
28. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
29. Jacquet, M., Kalekine, M. & Boy-Marcotte, E. (1985) *Biochimie* **67,** 583–588.
30. McCarroll, R., Olson, G. J., Stahl, Y. D., Woese, C. R. & Sogin, M. L. (1983) *Biochemistry* **22,** 5858–5868.
31. Potter, R. L. & Taylor, S. S. (1979) *J. Biol. Chem.* **254,** 2413–2418.
32. Rannels, S. R., Cobb, C. E., Landiss, L. R. & Corbin, J. D. (1985) *J. Biol. Chem.* **260,** 3423–3430.
33. Reimann, E. M. (1986) *Biochemistry* **25,** 119–125.
34. Scott, J. D., Fischer, E. H., Takio, K., Demaille, J. G. & Krebs, E. G. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 5732–5736.
35. Bubis, J. & Taylor, S. S. (1985) *Biochemistry* **24,** 2163–2170.
36. Rannels, S. R. & Corbin, J. D. (1980) *J. Biol. Chem.* **255,** 7085–7088.
37. de Gunzburg, J. & Véron, M. (1982) *EMBO J.* **1,** 1063–1068.
38. Constantinou, A. I., Squinto, S. P. & Jungmann, R. A. (1985) *Cell* **42,** 429–437.
39. Lacombe, M. L., Podgorski, G. J., Franke, J. & Kessin, R. H., *J. Biol. Chem.*, in press.