Amino Acid Sequence of the Regulatory Subunit of Bovine Type II Adenosine Cyclic 3′,5′-Phosphate Dependent Protein Kinase†

Koji Takio, Stephen B. Smith,‡ Edwin G. Krebs, Kenneth A. Walsh, and Koiti Titani*

ABSTRACT: Evidence is presented that establishes the amino acid sequence of the regulatory subunit of type II cAMP-dependent protein kinase from bovine cardiac muscle. Complementary sets of overlapping peptides were generated primarily by tryptic digestion and by chemical cleavage at methionyl residues. The analysis was augmented by chemical cleavage at a single tryptophanyl residue and at three of the four aspartyl–proline bonds. Several large fragments generated by limited proteolysis contributed to the proof of structure. The subunit is a single chain of 400 residues corresponding to a molecular weight of 45 004. An amino-terminal segment of about 100 residues is believed to include the region responsible for oligomeric association. The remainder of the molecule consists of two tandem homologous domains, each of which is thought to bind a single molecule of cAMP. Comparison of the three domains with corresponding regions of the type I isozyme, of the *Escherichia coli* catabolite gene activator protein, and of cGMP-dependent protein kinase indicates extensive regions of homology and as much as 50% identity with the sequence of an internal segment of the type I isozyme.

Adenosine cyclic 3′,5′-phosphate (cAMP) dependent protein kinase (EC 2.7.1.37; ATP:protein phosphotransferase) plays important roles in mediating the second messenger effects of cAMP in metabolic and other aspects of cellular control (Krebs & Beavo, 1979; Taylor et al., 1981; Flockhart & Corbin, 1982). Two major types of enzyme have been recognized and studied, namely, type I which predominates in skeletal muscle and type II in cardiac muscle. In each case, the inactive molecule consists of two regulatory and two catalytic subunits. Upon binding of cAMP to the regulatory subunits, the catalytic subunits dissociate and display phosphorylating activity whereas the regulatory subunits remain as a dimer.

Current analyses in our laboratory and others of the amino acid sequence of both types of cAMP-dependent protein kinase are providing information concerning homologies among related families of proteins (Barker & Dayhoff, 1982; Weber et al., 1982; Takio et al., 1983; Titani et al., 1983) and about sites or domains associated with specific aspects of function (Taylor et al., 1981). The amino acid sequence of the catalytic subunit from bovine cardiac muscle has been described in detail (Shoji et al., 1981, 1983), and a preliminary report of the sequence of the regulatory subunit (R_{II})[1] has been published (Takio et al., 1982). The present report presents the experimental basis of the proof of the amino acid sequence of R_{II}.

Materials and Methods

Proline-specific endopeptidase and the glutamyl-directed V8 protease from *Staphylococcus aureus* were obtained from Miles Laboratories. Rat mast cell protease II was a generous gift of Dr. N. Katunuma (Tokushima University, Japan).

The regulatory subunit of cyclic AMP-dependent protein kinase (R_{II}) was prepared from beef heart as described (Dills et al., 1979) and either stored frozen or reduced and S-carboxymethylated (Takio et al., 1980).

Several fragments were generated by limited proteolysis of native protein (Table I), precipitated with 10% trichloroacetic acid, redissolved, reduced, carboxymethylated, and separated on Sephacryl S-200 in acidified 7 M urea as described previously (Takio et al., 1980).

^{14}C-Methylation of methionyl residues in fragment Tβ followed the procedure of Link & Stark (1968). Tβ (100 nmol) was treated with 5 μmol (250 μCi) of $^{14}CH_3I$ in 1 mL of 6 M guanidine hydrochloride, 0.1 M KNO_3, and 0.1 M acetate (pH 5) at room temperature for 20 h in the dark. The protein was separated from the mixture by gel filtration on Sephadex G-25 in 9% formic acid. [^{14}C]Methionyl residues were then generated by treatment with 0.5 M dithiothreitol in 8 M guanidine hydrochloride, pH 10.5 at 37 °C for 20 h, and desalted as above.

Citraconylation followed the procedure of Habeeb & Atassi (1970). In the case of fragment Tβ, 24 nmol of methyl-^{14}C-labeled protein (above) was mixed with 500 nmol of unlabeled Tβ, treated with 200 μL of citraconic anhydride in 2 mL of 8 M guanidine hydrochloride in a pH stat at pH 8.5, and then extensively dialyzed at pH 8.8.

Small peptides were purified by reversed-phase HPLC with a Varian 5000 liquid chromatograph on columns of μBondapak C18 (Waters Associates), ODS-HC/SIL-X-1 (Perkin Elmer), or SynChropak RP-P (SynChrom) using an acetonitrile gradient in aqueous trifluoroacetic acid (Mahoney & Hermodson, 1980). In some cases gel filtration columns were monitored by using fluorescamine by the method of Nakai et al. (1974).

Amino acid analyses, sequenator analyses, and identifications of phenylthiohydantoins followed procedures described previously by Takio et al. (1983).

Segments of protein sequences were examined for homology by using the ALIGN program of Dayhoff et al. (1983) with a gap penalty of 10 and the mutation data matrix. Alignment scores are expressed in units of standard deviation from those

† From the Howard Hughes Medical Institute and the Departments of Biochemistry and Pharmacology, University of Washington, Seattle, Washington 98195. *Received April 23, 1984.* This work was supported by a grant from the National Institutes of Health (GM 15731).

* Address correspondence to this author at the Department of Biochemistry.

‡ Present address: Department of Biomedicinal Chemistry, School of Pharmacy, John Stauffer Pharmaceutical Science Center, University of Southern California, Los Angeles, CA 90035.

[1] Abbreviations: R_{II} and R_I, regulatory subunits of cAMP-dependent protein kinase, types II and I; CAP, catabolite gene activator protein of *E. coli*; cGK, cGMP-dependent protein kinase; SDS–PAGE, polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate; HPLC, high-performance liquid chromatography.

0006-2960/84/0423-4200$01.50/0

Table I: Products of Limited Proteolysis of Native R_{II}[a]

fragment derived	digesting enzyme	ratio of enzyme:substrate	time of digestion	yield (%)	residues	previous name[b]
Tα	trypsin	1:1000	5 min	95	1–92	S_t
Tβ	trypsin	1:1000	5 min	90	93–400	L_t
Tγ	trypsin[c]	1:10	45 min	50	315–400	L_{tC}
Tδ	trypsin[c]	1:10	45 min	40	93–314	L_{tN}
Tε	trypsin[c]	1:10	45 min	50	1–43	S_{tN}
T1	trypsin[c]	1:10	45 min	35	1–22	
T2	trypsin[c]	1:10	45 min	35	23–38	
T6	trypsin[d]	1:50	3 h	65	46–89	
MCα	mast cell protease II	1:100	20 h	85	1–97	S_{mcp}
Eα	*S. aureus* protease	1:100	2 h	95	1–99	S_{sp}

[a] Digested, S-carboxymethylated, and purified on Sephacryl S-200 as described by Takio et al. (1980), except for peptide T6[d]. [b] Nomenclature used by Takio et al. (1980, 1982). [c] See Figure 2. [d] Peptide T6 was recovered from the 10% trichloroacetic acid supernatant after precipitation of the products of limited tryptic hydrolysis (Takio et al., 1980). The supernatant was desalted on Sephadex G-25 and then fractionated by HPLC.

of randomly generated sequences. Usage of this program in our laboratory is described in greater detail by Reimann et al. (1984).

Results

The general strategy of analysis involved principally two overlapping sets of peptides, one generated by cleavage at methionyl residues and another by tryptic digestion at arginyl residues. In addition, our earlier study of the sensitivity of R_{II} to limited proteolysis (Takio et al., 1980) facilitated the generation of several large fragments (Table I) which proved to be useful as targets for both Edman degradation (Figure 1) and subdigestion (Table II). The results of Edman degradations of these and other fragments are summarized in Figure 1, which illustrates the overlapping matrix of data constituting the proof of sequence of R_{II}.

Limited proteolytic digestion was applied to the native molecule in a variety of ways (Table I). As reported previously (Takio et al., 1980), mild treatment with trypsin, *S. aureus* protease, or mast cell protease each broke the chain in a single location to give a blocked amino-terminal fragment Tα, Eα, or MCα, respectively, and the corresponding larger carboxyl-terminal fragment beginning at residues 93, 98, or 100. More intensive treatment with trypsin led to further fragmentation of Tα (residues 1–92) to a smaller peptide Tε (residues 1–43), previously referred to as S_{tN} (Takio et al., 1982), or to the limit digestion products T1–T7 (Figures 1 and 2; Tables I and II). Similarly, more intensive treatment with trypsin replaced the large carboxyl-terminal product of limited digestion, Tβ (residues 93–400), with fragments Tδ (residues 93–314) and Tγ (residues 315–400).

Cleavage with CNBr. A complete set of fragments was generated by cleavage of reduced, S-carboxymethylated R_{II} with CNBr and separation of this mixture as illustrated in Figure 3. Several peptides required further purification, as indicated in Table III. Their compositions are listed in Table III, and most of their sequences are included in the summary in Figure 1.

Alignment of Residues 93–400. Of the 10 purified fragments, nine (M2–M10) each yielded a single sequence during Edman degradation, and only the largest (M1) was refractory to this procedure. Since the native protein is blocked, it was assumed that fragment M1 occupied the amino-terminal 151 residues including the hinge region and the autophosphorylation site previously identified (Takio et al., 1980) at the amino-terminal region of the large tryptic fragment Tβ (residues 93–400).

The carboxyl terminus of the hinge region (DTDPRVIH) contains an acid-labile Asp–Pro bond (residues 111 and 112 in Figure 1). Its cleavage product with the amino-terminal

Table II: Peptides Generated by Chemical or Enzymatic Cleavage

peptide	source fragment	cleavage[a] technique	yield[b] (%)	residues
T1-C1	T1	chymotrypsin	45	1–12
T1-C2	T1	chymotrypsin	48	1–13
T1-P2	T1	prolyl endopeptidase	45	8–22
Tε-PP	Tε	pepsin[c]	60	20–30
MCα-C3	MCα	chymotrypsin	40	22–36
MCα-C4	MCα	chymotrypsin	46	37–66
Eα-DP2	Eα	Asp–Pro cleavage[d]	30	62–111
T7	Tα	trypsin	60	90–92
Tβ-DP2	Tβ	Asp–Pro cleavage[e]	35	112–140
M1-E9	M1	*S. aureus* protease	25	129–151
Tβ-DP3	Tβ	Asp–Pro cleavage[e]	30	141–397
Tβ-R8	Tβ[f]	trypsin	60	192–213
W2	R_{II}[g]	Trp cleavage[h]	35	227–400
Tβ-R13	Tβ[f]	trypsin		246–266
Tβ-R14	Tβ[f]	trypsin	60	267–283
M5-R15	M5[i]	trypsin	45	284–325
Tγ-R19	Tγ[i]	trypsin	38	377–400

[a] Unless otherwise indicated, enzymatic digests contained 1:30–1:100 weight ratios of enzyme:substrate in 0.1 M NH_4HCO_3, at 37 °C, 3–4 h. The digests were separated by HPLC on μBondapak C18 (see Materials and Methods). [b] Estimated by amino acid analysis. [c] Digested in 10 mM HCl. [d] Cleaved in 90% formic acid, 170 h, at room temperature, and separated on DEAE-Sephacel (pH 8.5) by using a gradient of ammonium bicarbonate (0.05–1 M). [e] Incubated 90 h at 37 °C in 10% acetic acid–7 M guanidine hydrochloride and adjusted to pH 2.5 with pyridine. Separation was effected on Sephacryl S-200 in 7 M urea and 10 mM HCl. [f] ^{14}C-Methylated and citraconylated (Figure 4). [g] The reduced, S-carboxymethylated form of the regulatory subunit was used. [h] Cleaved by the method of Mahoney & Hermodson (1979) and separated on Sephacryl S-200 (1.5 × 171 cm) in 7 M urea–10 mM HCl. [i] The fragment was citraconylated before tryptic digestion.

sequence PRVIHPKT··· (fragment Tβ-DP2 in Table III) was found to overlap M1-E9 (Figure 1) which, in turn, overlapped another Asp-Pro cleavage product, Tβ-DP3, thus providing a continuous sequence through Met-151 which aligned M1 and M2.

Proof of the structure from this overlapped region (residues 93–160) to the carboxyl terminus of the whole protein was derived primarily by overlapping the set M2 to M10 with tryptic peptides derived (Figure 4) from ^{14}C-methylated, citraconylated Tβ and from citraconylated Tγ (Tβ-R8 through Tγ-R19 in Figure 1). The search for peptides that overlapped the CNBr fragments was assisted by the ^{14}C radiolabel introduced into the methionyl residues by the methylation/demethylation procedure. Two additional peptides provided necessary overlaps (Figure 1). The sequence of W2, a product of chemical cleavage of the single tryptophanyl bond at residue 226, aligned M3 and M4. An extended sequence analysis of Tγ aligned M5–M8. M10 was the only product of CNBr cleavage that lacked homoserine; hence, it must represent the

Table III. Amino Acid Compositions[a] of Peptides Isolated from Subunit R_{II}

Peptide Source[b] Residues No.		M1 1-151	M2 HPLC 152-208	M3 HPLC 209-247	M4 G-25 248-267	M5 268-358	M6 359-371	M7 G-25 372-374	M8 G-25 375-388	M9 G-25 389-393	M10 394-400	Hydrolysis of R_{II}	Sequence of R_{II}	T1 Table I 1-22	Tγ-R19 Table II 377-400	T1-C1-A1 see text 1-12	T1-C2 Table II 1-13	T1-C2-Th2 see text 3-8	Tε-PP Table II 20-30	Tε Table I 1-43
Asp/Asn	(D/N)	18.8(20)	9.6(11)	4.2(4)	0.9(0)	8.9(8)	1.4(1)	1.1(1)	1.0(1)	0.5(0)	1.9(2)	46.7	(48)		2.7(3)				1.7(2)	3.2(3)
Thr	(T)	6.9(7)	3.6(4)	3.6(4)		3.4(3)						18.0	(18)	1.9(2)		1.1(1)	1.0(1)			2.9(3)
Ser	(S)	8.3(9)	1.8(2)	1.9(2)	3.2(4)	3.9(5)		0.5(0)	1.2(1)	1.9(2)		25.1	(25)	1.1(1)	2.4(3)	1.5(1)	1.0(1)			1.2(1)
Glu/Gln	(E/Q)	23.4(23)	7.0(7)	1.4(1)	4.9(4)	12.0(12)	2.0(2)		2.6(3)		1.1(1)	48.8	(53)	4.5(4)	4.1(4)	2.4(2)	2.4(2)	1.1(1)	1.4(1)	6.5(6)
Pro	(P)	11.9(15)		1.0(1)	1.1(1)	1.5(1)	0.9(1)				0.8(1)	18.9	(20)	2.1(2)	0.7(1)	1.5(2)	2.2(2)	2.0(2)	2.2(2)	4.1(4)
Gly	(G)	7.6(6)	5.6(6)	2.1(2)	0.5(0)	6.7(8)	1.4(1)			1.2(1)	1.2(1)	25.0	(25)	2.2(2)	1.8(2)	1.7(1)	1.1(1)	1.0(1)		2.2(2)
Ala	(A)	12.0(12)	1.3(1)	4.0(4)	0.5(0)	5.8(7)	0.9(1)					25.2	(25)		0.6(0)					2.0(2)
CMCys	(C)	2.0(3)				1.6(2)	0.2(1)					5.6	(6)							
Val	(V)	11.5(11)	5.5(6)	2.9(3)	2.2(2)	8.5(10)	0.9(1)		0.9(1)			32.3	(34)	2.3(2)	1.2(1)				2.0(2)	3.6(4)
Met[c]	(M)	0.3(1)	1.0(1)	0.8(1)	1.0(1)	0.7(1)	0.8(1)	1.0(1)	0.8(1)	1.4(1)		7.4	(9)		0.6(2)					
Ile	(I)	5.5(5)	2.8(3)	2.6(3)	1.1(1)	5.7(9)		1.0(1)	0.9(1)		0.9(1)	22.3	(24)	1.9(2)	1.8(2)	2.0(2)	2.0(2)	1.8(2)		2.0(2)
Leu	(L)	13.4(14)	3.1(3)	2.2(2)	3.0(3)	4.5(4)	2.0(2)		1.0(1)		1.0(1)	30.0	(30)	4.0(4)	2.0(2)	2.0(2)	3.0(3)		2.0(2)	6.0(6)
Tyr	(Y)	1.9(3)	2.6(3)	1.1(1)		3.2(4)			0.8(1)			8.8	(12)	0.8(1)	0.8(1)					0.9(2)
Phe	(F)	5.5(5)	2.8(3)	1.1(1)	2.2(2)	2.4(2)	0.9(1)			1.0(1)		14.5	(15)		1.0(1)					1.9(2)
His	(H)	2.9(2)	1.3(2)			0.6(1)			0.8(1)			6.6	(6)	0.6(1)	0.7(1)	1.0(1)	1.0(1)			1.1(1)
Lys	(K)	5.1(3)	2.1(2)	3.9(4)	1.3(1)	8.5(11)			1.7(2)			23.3	(23)		1.0(1)					
Arg	(R)	12.9(12)	3.0(3)	4.9(5)	1.0(1)	3.7(3)	0.8(1)		1.0(1)			26.4	(26)	0.9(1)					1.8(2)	4.4(5)
Trp[d]	(W)												(1)							
Total			(151)	(57)	(39)	(20)	(91)	(13)	(3)	(14)	(5)	(7)		(400)	(22)	(24)	(12)	(13)	(6)	(11)	(43)
Yield (%)			62	35	30	61	90	52	15	65	7	69			61	38	27	48	80	75	50

[a] Residues per peptide by amino acid analysis (values < 0.4 not reported, except Met) or, in parentheses, from the sequence (Fig. 1).

[b] Peptides M1 through M10 are from cleavage at methionyl residues. The primary separation of fragments (Fig. 3) was followed where indicated by HPLC, or gel filtration on Sephadex G-25. The seven peptides on the right were derived from appropriate enzymatic digests (Table I, Table II).

[c] Methionine measured as homoserine in the peptides M1 through M10.

[d] Not determined.

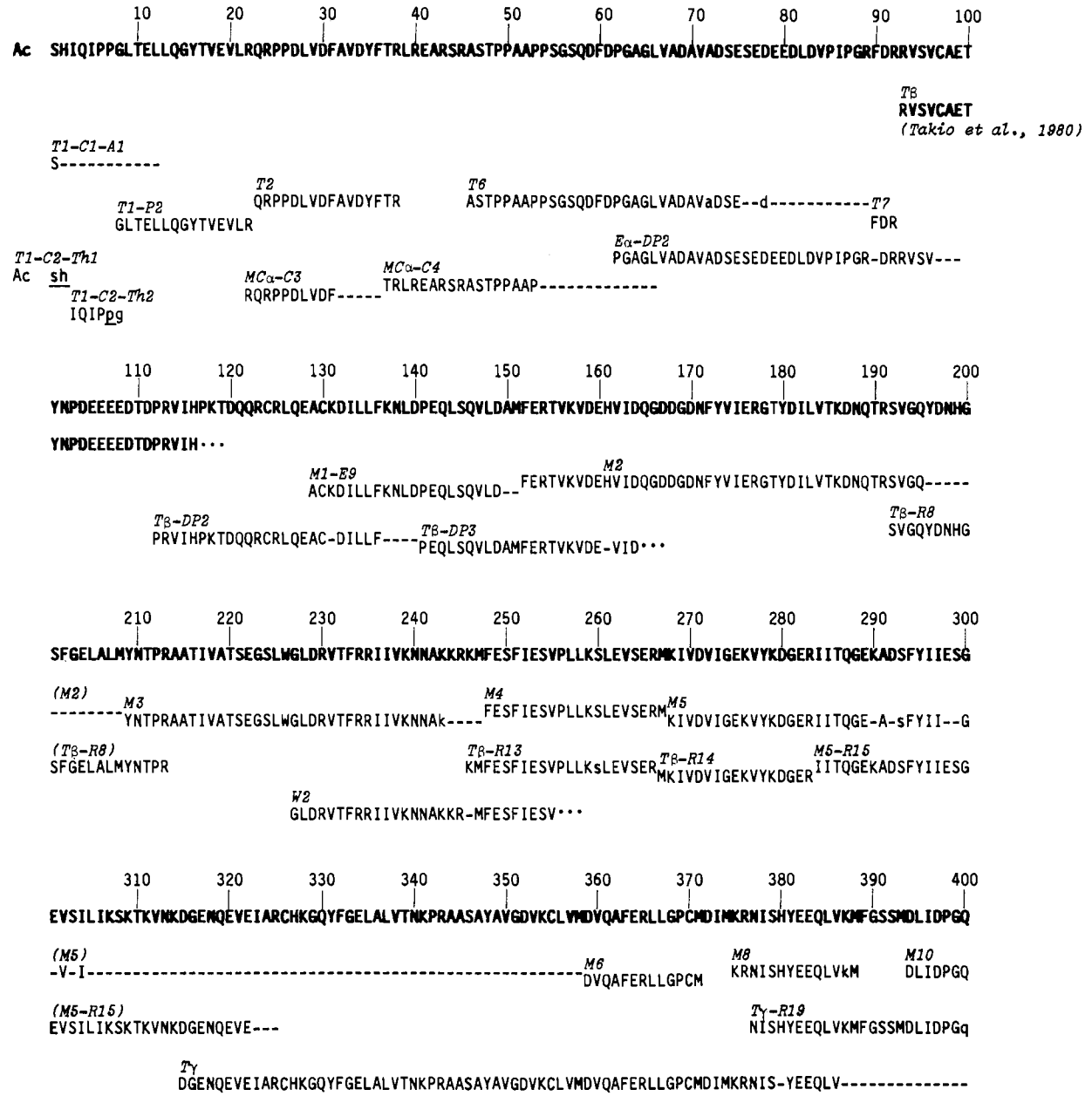

FIGURE 1: Summary of the proof of structure of the regulatory subunit. The proven sequence of specific peptides (named in italics) are given in one letter code below the summary sequence (bold type). Peptides generated by limited proteolysis (Table I) are identified with a prefix identifying the digesting enzyme (T, trypsin; MC, mast cell protease; E, *S. aureus* protease), followed by a greek letter (e.g., Tβ). Peptides with prefixes of M and W were generated by cleavage at methionyl (Figure 3) or tryptophanyl residues, respectively. Hyphenated names of peptides indicate subdigestion (Table II) with an agent identified in the suffix using the following code: C, chymotrypsin; Th, thermolysin; P, prolyl endopeptidase; A, 12 N HCl; DP, mild acid; R, trypsin (after citraconylation). Peptide sequences written in upper case letters are proven by Edman degradation. A lower case letter indicates that the identification is considered to be tentative. A hyphen denotes an unidentified residue. The symbol (···) indicates that the peptide is much longer but that no further residues were identified by Edman degradation. The underlined residues are placed primarily by amino acid composition. Ac denotes an acetyl group.

carboxyl terminus of the protein. This is confirmed by the analogous argument concerning the exclusive lack of arginine in Tγ-R19.

Within this segment of the molecule (residues 93–400), all overlaps span at least three residues except for that between M4 and M5 (residues 265–269) where a Glu-Arg-Met-Lys-Ile sequence yielded only single residue overlaps. However, the only other Met-Lys sequence is already well placed in Tγ (residues 374 and 375), and only M4 could, on tryptic digestion of citraconylated protein, yield an amino-terminal methionine.

The carboxyl terminus of M3 (Lys-Met) is also somewhat tentative because the phenylthiohydantoin of Lys-246 was not identified in M3 or in W2, and its identification within Tβ-R13 lacked rigor because the peptide was cosequenced in a 3:2 mixture with Tβ-R19.

Alignment of Residues 1–92. Alignment of peptides within the amino-terminal segment was difficult because the amino terminus of the protein is blocked. Good overlaps among MCα-C4, T6, Eα-DP2, T7, and Tβ led from the hinge region back to Thr-37 (Figure 1). Peptide T2 overlaps that sequence by two residues and with MCα-C3 extends the sequence to Arg-Gln-Arg-Pro (residues 22–25).

Determination of the sequence of T1 (residues 1–22) presented a special problem because the peptide was blocked. Furthermore, its sequence overlapped MCα-C3 by only a single arginyl residue, a weak point in the proof of structure of R_{II}. The overlap was strengthened by the composition of a peptic peptide (Tϵ-PP; Table III), which corresponded to residues 20–30, and by that of Tϵ itself, which fitted residues 1–43. Much of the sequence of T1 was provided by analysis

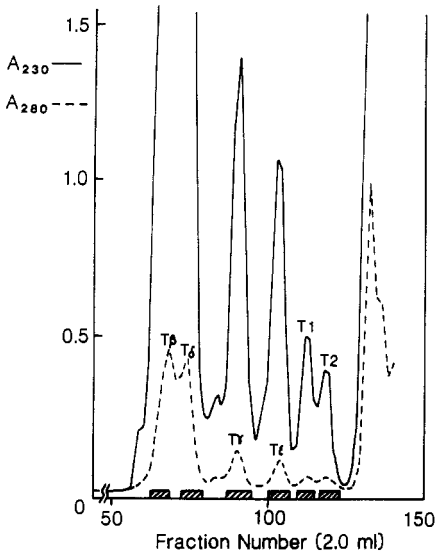

FIGURE 2: Separation of the products of limited tryptic hydrolysis of 20 mg of regulatory subunit. The digestion products were precipitated with 10% trichloroacetic acid and then ^{14}C-labeled S-carboxymethylated as described by Takio et al. (1980). The mixture was applied to a column (1.5 × 170 cm) of Sephacryl S-200 in 7 M urea–10 mM HCl and eluted at 12 mL/h. The large fragments Tβ, Tδ, Tγ, and Tε were pooled as indicated and desalted on Sephadex G-25 in 9% formic acid. Two small peptides (T1 and T2) were repurified on HPLC.

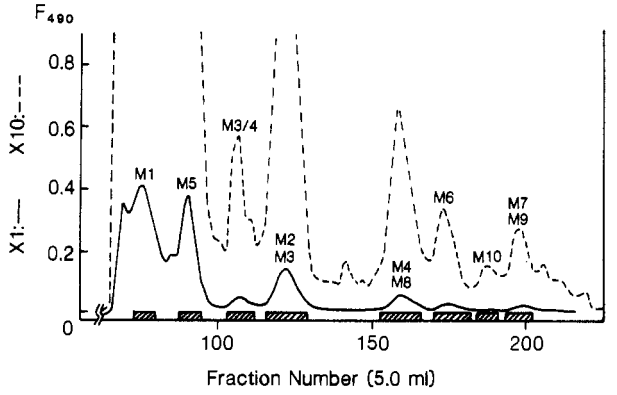

FIGURE 3: Primary separation of peptides after cleavage at methionyl bonds. ^{14}C-Labeled S-carboxymethylated regulatory subunit (50 mg) was treated with 50 mg of CNBr in 5 mL of 70% formic acid for 3 h. After lyophilization, the peptides were dissolved in 9% formic acid, applied to a column (2.5 × 186 cm) of Sephadex G-50 superfine, and eluted with 9% formic acid at 30 mL/h. Pooled fractions contained the indicated fragments which were further purified, if necessary, as indicated in Table III.

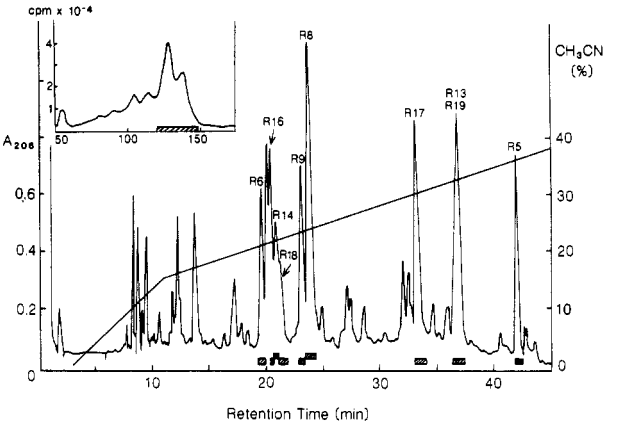

FIGURE 4: Separation of peptides derived by tryptic digestion of 524 nmol of methyl-^{14}C-labeled, N-citraconylated fragment Tβ. Inset: gel filtration on a column (1.5 × 145 cm) of Sephadex G-50 superfine in 9% formic acid at 8.2 mL/h. The broad pooled fraction was then separated by HPLC on an ODS-HC/SIL-X-1 column (Perkin-Elmer) with the indicated gradient of acetonitrile into 0.1% aqueous trifluoroacetic acid. The flow rate was 2 mL/min. Fractions pooled as indicated yielded two pure peptides used in Figure 1 (R8, R14), a mixture of R13 and R19 which was subjected to Edman degradation as that mixture, and six other peptides with compositions suggesting the following loci: R5 (residues 126–154), R6 (155–177), R9 (214–230), R16 (326–343), R17 (344–365), and R18 (366–376).

Table IV: Test for Homology[a] of Segments[b] of the Sequence of R_{II} to Segments of Other Sequences

		R_{II} residues		
	residues	1–134	135–256	257–400
R_I	1–134	3.5 (20%)		
	135–252		20.9 (50%)	12.5 (35%)
	253–379		10.7 (30%)	17.8 (41%)
CAP[c]	1–130		6.4 (18%)	7.0 (26%)
cGK[d]	1–101	0.5 (10%)		
	102–219		15.5 (31%)	8.5 (31%)
	220–340		6.9 (26%)	10.2 (36%)
R_{II}	135–256			12.7 (34%)

[a] The significance of a homologous relationship is expressed as an alignment score in standard deviation units and, in parentheses, as percentage identity. In the analyses alignments were optimized with a computer program using a penalty for the introduction of a gap (see Materials and Methods). Alignment scores are based on similarities in sequence and provide a statistical measure of the probability that the homology could not be the product of mere chance. The specific residue alignments are presented in Takio et al. (1984). [b] The three segments defined here within R_{II}, R_I, and cGK correspond in each protein to the set of segments defined as A, B, and C in Takio et al. (1984) and to the three domains discussed by Takio et al. (1982). [c] Weber et al. (1982). [d] Takio et al. (1984).

of T1-P2, the product of digestion with the prolyl endopeptidase. A chymotryptic digest of T1 yielded two blocked peptides differing by the presence of leucine residue 13 in T1-C2 but not in T1-C1. Comparison of their compositions (cf. T1-C2 in Table III) with that of T1, and subtraction of the carboxyl-terminal sequence (T1-P2), indicated that T1 contained the residues Ser, His, Ile_2, Glx, and Pro_2 between the blocked amino terminus and the Gly-Leu-Thr sequence (residues 8–10). Thermolytic digestion of T1-C2 yielded two peptides, a blocked dipeptide T1-C2-Th1 (Ser, His), and a hexapeptide T1-C2-Th2 (Table III) which washed out of the spinning cup badly during degradation. A tentative carboxyl-terminal sequence of T1-C2-Th2, Pro-Gly, was deduced to overlap the amino-terminal glycyl residue of T1-P2 on the grounds of both composition and the specificity of the prolyl endopeptidase. Thus, the sequence of residues 3–22 was established. Proton NMR of T1-C2-Th1 was reported previously to have identified an acetyl group as the amino-terminal block (Takio et al., 1982). It was partially removed by an N to O shift during overnight treatment of T1-C1 with 12 N HCl at room temperature. The product of deacylation, purified on HPLC, contained both serine and histidine (Table III), and the amino terminus was identified as serine by Edman degradation.

On the basis of our previous reports of internal homology (Takio et al., 1982) and of homology with R_I, cGK, and CAP (Weber et al., 1982; Takio et al., 1983), we applied a computerized program (cf. Materials and Methods) to estimate the similarity of three segments of R_{II} with the corresponding segments of the other proteins. The specific residue alignments are given elsewhere (Takio et al., 1984), but the alignment scores and percent identities are listed in Table IV. Alignment scores greater than 5 indicate that the relatedness of two segments differs from mean scores of scrambled sequences of 5 standard deviations and that the segments are homologous. An alignment score of 1 is 1 standard deviation from the mean

score of scrambled sequences and thus not diagnostic of homology. These data provide a more quantitative basis for prior conclusions of internal homology (residues 135–256 vs. 257–400) and of homology with CAP. In addition, these data indicate that the most similar segments are the middle segments of R_{II} and R_I and that the amino-terminal segment of R_{II} is marginally related to that of R_I but not detectably to that of cGK.

Discussion

In general the proof of the sequence of the regulatory subunit of bovine type II cAMP-dependent protein kinase (Figure 1) involved two primary sets of overlapping peptides derived by cleavage at either methionyl or arginyl residues. In addition, advantage was taken of limited proteolysis in a "hinge" region (Takio et al., 1980) and of an anomalous tryptic cleavage at a Lys–Asp bond (residues 314–315). The susceptibility of the protein to limited proteolysis in the hinge region has been observed before in several laboratories (Flockhart et al., 1980; Takio et al., 1980; Potter & Taylor, 1980), but a rationale for the unexpected lability of a Lys–Asp bond only becomes clear on comparison with another cAMP-binding protein of known three-dimensional structure (Weber et al., 1982), i.e., CAP as discussed below. In addition to this general strategy, chemical cleavage at the single tryptophanyl residue and acid cleavage at three of the four Asp–Pro bonds provided fragments useful for the proof of sequence. In a search for methionyl overlaps, the methionyl residues were exogenously labeled with [^{14}C]methyl iodide to facilitate their location among the many tryptic peptides. Finally, tactics for determination of the amino-terminal region were complicated by the blocking α-*N*-acetyl group which necessitated subdigestion with several enzymes, including a prolyl endopeptidase that recently became available (Yoshimoto et al., 1980).

Approximately 75% of the reported sequence was separately determined in more than one set of overlapping peptides. The proof is further strengthened by the fact that all analyses of phenylthiohydantoins were performed independently in two separate and complementary HPLC systems. The weakest data in the analysis are found in the amino-terminal nonapeptide region where compositional and even specificity arguments were necessary for proof of structure. Other weak points are seen in the single residue overlaps at residues 21–23 (Leu-Arg-Gln) and residues 266–268 (Arg-Met-Lys).

We have discussed previously (Takio et al., 1982) the location of an amino-terminal domain involved in dimerization and of duplicated internal domains identified with binding of cAMP, so this report will largely summarize those aspects of the molecule and add observations which have become available since that earlier report.

Residues 1–134 constitute the amino-terminal domain (Takio et al., 1982) which contains sites involved in the dimerization of the regulatory subunits and in binding to the catalytic subunit which "autophosphorylates" serine-95 of the regulatory subunit. The existence of a second phosphorylation site was established by Flockhart et al. (1980) but not located in the sequence. More recent studies of Carmichael et al. (1982) suggest that this site is a serine residue which can be labeled by casein kinase II with [γ-^{32}P]ATP. These investigators carried out limited proteolysis and fragment separation as in Figure 2, then subdigested the amino-terminal fragment more extensively, and identified a labeled peptide (residues 67–90) with serine residues corresponding to positions 74 and 76 in Figure 1. These two residues were not distinguished from each other as phosphorylation sites, but in a similar study Hemmings et al. (1982) showed that both serines were phosphorylated by glycogen synthase 5. It is interesting that both these serine residues lie within an acidic region characteristic of casein kinase substrates, in contrast to the basic region near serine-95 characteristic of phosphorylation by cAMP-dependent protein kinase itself. In addition to these three serine residues which can become phosphorylated, two others at residues 44 and 47 are phosphorylated by glycogen synthase 3 (Hemmings et al., 1982). At this time, it is not clear that the phosphorylation of any of these sites plays a significant role in the regulation of function of the enzyme.

Prior alignment of the sequence of R_{II} with those of its isozyme R_I and of the cGMP-dependent enzyme (cGK) in the region of their phosphorylation sites (Takio et al., 1983) suggested homology in a 45-residue segment surrounding Ser-95. However, when longer segments, including the amino termini, are compared (residues 1–134 in Table IV), a homologous relationship to cGK is not evident, although that to R_I is significant. Thus, constraints on conservation of sequence in this region may be quite localized.

In contrast, the two carboxyl-terminal domains (residues 135–256 and 257–400) are clearly homologous with R_I, with cGK, with each other, and with CAP, in decreasing order of significance (Table IV). Each of these relationships is more significant than any in the amino-terminal region, suggesting conservation of a duplicated functional domain during several gene duplication events. It is well established that the carboxyl-terminal two-thirds of the regulatory subunit serves to bind two molecules of cAMP during the regulatory process (Corbin et al., 1978; Weber & Hilz, 1979; Taylor et al., 1981). We discussed previously the internal homology of R_{II} in terms of the two tandem domains binding cAMP (Takio et al., 1982). Each of these domains has since been shown to be homologous with the cAMP-binding domain of CAP, the catabolite gene activator protein of *Escherichia coli* (Weber et al., 1982), a measure of which is given in Table IV. Under the assumption that the three-dimensional structure of the cAMP-binding domain of CAP is a guide to homologous substructures within R_{II}, details of the hypothetical binding site in R_{II} can be examined. Photoaffinity labeling of porcine R_{II} by Kerlavage & Taylor (1980) with a cAMP analogue identified a tyrosine residue at its binding site which corresponds to Tyr-381 in homologous bovine R_{II}. However, comparison of the three-dimensional structure of the binding domain of CAP with the homologous carboxyl-terminal domain of R_{II} suggests that Tyr-381 is proximal to, rather than a contributory component of, the binding site. It also becomes evident that the trypsin-susceptible Lys-Asp dipeptide (residues 314–315) mentioned above would correspond to a loop between two strands of β-structure in CAP, perhaps contributing to the anomalous lability of this bond to trypsin in R_{II}.

Majerfeld et al. (1984) have recently studied cAMP-dependent protein kinase isolated from the primitive eukaryote *Dictyostelium discoideum*. Its regulatory subunit appears to be the functional counterpart of the bovine regulatory subunit since it inhibits bovine catalytic subunits and only in the absence of cAMP (Leichtling et al., 1981). Interestingly, on SDS–PAGE the *D. discoideum* regulatory subunit appears to be about 10 000 daltons smaller than the bovine subunit (Majerfeld et al., 1984) and to bind only one cAMP (de Wit et al., 1982). It would be intriguing to compare the primary structure of this protein with that of the bovine subunit. It is possible that the region which appears to have been duplicated in a tandem manner in the bovine genome would occur as a single copy in this more primitive organism. However,

on SDS–PAGE the bovine subunits have anomalous mobilities suggesting molecular weights 5000–9000 too large (Takio et al., 1982; Titani et al., 1984). These anomalies appeared to be related to asymmetries in the dimerization domains. If the *D. discoideum* subunit lacks this anomaly, it could be as large as the bovine subunits and contain a cryptic binding site as well as an effective one.

Much more can be deduced from the homologous relationships among R_{II}, R_I, CAP, and cGK, and these considerations are presented in another report (Takio et al., 1984) together with proof of the sequence of cGK. The proof of sequence of R_I is presented in another report (Titani et al., 1984).

Acknowledgments

We thank Floyd E. Kennedy, Barbara Flug, Richard R. Granberg, Roger D. Wade, Santosh Kumar, and Brad McMullen for their skillful assistance with various aspects of this work.

References

Barker, W. C., & Dayhoff, M. O. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2836.

Carmichael, D. F., Geahlen, R. L., Allen, S. M., & Krebs, E. G. (1982) *J. Biol. Chem. 257*, 10440.

Corbin, J. D., Sugden, P. H., West, L., Flockhart, D. A., Lincoln, T. M., & McCarthy, D. (1978) *J. Biol. Chem. 253*, 3997.

Dayhoff, M. O., Barker, W. C., & Hunt, L. T. (1983) *Methods Enzymol. 91*, 524.

de Wit, R. J. W., Arents, J. C., & van Driel, R. (1982) *FEBS Lett. 145*, 150.

Dills, W. L., Goodwin, C. D., Lincoln, T. M., Beavo, J. A., Bechtel, P. J., Corbin, J. D., & Krebs, E. G. (1979) *Adv. Cyclic Nucleotide Res. 10*, 199.

Flockhart, D. A., & Corbin, J. D. (1982) *CRC Crit. Rev. Biochem. 12*, 133.

Flockhart, D. A., Watterson, D. M., & Corbin, J. D. (1980) *J. Biol. Chem. 255*, 4435.

Habeeb, A. F. S. A., & Atassi, M. Z. (1970) *Biochemistry 9*, 4939.

Hemmings, B. A., Aitken, A., Cohen, P., Raymond, M., & Hofmann, F. (1982) *Eur. J. Biochem. 127*, 473.

Kerlavage, A. R., & Taylor, S. S. (1980) *J. Biol. Chem. 255*, 8483.

Krebs, E. G., & Beavo, J. A. (1979) *Annu. Rev. Biochem. 48*, 923.

Leichtling, B. H., Spitz, E., & Rickenberg, H. V. (1981) *Biochem. Biophys. Res. Commun. 100*, 515.

Link, T. P., & Stark, G. R. (1968) *J. Biol. Chem. 243*, 1082.

Mahoney, W. C., & Hermodson, M. A. (1980) *J. Biol. Chem. 255*, 11199.

Majerfeld, I. H., Leichtling, B. H., Meligoni, J. A., Spitz, E., & Rickenberg, H. V. (1984) *J. Biol. Chem. 259*, 654.

Nakai, N., Lai, C. Y., & Horecker, B. L. (1974) *Anal. Biochem. 58*, 563.

Potter, R. L., & Taylor, S. S. (1980) *J. Biol. Chem. 255*, 9706.

Reimann, E. M., Titani, K., Ericsson, L. H.., Wade, R. D., Fischer, E. M., & Walsh, K. A. (1984) *Biochemistry* (accompanying paper in this issue).

Shoji, S., Parmelee, D. C., Wade, R. D., Kumar, S., Ericsson, L. H., Walsh, K. A., Neurath, H., Long, G. L., Demaille, J. G., Fischer, E. H., & Titani, K. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 848.

Shoji, S., Ericsson, L. H., Walsh, K. A., Fischer, E. H., & Titani, K. (1983) *Biochemistry 22*, 3702.

Takio, K., Walsh, K. A., Neurath, H., Smith, S. B., Krebs, E. G., & Titani, K. (1980) *FEBS Lett. 114*, 83.

Takio, K., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2544.

Takio, K., Smith, S. B., Walsh, K. A., Krebs, E. G., & Titani, K. (1983) *J. Biol. Chem. 258*, 5531.

Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1984) *Biochemistry* (following paper in this issue).

Taylor, S. S., Kerlavage, A. R., Zoller, M. J., Nelson, N. C., & Potter, R. L. (1981) *Cold Spring Harbor Conf. Cell Proliferation 8*, 3.

Titani, K., Takio, K., Sasagawa, T., Kumar, S., Wade, R. D., Ericsson, L. H., Walsh, K. A., Smith, S. B., & Krebs, E. G. (1983) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 42*, 2256.

Titani, K., Sasagawa, T., Ericsson, L. H., Kumar, S., Smith, S. B., Krebs, E. G., & Walsh, K. A. (1984) *Biochemistry* (preceding paper in this issue).

Weber, I. T., Takio, K., Titani, K., & Steitz, T. A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7679.

Weber, W., & Hilz, H. (1979) *Biochem. Biophys. Res. Commun. 90*, 1073.

Yoshimoto, T., Walter, R., & Tsuru, D. (1980) *J. Biol. Chem. 255*, 4786.