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

Koiti Titani,* Tatsuru Sasagawa,‡ Lowell H. Ericsson, Santosh Kumar, Stephen B. Smith,§ Edwin G. Krebs, and Kenneth A. Walsh

ABSTRACT: The complete amino acid sequence of the regulatory subunit of type I cAMP-dependent protein kinase from bovine skeletal muscle is presented. The S-carboxymethylated protein was cleaved with cyanogen bromide to provide a complete set of nonoverlapping fragments. These fragments were overlapped and aligned by using peptides generated by proteolytic cleavage. The protein contains 379 amino acid residues corresponding to a molecular weight of 42 804. As in the type II regulatory subunit of cAMP-dependent protein kinase, a pattern of internal gene duplication is observed, which is consistent with two cAMP-binding domains. The two types of regulatory subunit from type I and type II kinase display similarities in domain substructure and in amino acid sequence, which provide a molecular basis for new insight into their regulatory roles. Detailed analyses of the homology of the regulatory subunits of type I and type II cAMP-dependent protein kinase and of similar relationships to cGMP-dependent protein kinase and *Escherichia coli* catabolite gene activator protein are presented in accompanying reports from this laboratory [Takio, K., Smith, S. B., Krebs, E. G., Walsh, K., & Titani, K. (1984) *Biochemistry* (second paper of three in this issue); Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1984) *Biochemistry* (third paper of three in this issue)].

The wide variety of physiological effects of adenosine cyclic 3′,5′-phosphate (cAMP) in eukaryotic organisms is mediated primarily by the phosphorylation of specific cellular proteins by cAMP-activated protein kinases (ATP:protein phosphotransferase, EC 2.7.1.37) (Walsh et al., 1968). The enzyme is widely distributed (Kuo & Greengard, 1969) and well recognized for its general role in the regulation of metabolism and other cellular functions. It is composed of two regulatory (R) and two catalytic (C) subunits that together constitute an inactive holoenzyme (R_2C_2) [reviewed by Krebs & Beavo (1979)]. When two molecules of cAMP bind to each regulatory subunit, active catalytic subunits dissociate from the tetramer, leaving a dimer of regulatory subunits with four molecules of bound cAMP.

In mammalian tissues, the enzyme is present in two isozyme forms, type I, the predominant form in skeletal muscle, and type II, the predominant form in cardiac muscle (Reimann et al., 1971; Rubin et al., 1972). These isozymes have identical substrate specificity but can be separated from each other on DEAE-cellulose, and the differences between them are attributed to differences in the structure of the R subunits (Rubin & Rosen, 1975; Corbin et al., 1975). The C subunits are thought to be identical (Hoffman et al., 1975; Zoller et al., 1979). The two isozymes have different tissue distributions (Corbin et al., 1975) and intracellular localizations (Krebs & Beavo, 1979) and vary in their relative amounts during development (Lee et al., 1976) and in the cell cycle (Costa et al., 1976), suggesting that they may have different physiological functions. The type I isozyme does not undergo the autophosphorylation characteristic of type II, but it can be phosphorylated in vitro in a slow reaction catalyzed by cGK,[1] the cGMP-dependent protein kinase (Geahlen & Krebs, 1980; Hashimoto et al., 1981).

A growing body of structural information is becoming available concerning the two types of cAMP-dependent protein kinase and a related enzyme, cGMP-dependent protein kinase. The amino acid sequences of both the catalytic and the regulatory subunits of bovine type II enzyme have been reported (Shoji et al., 1983; Takio et al., 1982, 1984a) and that of cGK is presented in an accompanying report (Takio et al., 1984b). In addition it has been demonstrated in several laboratories that the regulatory subunits of both isozymes are readily cleaved by limited proteolysis into separable amino-terminal and carboxy-terminal domains at "hinge" regions (Corbin et al., 1978; Potter & Taylor, 1979a,b; Flockhart et al., 1980; Takio et al., 1980, 1983). The site on the type I regulatory subunit (R_I) that is phosphorylated by cGMP-dependent protein kinase is in this hinge region (Hashimoto et al., 1981), as are the autophosphorylation sites on R_{II}, and cGK (Potter & Taylor, 1979b; Takio et al., 1980, 1983). The binding sites for cAMP on both R_I and R_{II} are located in their carboxyl-terminal domains (Rannels & Corbin, 1979; Kerlavage & Taylor, 1980).

The present report describes the complete amino acid sequence of bovine skeletal muscle regulatory subunit of type I. This 379 residue structure includes the 55-residue segment from the hinge region sequenced by Hashimoto et al. (1981) and indicates molecular aspects of the similarity of the type I regulatory subunit to its isozyme R_{II}, to the catabolite gene activator protein of *Escherichia coli* (CAP), and to cGK, the cGMP-dependent protein kinase of bovine lung.

Materials and Methods

The regulatory subunit of bovine skeletal muscle type I cAMP-dependent protein kinase was purified to homogeneity

† 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: Toyo Soda Co. Ltd., 2743-1 Hayakawa Ayase-shi, Kanagawa-Ken, Japan.

§ 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_I and R_{II}, regulatory subunits of cAMP-dependent protein kinases, types I and II; cGK, cGMP-dependent protein kinase; CM, *S*-carboxymethyl; SDS–PAGE, polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate; TPCK, tosylphenylalanine chloromethyl ketone; HPLC, high-performance liquid chromatography.

0006-2960/84/0423-4193$01.50/0

as described by Dills et al. (1979). α-Chymotrypsin, elastase, and TPCK–trypsin were obtained from Worthington, lima bean trypsin inhibitor was from Millipore, *Staphylococcus aureus* V8 protease was from Miles, endoproteinase Lys-C was from Boehringer, and thermolysin was from Calbiochem. The *Astacus* protease was a generous gift of Dr. Robert Zwilling, University of Heidelberg. Radiolabeled iodo[^{14}C]acetic acid and [^{14}C]methyl iodide were products of New England Nuclear Corp. L-Alanyl-L-serine and L-seryl-L-alanine (Sigma products) were N-acetylated with acetic anhydride in 50% pyridine and, after removal of excess reagents by lyophilization, purified by reversed-phase HPLC.

The protein was reduced and S-carboxymethylated as described by Koide et al. (1978). Citraconylation followed the procedure of Habeeb & Atassi (1970).

Methionyl residues were extrinsically radiolabeled by a methylation/demethylation procedure which facilitated the isolation of tryptic peptides by the procedure of Sasagawa et al. (1983). CM protein (100 nmol) was radiolabeled with [^{14}C]methyl iodide, digested with trypsin, lyophilized, and separated into fractions soluble and insoluble in 9% formic acid. The soluble fraction was further separated by reversed-phase HPLC. Radioactive fractions were then demethylated as described and purified by rechromatography to yield methionine-containing peptides. The acid-insoluble fraction was further digested with *S. aureus* protease, and constituent methionine-containing peptides were isolated in a similar manner.

The CM protein was cleaved at methionyl residues in 70% formic acid as described by Gross & Witkop (1962). With the following exceptions, enzymatic digestions were performed at 37 °C for 1–6 h in 0.1 M ammonium bicarbonate, pH 8.0. In the case of the protease from *S. aureus*, 0.05 M phosphate buffer, pH 7.8, was used overnight. The chymotryptic digest included 10 mM *p*-aminobenzamidine.

HPLC was performed on a Varian 5000 liquid chromatograph using a SynChropak RP-P (SynChrom) column with an acetonitrile gradient in aqueous trifluoroacetic acid (Mahoney & Hermodson, 1980). Amino acid analysis, sequenator analyses, and identifications of phenylthiohydantoins were performed as described previously (Shoji et al., 1983).

Mass spectrometric analysis was performed after permethylation. The blocked dipeptide and synthetic acetyl-L-alanyl-L-serine and acetyl-L-seryl-L-alanine (each approximately 10 nmol) were permethylated according to the method of Leclercq & Desiderio (1971) and purified by reversed-phase HPLC. The lyophilized sample was analyzed with a VG 7070H mass spectrometer in the chemical ioinization mode by William N. Howald (Laboratory of Medicinal Chemistry, University of Washington).

Alignments and comparisons of related sequences were accomplished with the computer program ALIGN of Dayhoff et al. (1983) as described (Takio et al., 1984a).

Results

General Strategy. Sequenator analysis of the intact CM protein (10 nmol) yielded no phenylthiohydantoins in 10 cycles of degradation, clearly indicating that the amino terminus of the protein is blocked. The blocking group was later identified to be acetyl by mass spectrometric analysis of a blocked dipeptide derived from the amino terminus of the CM protein.

Most of the sequence was obtained by analysis of eight nonoverlapping cyanogen bromide fragments (M1–M8) and their subpeptides. The remainder of the sequence and overlaps of the fragments were obtained largely by analysis of tryptic peptides generated by cleavage at arginyl bonds.

Our results (Figure 1), combined with the 55-residue sequence near the phosphorylation site (Hashimoto et al., 1981), provide the complete 379-residue sequence of the type I subunit.

Cleavage at Methionyl Bonds. Fifty milligrams of the [^{14}C]CM protein (1200 nmol) was cleaved with cyanogen bromide, and the mixture was separated into six fractions on a Sephadex G-50 superfine column (1.5 × 200 cm) in 9% formic acid as shown in Figure 2A. Fraction 1, and material insoluble in 9% formic acid prior to gel filtration, contained mainly fragment M7, which was further purified on a Sephacryl S-200 column (1.5 × 200 cm) in 7 M urea–10 mM H_3PO_4 (Figure 2B). Fractions 4 and 6 each contained two fragments, which were purified by reversed-phase HPLC as shown in Figure 2C,D. Some of fragment M4 eluted in an anomalous manner from the Sephadex G-50 column in acid. With its 29 residues, it would have been expected to elute in pooled fraction 5 with fragment M3, but some was recovered from pooled fraction 1, indicating extensive aggregation, and the rest from pooled fraction 6 where it coeluted with the nonapeptide M6.

Compositions and yields of the eight primary fragments thus isolated are listed in Table I, together with the amino acid analysis of the intact CM protein.

Cleavage at Arginyl Bonds. The [^{14}C]CM protein (600 nmol) was N-citraconylated at pH 8.8 in the presence of 7 M guanidine hydrochloride, dialyzed for 2 days against 0.1 M NH_4HCO_3 at the same pH, and then digested at 37 °C with TPCK–trypsin (276 μg) for 2 h. Digestion was terminated by the addition of lima bean trypsin inhibitor (276 μg). The mixture was further incubated at 37 °C for 15 min and then separated on a column of Sephadex G-50 superfine as shown in Figure 3. Each fraction was lyophilized and, after decitraconylation in 9% formic acid at 37 °C for 2 h, further purified on a SynChropak column by reversed-phase HPLC in a similar manner to those shown in Figure 2C,D. Pooled fraction 1 contained mainly R7 (residues 114–209) and was used for sequence analysis without further purification. Although peptides accounting for the entire protein were thus isolated, Figure 3 indicates only the locations of those used in the sequence analysis summarized in Figure 1.

Sequence Analysis. Of the cyanogen bromide fragments, only M1 was N^{α} blocked, as was the intact protein, indicating that this fragment was derived from the amino-terminal portion of the protein.

Various subpeptides of M1 were generated (in 50 nmol digests) with trypsin, chymotrypsin, *S. aureus* protease, elastase, and/or *Astacus* protease and analyzed to elucidate the blocked amino terminus. Trypsin generated four subpeptides, M1-T1 (residues 1–11; identical with R1) and three others corresponding to cleavage at lysine residues 22 and 30. Each was isolated by reversed phase HPLC. Only M1-T1 was N^{α} blocked. Elastase generated three peptides from M1-T1 or R1, of which only one, with the composition (Ala, Ser), was N^{α} blocked. Mass spectrometric analysis of this peptide (10 nmol), after permethylation, identified it as acetylalanylserine. This spectrum was essentially identical with that of synthetic acetyl-L-alanyl-L-serine but not with that of synthetic acetyl-L-seryl-L-alanine. Sequencer analysis of the other two elastase-derived peptides Gly-Thr and Thr-Ala-Ser-Glu-Glu-Glu-Arg provided the remainder of the sequence of R1 (or M1-T1). Although there is no direct evidence for overlap of the three elastase peptides, the blocked peptide must be amino terminal and the arginine-containing peptide must be carboxy terminal, completing the sequence of R1. Digestion of M1

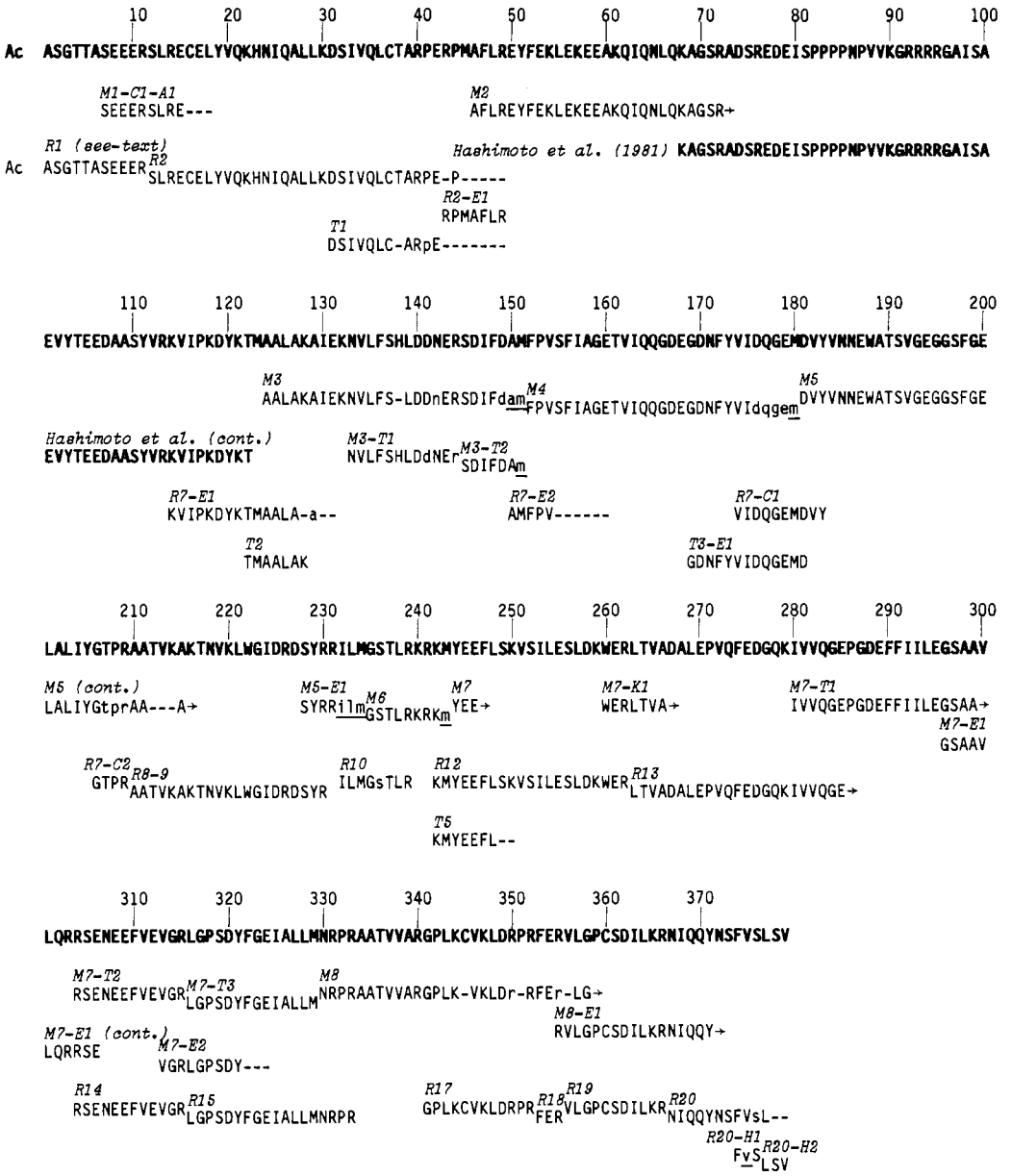

FIGURE 1: Summary of the sequence proof of the R_I subunit. The proven sequences of specific peptides (named in italics) are given in one-letter code below summary sequence (bold type). Prefixes M and R denote peptides generated by cleavage of the CM protein at methionyl and arginyl bonds, respectively. Their subpeptides, generated by digestion with *Astacus* protease, α-chymotrypsin, *S. aureus* protease, thermolysin, endoproteinase Lys-C, or trypsin, are designated by prefixes A, C, E, H, K, and T. Peptide sequences written in upper case letters are proven by Edman degradation; those in lower case letters indicate tentatively identified residues. Underlined lower case letters denote residues deduced only from compositional analysis. Those not identified in a peptide are shown by dashes. An arrow indicates that the peptide is much longer but that no further residues were identified by Edman degradation. Ac denotes an acetyl group.

by chymotrypsin generated another blocked peptide, M1-C1 (residues 1–19), which was further digested with *Astacus* protease to generate M1-C1-A1 (residues 7–19). Analysis of this peptide provided an overlap for R1 and R2.

Analysis of R2 yielded the sequence of residues 12–42 and identified residue 44 as proline. Analysis of R2-E1, generated by digestion of R2 with *S. aureus* protease and isolated by reversed-phase HPLC, completed the sequence of R2 and provided an overlap for M1 and M2.

Analysis of M2 yielded the sequence of residues 46–72, which overlapped by 5 residues the 55-residue sequence previously reported by Hashimoto et al. (1981) as the region including the site (Ser-99) phosphorylated by cGMP-dependent protein kinase. The composition of M2 was compatible with the sequence of residues 46–123, placing the 55-residue segment of Hashimoto et al. at residues 68–122 and tentatively placing methionine at residue 123. This placement was confirmed, and the "Hashimoto segment" overlapped to M3 by analysis of peptide R7-E1, isolated by HPLC from an *S. aureus* protease digest of R7.

R7 was the largest peptide in the tryptic series and appeared to be derived from residues 114–209. Although sequencer analysis of the peptide yielded the sequence beginning with Lys-Val-Ile (residues 114–116), it became rapidly ambiguous, due to its size and lack of purity (only 70–80%). Digestion of R7 with *S. aureus* protease or with chymotrypsin yielded, by reversed-phase HPLC, three methionine-containing peptides, R7-E1 (above), R7-E2, and R7-C1, and the carboxyl-terminal tetrapeptide, R7-C2. Peptides R7-E2 and R7-C1 provided overlaps for M3–M5. M3 and M4 were each sequenced almost to the carboxyl terminus, although with some ambiguity, which was later clarified by analysis of subpeptides.

Sequencer analysis of M5 extended the sequence to residue 211, although the sequence Thr-Pro-Arg at residues 207-209

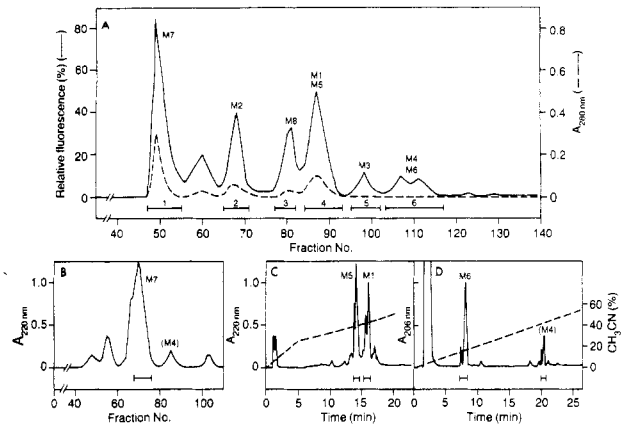

FIGURE 2: Separation of a cyanogen bromide digest of the CM protein (1200 nmol). (A) Primary fractionation on a column (1.5 × 200 cm) of Sephadex G-50 superfine in 9% formic acid. Fractions of 2 mL were collected at a flow rate of 10 mL/h and pooled as indicated by horizontal bars. (B) Further purification of fraction 1 on a column (1.5 × 200 cm) of Sephacryl S-200 in 7 M urea–10 mM H_3PO_4. Fractions of 2 mL were collected at a flow rate of 8 mL/h and pooled as indicated by horizontal bars. (C and D) Further purification of fractions 4 and 6 in (A) on a SynChropak RP-P column (4.1 × 240 mm) by reversed-phase HPLC.

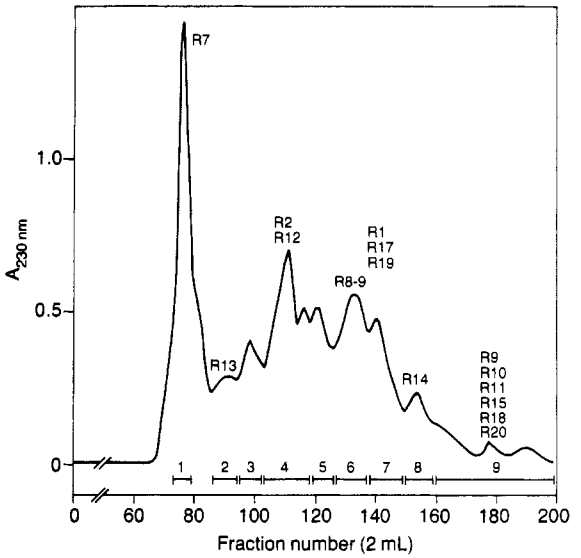

FIGURE 3: Primary fractionation of a tryptic digest of citraconylated CM protein (600 nmol) on a column (1.5 × 200 cm) of Sephadex G-50 superfine in 0.1 M NH_4HCO_3 adjusted to pH 8.8 with NH_4OH.

was only tentatively assigned. This sequence was confirmed by analysis of R7-C2, which was derived from the carboxyl terminus of R7 by chymotrypsin. The carboxyl-terminal sequence of M5 was provided by analysis of R8-9, beginning with Ala-Ala----- (residues 210–230), and M5-E1 generated by *S. aureus* protease. The latter was not retained well in the Sequencer cup, and the carboxyl-terminal portion of M5 (Ile, Leu) Met is based on its composition. Thus, the sequence of R10 appears to link M5 to M6.

M6 was sequenced almost to its carboxyl terminus and linked to M7 by analysis of R12. M7 was not homogeneous but contaminated with 0.3–0.4 mole fraction of a fragment overlapping M6 and M7 (residues 235–329) as estimated by observation of the minor sequence Gly-Ser-Thr---, which interfered with extensive analysis of the major sequence beginning with Try-Glu-Glu-----. The amino-terminal sequence of M7 was confirmed by analysis of R12. The remainder of the sequence of M7 was provided by analysis of subpeptides generated by digestion with trypsin, *S. aureus* protease, or endoproteinase Lys-C and isolated by reversed-phase HPLC. M7-K1, a subpeptide generated with the endoproteinase Lys-C, provided the overlap for R12 and R13. R15 confirmed the carboxyl-terminal sequence of M7 identified by analysis of M7-T3 and overlapped by four residues the carboxyl-terminal fragment M8.

Analysis of M8 yielded the sequence of residues 330–358, although several identifications were only tentative. Analyses of R17, R18, and R19 removed these uncertainties. Analysis of M8-E1 (a subpeptide generated by digestion of M8 with *S. aureus* protease) extended the sequence to tyrosine-371. The carboxyl-terminal portion of M8 was difficult to determine. Repeated analysis of R20 extended the sequence to Leu-377, but Ser-376 was not clearly identified, and the peptide washed out of the Sequencer cup very rapidly. R20 was then digested with thermolysin to generate R20-H1 and R20-H2. Analysis of R20-H1 confirmed that residue 376 was serine. Analysis of R20-H2 identified the carboxyl-terminal sequence of M8 as Leu-Ser-Val. Acid hydrolysis of R20 for 96 h yielded only two valyl residues per mol on the basis of one leucyl residue per mol, ruling out the possibility of a carboxyl-terminal Val-Val. Since R20 is the only peptide lacking arginine in the tryptic digest of citraconylated protein, it is assumed to represent the carboxyl terminus of the protein. This is consistent with the analogous unique lack of homoserine in M8. However, the composition of M8 is poor, especially in its threonine, glycine, and lysine residues. Thus, the carboxyl terminus of the whole protein is inferred to be Leu-Ser-Val, but the rigor of the proof is marginal.

Finally, to confirm the alignment of the eight CNBr fragments, the methionine residues of the CM protein were extrinsically ^{14}C-radiolabeled, and acid-soluble tryptic peptides (T1–T7) were isolated by reversed-phase HPLC (Sasagawa et al., 1983) and analyzed. T1 appeared to correspond to residues 31–49, confirming the overlap for M1 and M2. Similarly, T2 overlapped M2 and M3. T4 and T7 were identical with R10 and R15, respectively, corroborating the alignment of M5–M6 and of M7–M8. T5 and T6 overlapped each other (T6 lacked lysine 242) as well as M6 and M7. A large, acid-insoluble peptide, T3 (residues 145–214), was subdigested with *S. aureus* protease to yield T3-E1, which overlapped M4 and M5 (by one residue). The overlap of M3 and M4 was not recovered from this digest.

Together, these data provide a unique alignment of the 379 residues in the regulatory subunit. As is usual in a proof of sequence, some regions are established with greater certainty than others; the reliability of this proof is discussed below.

Discussion

In general, the proof of the structure of subunit R_I of bovine skeletal muscle cAMP-dependent protein kinase involved two primary sets of overlapping peptides derived by cleavage at either arginyl or methionyl residues. A special problem was presented by the acetylated amino terminus whereas the prior knowledge of the 55-residue interior sequence, which includes the site of phosphorylation by cGMP-dependent protein kinase (Hashimoto et al., 1981), facilitated resolution of the total sequence. The present study also benefited from an application of the method of Sasagawa et al. (1983) to specifically isolate methionine-containing peptides by hydrophobicity modulation. As in studies of the catalytic and regulatory subunits of the type II enzyme (Shoji et al., 1983; Takio et al., 1982, 1984a), every phenylthiohydantoin was identified in two complementary HPLC systems, and more than half of the sequence was confirmed by redundant analyses of peptides from complementary sets.

Table I. Amino Acid Compositions[a] of Peptides Isolated after Cleavage with CNBr

Peptide		M1	M2	M3	M4	M5	M6	M7	M8	Total from sequence	Total from amino acid analysis of R_I
Residues		1-45	46-123	124-151	152-180	181-234	235-243	244-329	330-379		
Asp/Asn	(D/N)	2.4 (2)	6.7 (6)	5.8 (6)	3.9 (4)	5.8 (6)		7.5 (6)	5.4 (5)	(35)	31.5
Thr	(T)	2.7 (3)	2.4 (2)	0.6 (0)	1.0 (1)	3.5 (4)	0.9 (1)	2.1 (1)	1.7 (1)	(13)	13.7
Ser	(S)	3.5 (4)	5.4 (5)	2.4 (2)	1.3 (1)	2.9 (3)	0.9 (1)	5.7 (6)	3.9 (4)	(26)	23.6
Glu/Gln	(E/Q)	8.6 (9)	13.0 (13)	2.7 (2)	5.7 (6)	3.3 (3)		16.3 (18)	2.9 (3)	(54)	55.6
Pro	(P)	2.0 (2)	5.9 (6)		0.8 (1)	1.1 (1)		3.6 (3)	3.9 (4)	(17)	17.3
Gly	(G)	1.3 (1)	4.1 (3)	1.1 (0)	4.3 (4)	5.7 (6)	1. (1)	7.8 (7)	2.9 (2)	(24)	26.3
Ala	(A)	4.0 (4)	8.1 (8)	4.7 (5)	1.4 (1)	4.8 (5)		6.2 (5)	3.0 (3)	(31)	32.0
1/2Cys	(C)	1.6 (2)							1.2 (2)	(4)	3.7
Val	(V)	1.9 (2)	4.4 (5)	1.6 (1)	2.3 (3)	4.8 (5)		7.8 (8)	4.9 (6)	(30)	28.8
Met	(M)[b]	0.4 (1)	0.3 (1)	0.5 (1)	0.4 (1)	0.3 (1)	0.3 (1)	0.2 (1)	0.0 (0)	(7)	6.6
Ile	(I)	1.9 (2)	3.8 (4)	1.9 (2)	2.6 (3)	2.8 (3)		4.6 (5)	2.1 (2)	(21)	18.5
Leu	(L)	5.0 (5)	3.0 (3)	3.0 (3)	0.4 (0)	4.0 (4)	1.0 (1)	10.0 (10)	5.0 (5)	(31)	31.0
Tyr	(Y)	0.9 (1)	3.6 (4)		1.0 (1)	2.7 (3)		2.5 (2)	1.2 (1)	(12)	10.5
Phe	(F)		2.4 (2)	2.0 (2)	3.0 (3)	1.3 (1)		6.0 (6)	2.1 (2)	(16)	16.0
His	(H)	1.0 (1)		1.0 (1)						(2)	2.2
Lys	(K)	2.0 (2)	7.7 (8)	2.4 (2)		3.0 (3)	2.0 (2)	3.3 (3)	4.0 (3)	(23)	21.7
Arg	(R)	4.1 (4)	8.3 (8)	1.5 (1)		4.2 (4)	2.0 (2)	4.4 (4)	6.9 (7)	(30)	31.2
Trp	(W)[c]	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(0)	N.D.(2)	N.D.(0)	N.D.(1)	N.D.(0)	(3)	N.D.
No. of Residues		45	78	28	29	54	9	86	50	(379)	
% Yield		35	65	40	15	40	65	50	35		

[a]Residues per peptide by amino acid analysis (values < 0.3 not reported except Met) or, in parentheses, from the sequence (Figure 1). Ratios are calculated to the integral values underlined.

[b]Methionine measured as homoserine.

[c]Not determined.

In spite of the rigor of the approach, there are regions of relatively weaker analytical proof such as the overlaps aligning M1 and M2 (Met-45), M3 and M4 (Met-150), and M6 and M7 (Met-243). Marginal assignments of sequence are also seen in the Thr-Pro-Arg placement at residues 208–210 and of the Ile-Leu-Met placement at residues 232–234. Finally, the proof of the carboxyl-terminal pentapeptide is somewhat cumbersome. Fortunately, Lee et al. (1983) have recently reported the isolation of a cDNA clone for the type I bovine regulatory subunit, and their unpublished cDNA sequence analysis confirms our assignments in each of the marginal areas identified above (D. C. Lee, H. Knickerbocker, and G. S. McKnight, personal communication).

The molecular weight of subunit R_I, calculated from the amino acid sequence (excluding the phosphoryl group), is 42 804. Analysis by SDS-PAGE had indicated a molecular weight of 47 000–48 000 for porcine, rabbit, and bovine subunits R_I (Potter et al., 1978; Hofmann et al., 1975; Hashimoto et al., 1981). It should be noted that similar discrepancies between measured and calculated molecular weights were seen with R_{II} (Takio et al., 1982) and ascribed to the asymmetric shape of the R_{II} dimer and a tendency to distort the determination of mass, even in the presence of SDS. A similar explanation may hold for the closely related R_I dimer.

On the basis of a short sequence reported to occur at the hinge region of porcine R_I (Potter & Taylor, 1979a), one can deduce that the porcine and bovine subunits are homologous (Figure 4A). Except for an unidentified porcine residue corresponding to Ser-110, the 15-residue sequence in the bovine protein (Residues 94–112) is identical with that in the porcine protein. Thus, it may be reasonable to correlate the function of the porcine enzyme with the amino acid sequence of the bovine enzyme. For example, it is reported that the porcine subunit and its proteolytically generated amino-terminal domain both behave as disulfide-linked dimers on SDS–PAGE (Taylor et al., 1981). Cysteine-16 or cysteine-37 in the bovine sequence could provide this linkage. However, it is not clear whether the disulfide is of physiological significance because it would be expected to be reduced during isolation of the enzyme in its R_2C_2 form. Conversely, a disulfide could be generated from cysteine residues under the conditions of limited proteolysis.

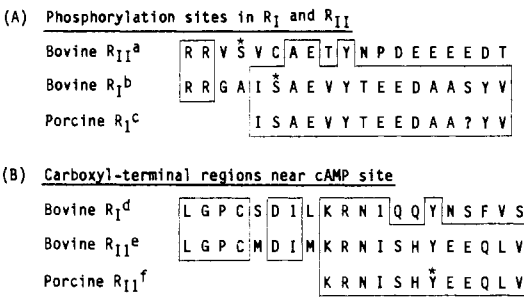

FIGURE 4: Homologous sequences in two regions of bovine and/or porcine R_I and R_{II}. (A) Sites of phosphorylation (S*) in bovine R_{II} and R_I aligned with a nonphosphorylated peptide from the hinge region of porcine R_I. (B) Carboxyl-terminal segment of bovine R_I aligned with a sequence 11 residues from the carboxyl terminus of bovine R_{II} and with a peptide from porcine R_{II} which was photoaffinity labeled (Y*) with a cAMP analogue. Boxes enclose identical residues. Superscripts to R_I and R_{II} identify sources of data, as follows: (a) Takio et al. (1980, 1982); residues 92–110. (b) Hashimoto et al. (1981); this work; residues 94–112. (c) Potter & Taylor (1979a). (d) This work; residues 357–379. (e) Takio et al. (1982), residues 367–389. (f) Kerlavage & Taylor (1980).

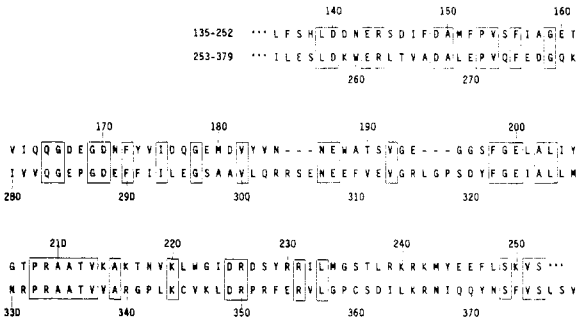

FIGURE 5: Internal sequence homology in R_I. Residues 135–252 are aligned with residues 253–379. Identical residues are enclosed in boxes; (---) gaps placed to optimize the homology; (···) indicates that the sequence continues.

It has been shown previously that the site of phosphorylation of bovine R_I by cGMP-dependent protein kinase is at a serine (Hashimoto et al., 1981) which can now be placed at residue 99. That report, and a later one from the same group (Takio et al., 1983), suggested that the surrounding sequence in R_I may be homologous to the corresponding sites of autophosphorylation in bovine R_{II} and in cGK. The suggested homology of the two isozymes and cGK within the same species of animal is less apparent than the virtual identity of the corresponding regions in bovine and porcine R_I (Figure 4A), indicating that the distinction between type I and type II kinases is a more highly conserved feature of these enzymes than might have been expected.

As in the case of bovine R_{II}, it is appropriate to consider subunit R_I in terms of three constituent domains, an amino-terminal 134 residue domain, which includes the sites associated with dimerization and phosphorylation (Hashimoto et al., 1981), and two other domains, which have characteristics of tandem internal gene duplication. This pair of segments in R_I is aligned in Figure 5 and reveals 37% identity and an alignment score of 11.7, strongly indicative of internal homology. It is well-known that these two domains are easily cleaved by limited proteolysis from the amino-terminal domain and that the paired domains exclusively bind 2 mol of cAMP (Rannels & Corbin, 1979; Potter & Taylor, 1980; Weber & Hilz, 1979). It is reasonable to assume that each of the domains within the pair binds a single cAMP. In these respects, subunit R_I resembles that from the type II isozyme (Takio et al., 1982), and Takio et al. (1984b) have shown that the similarity extends to cGK. Attempts to label in R_I a binding site for cAMP with an azido analogue of [^{32}P]cAMP (S. B. Smith and K. Titani, unpublished experiments) indicate that both M8 and R20 are labeled and provide preliminary evidence that tyrosine-371 is the site of that label. Analogous photoaffinity labeling of porcine R_{II} (Kerlavage & Taylor, 1980) identified a tyrosyl residue in a sequence that appears to be homologous to bovine R_I (Figure 4B). This is of some interest because it focuses attention on the carboxyl-terminal regions of all three subunits which may be in juxtaposition to the bound cAMP.

It is evident that there are parallels not only in the binding and regulatory functions of R_I and R_{II} but also in their domain organization and in their amino acid sequences. An analysis of the extent of homology between these two subunits is outlined in an accompanying report (Takio et al., 1984a), which compares the sequences of the three hypothetical domains in R_I with the corresponding domains in R_{II} which, in turn, are homologous with cGK. It was found that there is ample evidence of homology between R_I and R_{II} in their middle and carboxyl-terminal paired domains (residues 135–379 in R_I), whereas alignments of the amino-terminal domains are less convincing. A detailed analysis of the homology of R_I with cGMP-dependent protein kinase is reported by Takio et al. (1984b) as part of a demonstration of homologous alignments of seven related protein sequences. Since it is known that the cAMP-binding domains of R_{II} are homologous with the cAMP-binding catabolite gene activator protein from *E. coli* (Weber et al., 1982), it is apparent that the homologous relationship will extend to both domains of R_I and to similar domains within cGMP-dependent protein kinase (Titani et al., 1983; Takio et al., 1984b)

Finally, it should be mentioned that the differences between the amino acid sequences of R_I and R_{II} are easily sufficient to account for differences in physical and chemical behavior of the type I and type II isozymes. Zoller et al. (1979) have concluded that the catalytic subunits of the two porcine isozymes were "very homologous although apparently not identical". A preliminary comparison of tryptic fingerprints of the catalytic subunits derived from the two bovine enzymes in our laboratory (G. Marchildon and T. Sasagawa, unpublished data) indicates no discernible difference between them. Thus, it appears that a complete description of the molecular difference between the two isozymes lies in the difference between the amino acid sequences of subunits R_I and R_{II}.

Acknowledgments

We are particularly indebted to Dr. Koji Takio for valuable discussions throughout this study, to Dr. G. S. McKnight for access to the cDNA sequence data prior to publication, to Roger Wade, Floyd Kennedy, and Barbara Flug for able technical assistance in various aspects of this work and to Mary Woods for typing the manuscript.

References

Corbin, J. D., Keely, S. L., & Park, C. R. (1975) *J. Biol. Chem. 250*, 218.

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

Costa, M., Gerner, E. W., & Russell, D. H. (1976) *J. Biol. Chem. 251*, 3313.

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

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.

Geahlen, R. L., & Krebs, E. G. (1980) *J. Biol. Chem. 255*, 1164.

Gross, E., & Witkop, B. (1962) *J. Biol. Chem. 237*, 1856.

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

Hashimoto, E., Takio, K., & Krebs, E. G. (1981) *J. Biol. Chem. 256*, 5604.

Hofmann, F., Beavo, J. A., Bechtel, P. J., & Krebs, E. G. (1975) *J. Biol. Chem. 250*, 7795.

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

Koide, A., Titani, K., Ericsson, L. H., Kumar, S., Neurath, H., & Walsh, K. A. (1978) *Biochemistry 17*, 5657.

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

Kuo, J. F., & Greengard, P. (1969) *J. Biol. Chem. 244*, 3417.

Leclercq, P. A., & Desiderio, D. M., Jr. (1971) *Anal. Lett. 4*, 305.

Lee, D. C., Carmichael, D. F., Krebs, E. G., & McKnight, G. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3608.

Lee, P. C., Radloff, D., Schweppe, J. S., & Jungmann, R. A. (1976) *J. Biol. Chem. 251*, 914.

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

Potter, R. L., & Taylor, S. S. (1979a) *J. Biol. Chem. 254*, 2413.

Potter, R. L., & Taylor, S. S. (1979b) *J. Biol. Chem. 254*, 9000.

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

Potter, R. L., Stafford, P. H., & Taylor, S. S. (1978) *Arch. Biochem. Biophys. 190*, 174.

Rannels, S. R., & Corbin, J. D. (1979) *J. Biol. Chem. 254*, 8605.

Reimann, E. M., Walsh, D. A., & Krebs, E. G. (1971) *J. Biol. Chem. 246*, 1986.

Rubin, C. S., & Rosen, O. M. (1975) *Annu. Rev. Biochem. 44*, 831.

Rubin, C. S., Erlichman, J., & Rosen, O. M. (1972) *J. Biol. Chem. 247*, 36.

Sasagawa, T., Titani, K., & Walsh, K. A. (1983) *Anal. Biochem. 128*, 371.

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., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1984a) *Biochemistry* (second paper of three in this issue).

Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1984b) *Biochemistry* (third paper of three 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.

Walsh, D. A., Perkins, J. P., & Krebs, E. G. (1968) *J. Biol. Chem 243*, 3763.

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.

Zoller, M. J., Kerlavage, A. R., & Taylor, S. S. (1979) *J. Biol. Chem. 254*, 2408.