Homology of the γ Subunit of Phosphorylase *b* Kinase with cAMP-Dependent Protein Kinase†

Erwin M. Reimann, Koiti Titani, Lowell H. Ericsson, Roger D. Wade, Edmond H. Fischer, and Kenneth A. Walsh*

ABSTRACT: The complete amino acid sequence of the catalytic subunit (γ subunit) of rabbit skeletal muscle phosphorylase *b* kinase was determined. The γ subunit was purified by gel filtration in acidic 8 M urea after reduction and S-carboxymethylation in 7 M guanidine hydrochloride. Cleavage of the γ subunit at arginyl bonds gave a complete set of nonoverlapping peptides. Overlapping peptides were obtained by cleavage at methionyl, tryptophanyl, or glutamyl bonds and by selected subdigestion of two large peptides obtained by cleavage at methionyl bonds. Sequence analysis established that the protein contains 386 residues corresponding to a molecular weight (M_r) of 44 673. Comparison of the γ subunit with the catalytic subunit of bovine cAMP-dependent protein kinase and with tyrosine-specific kinases of viral origin revealed a significant degree of sequence identity among all of these proteins. These data suggest that calcium-dependent protein kinases may share a common ancestral gene and a common structural basis for catalytic function with a wide variety of other protein kinases which respond to different signals and control quite different processes.

A variety of protein kinases play important regulatory roles in the control of cellular function (Krebs & Beavo, 1979). These enzymes act as transducers that respond to chemical signals (e.g., the "second messengers", cAMP[1] or Ca^{2+}) by catalyzing the phosphorylation of key enzymes, which in turn control metabolic or physiological processes.

In a few cases, chemical details of the protein kinases have been presented. The prototype enzyme cAMP-dependent protein kinase, which catalyzes the phosphorylation of specific serine (and threonine, but not tyrosine) residues in its protein substrates, has been characterized in great detail at the chemical level (Shoji et al., 1983; Taylor et al., 1981; Takio et al., 1982). cGMP-dependent protein kinase has an analogous function [cf. review by Flockhart & Corbin (1982)], and an accompanying report indicates that it has a homologous structure (Takio et al., 1984). Recently, Barker & Dayhoff (1982) have shown that $pp60^{src}$, an oncogene product of the avian sarcoma virus that is known to be a tyrosine-specific protein kinase (Hunter & Sefton, 1980), bears a homologous relationship to the catalytic subunit of cAMP-dependent protein kinase in spite of large differences in both size and oligomeric nature.

Other protein kinases respond to calcium ions [cf. review by Flockhart & Corbin (1982)]. Of these, phosphorylase kinase (Malencik & Fischer, 1982) and myosin light chain kinase (Adelstein & Eisenberg, 1980) have been well characterized. They differ in their molecular weight and in their substrate specificity, but share with the cyclic nucleotide-dependent kinases the ability to catalyze the phosphorylation of serine residues in specific protein substrates. Detailed structural information has not been available to assess any possible relationship between the calcium-dependent kinases on one hand and either the cyclic nucleotide-dependent kinases or the tyrosine-specific kinases on the other. We chose to look for such a relationship in one of the more available calcium-dependent protein kinases, rabbit skeletal muscle phosphorylase *b* kinase, by determining the complete amino acid sequence of the γ subunit, which has been shown to have catalytic activity of its own (Skuster et al., 1980).

Phosphorylase *b* kinase resembles cAMP-dependent protein kinase in its ability to catalyze the phosphorylation of glycogen synthase at serine residue 7 (Embi et al., 1979, 1981) as well as its own α and β subunits and troponin I [cf. review by Carlson et al. (1979)]. However, the two kinases are clearly different with regard to phosphorylation of several other substrates (Carlson et al., 1979) and in their oligomeric structures. Whereas cAMP-dependent protein kinase is a 1.7×10^5 dalton tetramer consisting of two regulatory subunits and two catalytic subunits, phosphorylase *b* kinase has an M_r of 1.3×10^6 and contains four copies each of four types of subunits designated α, β, γ, and δ. The molecular weights of the subunits have been reported to be 118 000–146 000 for α, 108 000–128 000 for β, and 41 000–45 000 for γ (Hayakawa et al., 1973a; Cohen, 1973). The δ subunit has been shown to be identical with calmodulin (M_r = 16 680) and presumably confers calcium dependence on the enzyme (Cohen et al., 1978; Grand et al., 1981). The α and β subunits serve as substrates for kinases (Hayakawa et al., 1973b; Cohen, 1973) and thus have regulatory functions. Whereas activation of the cAMP-dependent protein kinase by cAMP involves dissociation of the enzyme into regulatory and catalytic subunits (Krebs & Beavo, 1979), phosphorylase kinase does not dissociate during activation (Carlson et al., 1979). Disaggregation of the enzyme in the presence of LiBr yielded αγδ and γδ complexes, both of which have catalytic activity (Skuster et al., 1980; Chan & Graves, 1982). The isolated γδ complex has a specific activity 5.6-fold greater than that of the native form, demonstrating that γ is in fact a catalytic subunit (Skuster et al., 1980). However, there is evidence that the β-subunit may also have catalytic activity [cf. reviews by Malencik & Fischer (1982) and Carlson et al. (1979)].

We report here the isolation of the γ subunit and the determination of its complete amino acid sequence. A large

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195 (K.T., L.H.E., R.D.W., E.H.F., and K.A.W.), and the Department of Biochemistry, Medical College of Ohio, Toledo, Ohio 43699 (E.M.R.). *Received February 24, 1984.* This work was supported by grants from the National Institutes of Health (GM 15731, AM 19231, and AM 07902), the Muscular Dystrophy Association (E.H.F.), and a Research Career Development Award (AM 00446 to E.M.R.).

[1] Abbreviations: HPLC, high-performance liquid chromatography; cAMP, adenosine cyclic 3′,5′-phosphate; cGMP, guanosine cyclic 3′,5′-phosphate; CMCys, (carboxymethyl)cysteine; $pp60^{src}$, transforming phosphoprotein of Rous sarcoma virus; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; TPCK, N^{α}-tosylphenylalanine chloromethyl ketone.

0006-2960/84/0423-4185$01.50/0

middle portion of the protein is shown to be remarkably similar to both the cAMP-dependent, serine-specific protein kinases and an oncogene-derived tyrosine-specific protein kinase. This homology is discussed in terms of presumed divergence from a common ancestor and of a common structural basis for the diverse processes that these enzymes regulate.

Materials and Methods

Preparation of the γ Subunit. Phosphorylase kinase was prepared from rabbit skeletal muscle essentially as described by Cohen (1973) except that acid precipitation was at pH 5.8 instead of pH 6.1 and DE52 was replaced by DEAE-Sephacel (Pharmacia). Prior to isolating the γ subunit, 50–120 nmol of kinase was reduced by incubation at room temperature for 1 h in 7 M guanidine hydrochloride (Heico), 32 mM dithiothreitol, 0.5 M Tris, and 13 mM EDTA, pH 8.6. Iodoacetic acid (72 mM, Sigma) was then added, and incubation was continued for 20 min in the dark to S-carboxymethylate the kinase. The mixture was adjusted to pH 3.3 and fractionated by gel filtration on Sephacryl S-300 (Pharmacia, 2.5 × 115 cm column) equilibrated in 8 M urea and 0.1 M H_3PO_4, pH 3.3 (adjusted with NH_4OH). The γ subunit emerged at 0.5 bed volume and was free of α, β, and δ subunits, as determined by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (Laemmli, 1970). Urea was removed immediately by dialysis or by gel filtration, and the protein was lyophilized. Prior to amino acid analysis the S-carboxymethylated γ subunit was further purified by reversed-phase HPLC.

Selective Cleavage. Arginyl bonds were cleaved by tryptic digestion of N-citraconylated protein with TPCK–trypsin (Worthington). [Trypsin was re-treated with N^{α}-tosylphenylalanine chloromethyl ketone (Carpenter, 1967) before use.] The γ subunit (130 nmol) was N-citraconylated (Atassi & Habeeb, 1972) in 2.4 mL of 7 M guanidine hydrochloride and 0.1 M Na_2HPO_4, pH 8.8, by the addition of 10 μL of citraconic anhydride over a period of 1 h while maintaining the pH at 8.8 on a pH stat at room temperature. The sample was dialyzed against 0.1 M NH_4HCO_3 (pH 8.8) and then treated with 40 μg of trypsin for 2 h at 40 °C. The reaction was terminated by adding 33 μg of lima bean trypsin inhibitor (Millipore). The mixture of peptides was fractionated immediately by gel filtraion on Sephadex G-50 at pH 8.9.

Glutamyl bonds were cleaved by digestion of N-succinylated (Yaoi et al., 1964) protein with *Staphylococcus aureus* V8 protease (Miles). (Succinylation increases the solubility of the γ subunit.) The protein (80 nmol) was incubated with the protease (1/100 w/w) for 26 h at 37 °C in 1 mL of 0.1 M NH_4HCO_3 at pH 8.0.

Methionyl bonds were cleaved by incubating 120 nmol of protein at 20 °C for 6 h in 1.0 mL of 70% formic acid containing 60 mg of CNBr (Gross, 1967). A large fragment from the cyanogen bromide digest, M11, was succinylated to increase its solubility and then digested with chymotrypsin (Worthington, treated with N^{α}-tosyllysine chloromethyl ketone) at pH 8 in 0.1 M NH_4HCO_3 and 10 mM *p*-aminobenzoic acid at 37 °C for 16 h. Cleavage at lysine residues in peptide M1 from the cyanogen bromide digest was accomplished by incubation of a mixture of M1 and M11 at 40 °C with endoproteinase Lys-C (Boehringer) at pH 8 in 0.1 M NH_4HCO_3 and 5 M urea. (Urea increases the solubility of M1.) Tryptophanyl bonds were cleaved by incubating 100 nmol of the γ subunit in 0.6 mL of 80% acetic acid at 20 °C in the dark with 1.5 mg of phenol and 1.7 mg of the skatole reagent (Pierce) described by Fontana et al. (1973). After 15 h the sample was diluted with 0.6 mL of H_2O, the excess reagent extracted twice with 0.6 mL of chlorobutane, and the aqueous phase lyophilized. Selective cleavage at aspartyl–proline bonds was achieved by incubating peptide M11 at 37 °C for 72 h in 70% formic acid (Landon, 1977).

Peptide Isolation. The four primary digests of the intact chain were fractionated first by gel filtration. The separation system varied: for cyanogen bromide digests, Sephadex G-50 superfine (Pharmacia) was used with 9% HCOOH or with 10 mM H_3PO_4 and 7 M urea; for digests with trypsin or *S. aureus* protease, Sephadex G-50 was used with 0.1 M NH_4HCO_3. In other cases (e.g., following cleavage at tryptophan) the primary separation was on an UltroPak TSK G2000SW high-pressure gel filtration column (LKB) in 6 M guanidine hydrochloride and 0.1 M NaH_2PO_4, pH 6.0. Subsequent purifications of fractions from gel filtration were achieved by reversed-phase HPLC on a SynChropak RP-P column (SynChrom, Inc., Linden, IN) using 0.1% trifluoroacetic acid in H_2O as solvent A and 0.08% trifluoroacetic acid in acetonitrile as solvent B. Separations were obtained with a flow rate of 2 mL/min and a linear gradient of 0–60% solvent B in 30–60 min. In some cases the peptides were dissolved in 6 M guanidine hydrochloride prior to HPLC.

In addition to the purification of primary fragments from the whole chain, several peptides were isolated after subdigestion of cyanogen bromide fragments M1 and M11. Reversed-phase HPLC was used to isolate peptide M1-K1 from those generated by treating a mixture of M1 and M11 with endoproteinase Lys-C. After cleavage of peptide M1 at arginyl residues, peptide R3 was isolated by gel filtration of the mixture on Sephadex G-50 equilibrated in 0.1 M NH_4HCO_3. After cleavage of the aspartyl–proline bond in M11, peptide M11-H1 was separated from the other fragment and from uncleaved M11 by gel filtration on Sephadex G-50 in 9% HCOOH. The chymotryptic peptide M11-C1, derived from M11, was isolated by reversed-phase HPLC.

Analytical Methods. Amino acid compositions were determined on a Dionex amino acid analyzer (Model D-500) after acid hydrolysis for 18 h unless otherwise noted. Automated sequence analyses were performed on a Beckman sequencer (Model 890C) in the presence of polybrene (Pierce) by the method of Edman & Begg (1967) with the Quadrol program of Brauer et al. (1975), except for peptide M11-C1, which was analyzed on a gas-phase instrument (Applied Biosystems, Inc.). Phenylthiohydantoin derivatives of amino acids were identified by reversed-phase HPLC using two complementary systems (Ericsson et al., 1977; Bridgen et al., 1976).

Search for Homology. The data bank of the National Biomedical Research Foundation (2222 sequences, June 1983 revision) was searched for homologous sequences with the SEARCH program described by Dayhoff (1979). Optimum alignments were then sought with the ALIGN program using the mutation data scoring matrix (MD + 6) and a gap penalty of 10. Alignment scores were calculated by using the ALIGN program as above. Scores are expressed according to the equation:

$$\text{alignment score} = (R - M)/(\text{SD})$$

where R = the real score for comparison of a pair of aligned sequences, M = the mean of 50 scores of randomly generated sequences of the same composition, and SD = the standard deviation from the mean of the randomly generated scores. Given a normal distribution of random scores, an alignment score of 2 corresponds to a probability of 0.023 that random chance would produce a better score. A score of 3 corresponds to a probability of 0.0014.

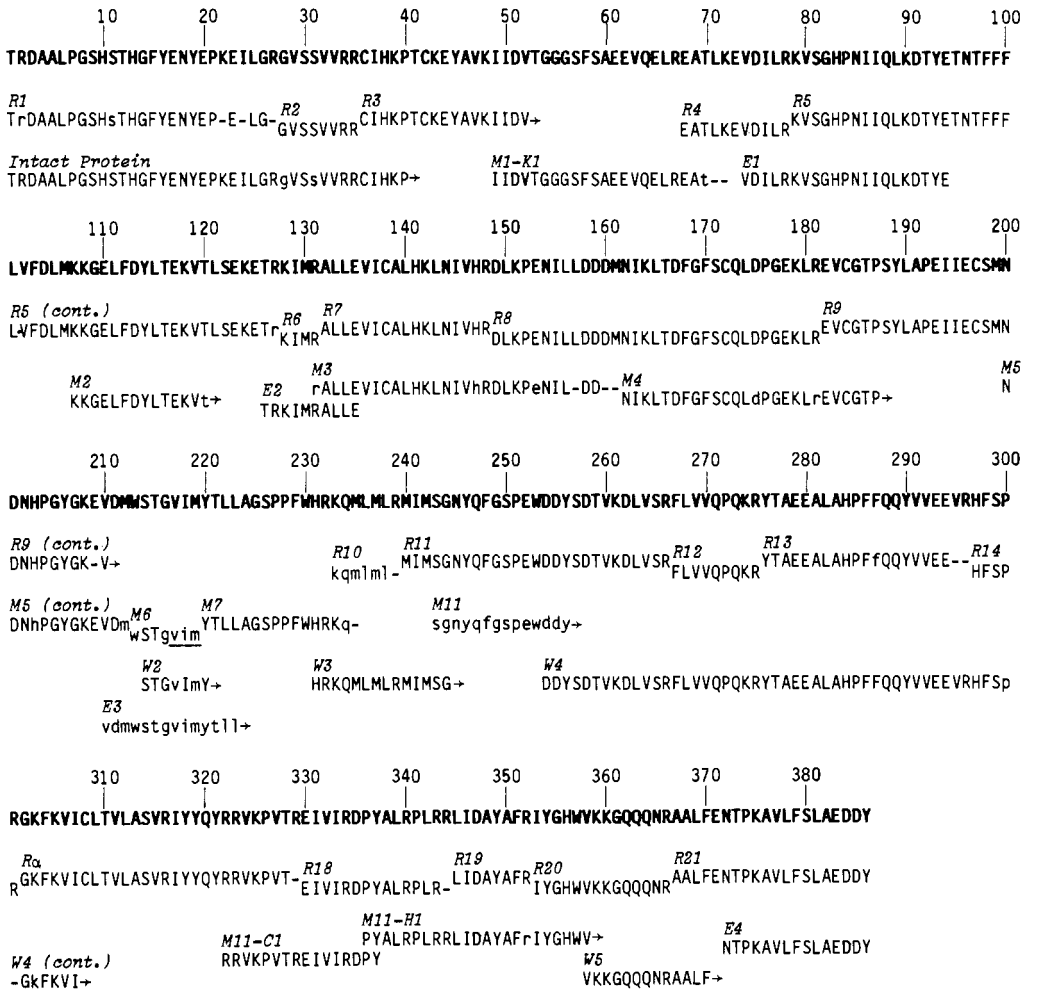

FIGURE 1: Summary of the proof of structure of the γ subunit. The proven sequences of specific peptides are given in one-letter code below the summary sequence (bold type). Prefixes E, M, R, and W denote peptides derived by cleavage of the γ subunit at glutamyl, methionyl, arginyl, and tryptophanyl residues, respectively. Peptide Rα arises from the incomplete cleavage at arginyl residues 316, 322, and 323. Peptide M1-K1 was obtained by subdigesting peptide M1 with endoproteinase Lys-C. Peptide M11-H1 was derived by aspartyl–proline cleavage of peptide M11. Peptide M11-C1 was derived by chymotryptic digestion of N-succinylated peptide M11. A lower case letter indicates that the identification of the residue is tentative. An underlined lower case letter denotes a residue deduced only from compositional analysis. An arrow indicates that the C-terminal residues were not analyzed by Edman degradation. A hyphen indicates an unidentified residue.

Results

The overall strategy used to prove the structure of the γ subunit was the determination of the amino acid sequence of a complete set of peptides generated by cleavage at arginyl bonds and then the alignment of these peptides with overlaps from other digests (Figure 1). Overlap peptides were obtained by cleavage of the whole subunit at tryptophanyl, methionyl, or glutamyl bonds and by selected subdigestions of the two large fragments isolated after cleavage at methionyl bonds.

Cleavage at Arginyl Bonds. Tryptic digestion of the N-citraconylated γ subunit yielded 19 unique and nonoverlapping peptides, representing all 386 residues of the protein (Table I). In addition, several other peptides were obtained as a result of cleavage at internal arginyl residues of some of these 19 peptides. The mixture of peptides was first separated into five pools (A–E) by gel filtration on Sephadex G-50 (Figure 2). Four of these pools were further fractionated by HPLC. Fraction E contained primarily Rα and smaller amounts of other peptides, which were separated by reversed-phase HPLC (not shown). Fraction D contained only peptide R5, which was analyzed by Edman degradation without further purification. Except as noted below, other peptides were pure after secondary separation by reversed-phase HPLC (Figure 2).

Peptide R20 was contaminated with 0.7 mol fraction of R18a (EIVIR) as estimated by composition and by observation of the minor sequence EIVI-. Likewise, R10 was contaminated with 0.37 mol fraction of R18b (DPYALPPLRR) as revealed by composition and by observation of its minor sequence. As shown by composition and sequenator analysis, R1 contained 0.67 mol fraction of a peptide corresponding to its shorter counterpart, residues 3–27. Despite the presence of these contaminants, unambiguous sequences were deduced for each of the peptides except R10, which is shown in lower case letters in Figure 1.

Yields of the 19 primary peptides obtained from 130 nmol of the γ subunit ranged from 18 to 100 nmol (Table I). In some cases the low yields were the result of poor recovery from the chromatographic systems. This is especially true for reversed-phase separation of the peptides in pool C (R3, R8, and R9). Additional R3 was isolated for sequence analysis by gel filtration of a tryptic subdigest of N-citraconylated peptide M1. Other yields were low because secondary peptides resulted from incomplete cleavage at internal arginyl residues. In peptides R1 and R18 cleavage at internal arginyl residues was incomplete due to adjacent prolyl or aspartyl residues (Allen, 1981).

Anomalous tryptic cleavages were observed in two portions of the molecule. In Rα neither the arginyl–isoleucine bond at residue 316 nor the arginyl–arginine bond at residue 322 was completely cleaved. In fact, neither of the putative pep-

Table I: Amino Acid Compositions[a] of Tryptic Peptides from S-Carboxymethylated and N-Citraconylated γ Subunit[b]

Peptide Residues No.	R1[c] 1-27	R2 28-35	R3[d] 36-67	R4 68-78	R5[e] 79-127	R6 128-131	R7 132-148	R8 149-181	R9 182-232	R10[f] 233-239	R11 240-266	R12 267-275	R13 276-296	R14 297-301	Rα 302-329	R18[g] 330-344	R19 345-352	R20[h] 353-366	R21 367-386	Total (from analysis of γ subunit)[i]	(from sequence in Fig. 1)
Source (Fig. 2 G50 Pool)	B	A	C	B	D	A	B	C	C	A	B	A	B	A	E	A	A	A	B		
Asp/Asn (D/N)	2.3(2)		1.9(1)	1.2(1)	4.6(5)		1.3(1)	7.9(8)	4.7(4)		4.8(5)				0.8(0)	1.2(1)	0.9(1)	1.3(1)	3.1(3)	32.4	(33)
Thr (T)	1.6(2)		1.8(2)	1.0(1)	4.9(6)			1.2(1)	3.0(3)		1.0(1)		1.0(1)		1.3(2)				1.0(1)	19.5	(20)
Ser (S)	2.0(2)	2.0(2)	2.2(2)		2.7(2)			1.4(1)	4.0(4)		3.3(4)			0.8(1)	1.3(1)				1.1(1)	21.1	(20)
Glu/Gln (E/Q)	3.6(3)	0.5(0)	4.4(5)	2.3(2)	6.4(6)		1.4(1)	3.6(3)	4.7(4)	1.3(1)	2.3(2)	2.0(2)	6.1(6)		1.7(1)	1.1(1)		3.0(3)	2.2(2)	44.3	(42)
Pro (P)	1.7(2)		N.D.[j](1)		1.4(1)			2.0(2)	4.8(5)		1.0(1)	0.9(1)	1.0(1)	0.7(1)	N.D.(1)	1.5(2)			0.9(1)	19.9	(19)
Gly (G)	2.8(3)	1.4(1)	3.0(3)		3.4(2)			2.5(2)	5.3(5)		1.9(2)				1.4(1)			2.1(2)		27.5	(21)
Ala (A)	2.0(2)		2.0(2)	1.1(1)	0.5(0)		2.1(2)		2.5(2)				2.9(3)		1.2(1)	1.0(1)	1.7(2)		3.6(4)	24.5	(20)
CMCys (C)			1.5(2)				0.7(1)	1.4(1)	2.5(2)						1.1(1)					4.9	(7)
Val (V)	0.5(0)	3.1(3)	2.5(3)	0.9(1)	3.1(3)		1.2(2)	0.9(0)	2.2.(3)		1.7(2)	1.3(2)	2.0(3)		3.1(5)	0.7(1)		N.D.(1)	0.9(1)	24.9	(30)
Met (M)					1.1(1)	0.9(1)		1.0(1)	2.4(3)	1.7(2)	1.7(2)									7.9	(10)
Ile (I)	1.0(1)		1.9(3)	0.9(1)	1.6(2)	1.0(1)	1.2(2)	1.5(2)	1.8(3)		0.9(1)				1.4(2)	1.4(2)	0.9(1)	N.D.(1)		19.7	(22)
Leu (L)	2.0(2)		1.4(1)	2.0(2)	6.0(6)		4.1(4)	5.1(6)	3.3(3)	1.9(2)	1.2(1)	1.0(1)	1.1(1)		2.0(2)	2.0(2)	1.0(1)		2.9(3)	37.0	(37)
Tyr (Y)	2.1(2)		1.4(1)		1.7(2)			0.7(0)	3.2(3)		1.9(2)		2.1(2)		1.3(3)	1.1(1)	0.8(1)	1.0(1)	1.0(1)	14.1	(19)
Phe (F)	N.D.(1)		0.7(1)		4.5(5)			1.5(2)	N.D.(1)		1.0(1)	0.8(1)	1.9(2)	0.8(1)	1.7(1)		0.9(1)		1.9(2)	17.7	(19)
His (H)	1.7(2)		1.2(1)		1.3(1)		1.8(2)		2.0(2)				1.0(1)	0.8(1)				1.1(1)		11.3	(11)
Lys (K)	1.1(1)		2.6(3)	1.0(1)	5.2(6)	1.0(1)	1.0(1)	3.0(3)	1.6(1)	1.0(1)	1.0(1)	0.9(1)			2.4(3)			2.0(2)	1.0(1)	24.7	(26)
Arg (R)	1.6(2)	2.0(2)	1.0(1)	1.0(1)	1.3(1)	1.2(1)	1.0(1)	1.1(1)	1.2(1)	1.0(1)	1.0(1)	1.0(1)	1.0(1)	1.0(1)	2.9(4)	3.3(4)	1.0(1)	1.0(1)	0.2(0)	24.8	(26)
Try (W)									N.D.(2)		N.D.(1)							N.D.(1)		N.D.	(4)
Yield(nmol)	19[k]	19[l]	18	72	100	48	40	28	24[m]	55	53	71	52	55	17[n]	24[p]	71	71	49		

[a]Residues per peptide by amino acid analysis (values < 0.5 not reported) or in parentheses from the sequence (Fig. 1).
[b]Peptides were isolated after digesting 130 nmol γ-subunit as described in Fig. 2.
[c]Amino acid analysis data is for a 3:2 mixture of R1:R1b (R1b = DAALPGSHSTHGFYENYEPKEILGR).
[d]For sequence analysis this peptide was isolated from a subdigest of M1.
[e]For compositional analysis this peptide was purified by reversed phase HPLC.
[f]Values are corrected for contamination with R18b (DPYALRPLRR). The ratio of R10:R18b was 2.7:1.
[g]This peptide appears to be contaminated with the related peptide having a single R at the C-terminus.
[h]Values are corrected for contamination with R18a (EIVIR). The ratio of R20:R18a was 1.4:1.
[i]Hydrolysis of S-carboxymethylated γ subunit was for 24, 72, and 120 h. Thr and Ser values were extrapolated to t = 0. Val and Ile values for 72 h are reported.
[j]N.D. = not determined.
[k]In addition, 48 nmol of R1a (TR) and 42 nmol of R1b[c] were recovered (compositional analysis only).
[l]In addition, 52 nmol of R2a (GVSSVVR) were recovered.
[m]In addition, 26 nmol (compositional analysis only) of R9a (EVC...VDM) and 19 nmol of R9b (YTL...WHR) were recovered.
[n]In addition 9 nmol of R17 (RVKPVTR) and 37 nmol of R17a (VKPVTR) were recovered.
[p]In addition 50 nmol R18a[h] and 20nmol R18b[f] were recovered.

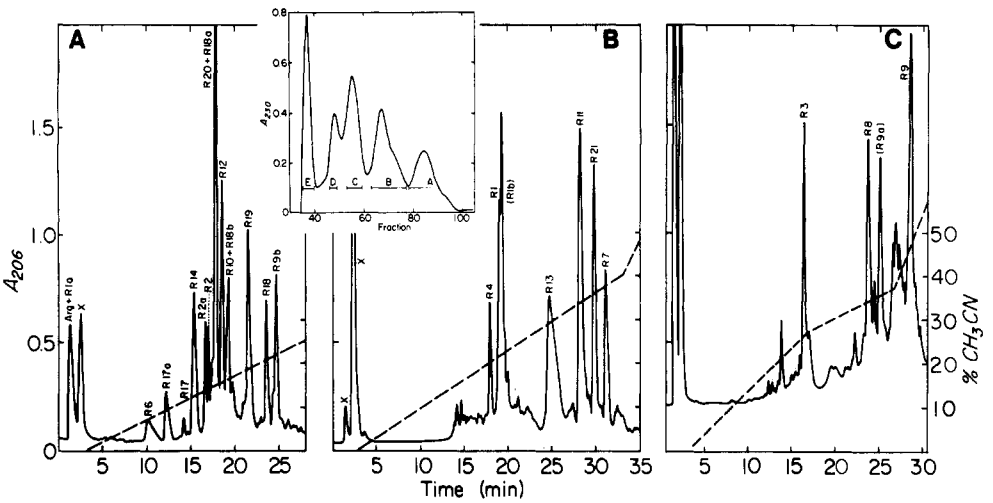

FIGURE 2: Isolation of peptides after cleavage at arginyl bonds. The peptide mixture obtained by digestion with trypsin was fractionated on a Sephadex G-50 superfine column (1.5 × 195 cm) equilibrated in 0.1 M NH_4HCO_3, pH 8.9 (insert). Fractions (3.2 mL) were collected at a flow rate of 10 mL/h. Pools A, B, and C were fractionated by reversed-phase HPLC (panels A, B, and C, respectively). Pool E was also fractionated by reversed-phase HPLC (not shown). Pool D contained a single peptide (R5). The nomenclature is indicated in Figure 1 for the major peptides and in Table I for the minor peptides.

tides R15 and R16 (residues 302–316, 317–322) was isolated. However, peptides corresponding to R17 (residues 323–329) were observed in yields totaling 35% (Table I). Perhaps the difficulty in generating these peptides relates to the hydrophobicity of this region. In this connection, although Rα has only 28 residues, it eluted in the void volume of the Sephadex G-50 column, indicating that it aggregated in 0.1 M NH_4HCO_3. In addition, anomalous cleavage of a tyrosyl–threonine bond (residues 220–221) in R9 produced 19 nmol of R9b (residues 221–232) and reduced the yield of R9. The composition of R9a indicated that it could comprise residues 182–213 (rather than residues 182–220), suggesting that the tryptophanyl–serine bond (residues 213–214) was also cleaved to generate the putative peptide STGVIMY (residues 214–220). This peptide was not recovered, but an unidentified peptide eluting at 14 min in Figure 2C has the same retention time as M6 (residues 213–219), which has one more tryptophan, one less tyrosine, and homoserine in place of methionine. These data suggest that trypsin cleaved at a tryptophanyl–serine bond in addition to the tyrosyl–threonine bond within R9. However, without tryptophan analyses of these peptides we cannot exclude the alternative possibility that the methionyl–tryptophan bond is cleaved rather than the tryptophanyl–serine bond. In either event, the region 212–220 is particularly sensitive to proteolysis, perhaps as a result of secondary structure which exists even after denaturation.

Cyanogen Bromide Digestion. Cyanogen bromide digestion also provided a virtually complete set of nonoverlapping peptides. Primary separation of the peptides on a TSK G2000 column in 7 M guanidine hydrochloride is shown in Figure 3. (Gel filtration on Sephadex G-50 in 9% HCOOH or in 7 M urea and 10 mM H_3PO_4 was somewhat less effective for separation of these peptides.) Eight of the 11 predicted peptides were recovered. The smallest peptides M8 (LM), M9 (LRM), and M10 (IM) were neither needed nor isolated from the small peak that follows M6. Further purification of peptides M2–M7 was achieved by reversed-phase HPLC. Peptides M3–M7 provided the information to overlap peptides R7–R10 and for part of the sequence of R9.

Peptides M1 and M11 were relatively insoluble and thus difficult to separate from each other. Preparations that were enriched in either M1 or M11 were used to determine their N-terminal sequences. M1 gave the same N-terminal sequence

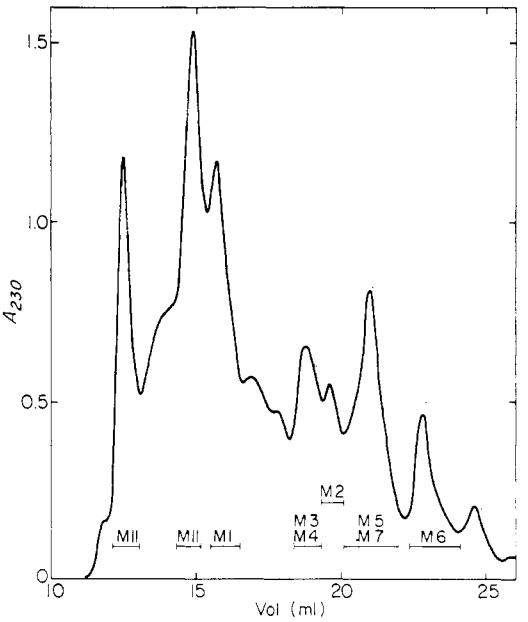

FIGURE 3: Primary separation of peptides after cleavage at methionyl bonds. The γ subunit (120 nmol) was digested in 1.0 mL of 70% formic acid containing 60 mg of CNBr. After lyophilization, the peptides were dissolved in 0.2 mL of 6 M guanidine hydrochloride and 0.1 M NaH_2PO_4, pH 6.0, and 0.1 mL was applied to a TSK G2000 column (7.5 × 600 mm) equilibrated in the same buffer. The flow rate was 0.2 mL/min.

as the intact γ subunit and M11 gave the N-terminal sequence beginning at residue 243 in Figure 1. Digestion of a mixture of M1 and M11 with endoproteinase Lys-C was used to generate peptide M1-K1, which overlapped R3/R4. Two other peptides were generated from enriched preparations of the parent fragment M11 in the following manner. Subdigestion of peptide M11 with chymotrypsin gave a peptide (M11-C1) corresponding to residues 322–337, giving the overlap for peptides Rα and R18. Treatment of M11 with mild acid cleaved the aspartyl–proline bond between residues 335 and 336. The resulting C-terminal fragment (M11-H1) provided information for overlapping peptides R18–R20.

Other Peptides. Cleavage of the γ subunit at tryptophanyl bonds yielded four useful peptides: W3 overlapped peptides M7, R10, and R11, W4 overlapped R11–Rα, W5 overlapped R20 and R21, and in addition, W2 provided confirmation for

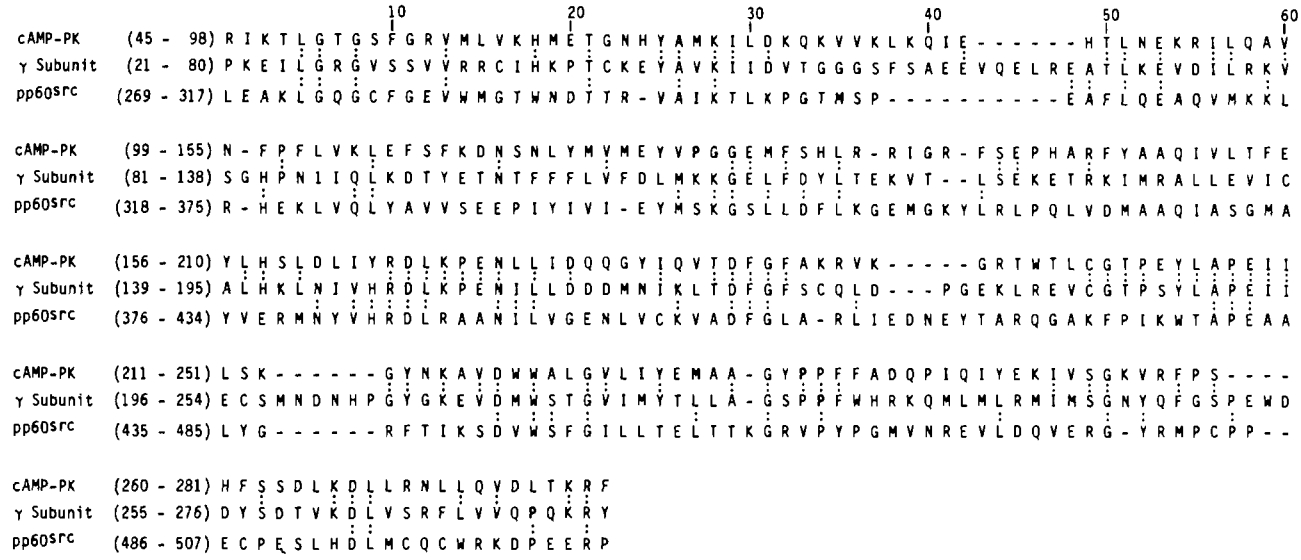

FIGURE 4: Homologous relationships among segments of bovine cAMP-dependent protein kinase, the γ subunit of phosphorylase kinase, and pp60[src], the transforming phosphoprotein of the Rous sarcoma virus (Prague C strain). Residue numbers are indicated in parentheses. Colons indicate paired identities. Hyphens identify gaps introduced to optimize the identities.

the sequence of residues 214–219. Thus the cleavages at tryptophanyl and methionyl bonds yielded peptides needed to overlap all of the arginyl peptides except for R1–R3 and R4–R7.

Completion of the proof of the structure was provided in the following manner. The N-terminal sequence of the whole protein linked R1, R2, and R3. Digestion of the whole protein at glutamyl residues (with the *S. aureus* enzyme) provided three key peptides, E1, E2, and E3. These peptides proved relatively difficult to purify, in part due to their acidic nature and in part due to the complexity of the mixture. Nonetheless, peptides E1 and E2 each gave a single sequence on analysis, overlapping R4/R5 and R5–R7, respectively. Peptide E3 was only about 65% pure, but the contaminants were readily recognized during sequenator analysis as two peptides beginning with the sequences corresponding to residues 197 (CSMN...) and 331 (IVIR...). The sequence of E3, deduced by subtraction, overlapped peptides M5–M7, to complete the sequence of R9 and thereby provide a unique solution for the structure of the γ subunit.

Evidence for the C-Terminus. Only one tryptic peptide lacking arginine (R21) was obtained in good yield (49 nmol from 130 nmol of the γ subunit). This peptide was sequenced to its C-terminus as indicated by agreement between the sequence and composition (Table I). Although another peptide lacking arginine (R9a) was isolated from the same digest in low yield (19 nmol), it was shown above to be an internal peptide. Digestion with *S. aureus* protease yielded a peptide, E4, which had both the sequence and composition expected for residues 372–386. W5 had the composition expected for residues 358–386 and M11-H1 had the composition expected for residues 336–386. Thus, there are four different peptides that indicate by their compositions and/or sequences that the C-terminus is Tyr-386.

Amino Acid Composition of the γ Subunit. The amino acid composition, as determined by sequence analysis and by acid hydrolysis, is shown in the final two columns of Table I. In general, there is agreement between the two sets of values. However, Ser, Glx, Gly, and Ala values for acid hydrolysis are somewhat high, whereas CMCys, Val, Met, Ile, Tyr, and Phe values are low. Hydrolysis for 18 h of six different preparations not purified by HPLC gave similar values except for Asx (36.0), Ser (21.1), Glx (46.7), Gly (24.4), CMCys (7.6), Met (9.0), and Ile (16.9). The reason for the lack of agreement between sequence data and hydrolysis data is not clear. Possibly there was contamination by low molecular weight peptides or free amino acids.

Table II: Alignment Scores for Segments[a] of Three Homologous Protein Kinases

	cAMP-PK (1–350)[b]	pp60[src] (251–526)[c]
γ subunit (1–276)	12.6	7.2
γ subunit (277–386)	0.3	0.3
cAMP-PK (1–350)		8.7

[a] Numbers in parentheses are residue numbers. [b] Catalytic subunit of bovine cAMP-dependent protein kinase. [c] Rous sarcoma virus (Prague C strain).

From the sequence an M_r of 44 673 was calculated. This may be compared with reported values of 41 000–45 000 based on polyacrylamide gel electrophoresis (Hayakawa et al., 1973a; Cohen, 1973).

Search for Homologous Proteins. A search of the National Biomedical Research Foundation data bank indicated that the γ subunit has a high degree of similarity to the catalytic subunit of bovine cAMP-dependent protein kinase (Shoji et al., 1983) and to several retroviral gene products (e.g., pp60[src]), which are known to be homologous with that protein (Groffen et al., 1983). Table II lists alignment scores for two large portions of the γ subunit that provide strong evidence of homology of the N-terminal 276 residues with two of these proteins, but no indication of homology of the C-terminal 110 residues with either. Specific alignments of residues 21–276 with homologous regions of the other two proteins are illustrated in Figure 4.

Discussion

The proof of the structure of he γ subunit of rabbit muscle phosphorylase kinase (Figure 1) involved two primary sets of overlapping peptides derived by cleavage at either arginyl or methionyl residues. The remaining overlaps were provided by digests specifically selected on the basis of partial sequences. In toto this led to considerable redundancy in the proof, with about 85% of the residues being observed in at least two peptides. Since each phenylthiohydantoin was identified in two complementary HPLC systems, four or more separate

Protein	Residue Numbers	Amino Acid Sequence	Reference
γ Subunit	43 - 54	K E Y A V K I I D V	This work
cAMP-PK	67 - 76	N H F A M K* I L D K	Zoller et al. (1981)
cGMP-PK	384 - 393	K T F A M K* I L K K	Hashimoto et al (1982)
pp60^{src}	290 - 299	T R V A I K T L K P	Barker & Dayhoff (1982)
pp90$^{gag-yes}$	574 - 583	T K V A I K T L K L	Kitamura et al. (1982)

FIGURE 5: Decapeptide sequences around the lysyl residues (K*) identified in the ATP-binding sites of porcine cAMP-dependent and bovine cGMP-dependent protein kinases (PK) and their implied counterparts in the γ subunit and in two tyrosine-specific viral kinases, pp60^{src} and pp90$^{gag-yes}$. pp90$^{gag-yes}$ is the product of the transforming gene of avian sarcoma virus Y73.

identifications have established the identity of the majority of the residues in the sequence.

The weakest portions of the proof are found in two regions, residues 211–220 and residues 301–322. In the latter region, the overlap of R14 and Rα depends upon analysis of cycles 48–54 of a single Edman degradation of fragment W4. In contrast, portions of the region of residues 211–220 have been observed in five different peptides, but none of them provided ideal data during Edman degradation. The N-terminal tryptophan in peptide M6 was apparently blocked (cf. Uphaus et al., 1959); furthermore, peptide W2 washed out of the spinning cup rapidly during degradation and was largely blocked also. The overlapping peptide E3 was sequenced as a mixture from which the major component (65%) was deduced after subtraction of two other recognizable sequences.

The amino acid composition of the γ subunit is not in total agreement with that of the proven sequence. Although the amino acid composition of this protein has been reported previously (Hayakawa et al., 1973b; Cohen, 1973; Chan & Graves, 1982), there is considerable variation among those reports for many of the amino acids. This variation may be the result of minor contamination with free amino acids or with proteins or peptides of very different composition. In any event the sequence shown in Figure 1 appears to be complete. This follows from the fact that both ends of the molecule are clearly established and the overlaps tie the intervening peptides into one chain. The N-terminus was determined by sequenator analysis of the whole protein and the C-terminus by the finding of a peptide in each of four digests that could, by specificity arguments, only be derived from the C-terminus. For example, both R21 and E4 ended in Glu-Asp-Asp-Tyr, thus placing them at the C-terminus by virtue of the enzymatic specificity of trypsin and of the *S. aureus* protease, respectively. The internal overlaps among the sets of peptides appear to leave no opportunity for alternative arrangements or missing peptides. Furthermore, every peptide isolated, many of which are not included in the minimum proof of Figure 1, fits the derived structure.

Any future attempt to label the active site of this protein could well take advantage of our observation that tryptic digestion of the citraconylated γ subunit gives a complete set of peptides of intermediate size (ranging from 4 to 51 residues), which are readily separable (Figure 2).

Homology with Other Proteins. The similarity of the amino acid sequence of cAMP-dependent protein kinase (catalytic subunit) with that of a major segment of phosphorylase *b* kinase (alignment score of 12.6 in Table II) provides strong evidence that the two enzymes are homologous. In the alignment of Figure 4, identical loci for 36% of the amino acid residues occur in the two proteins.

It is to be expected that this homology of the γ subunit with cAMP-dependent protein kinase would dictate homology of the γ subunit with pp60^{src} since the latter and cAMP-dependent protein kinase are homologous (Barker & Dayhoff, 1982). Comparison of the γ subunit with pp60^{src} shows a lower alignment score, 7.2 (Table II), and less identity (23%). Similar values are seen when cAMP-dependent protein kinase is optimally aligned with pp60^{src}. This suggests that the γ subunit and the catalytic subunit of cAMP-dependent protein kinase are more similar to each other than either is to pp60^{src}. One can only speculate whether the larger difference relates to contrasting enzymatic functions (e.g., of specificity toward serine vs. tyrosine) or to characteristics related to the capture of a eukaryotic gene by a retrovirus. Perhaps the gene found in the virus diverged from an ancestor of the serine-specific kinase genes prior to divergence of the latter from each other. In any case, it now appears that the two serine-specific protein kinases, one activated by cyclic nucleotides and the other by calcium ions, have a common ancestral relationship and presumably a common chemical basis for their catalytic function. Our laboratory has observed that cGMP-dependent protein kinase (Takio et al., 1984) is also homologous with cAMP-dependent protein kinase. In addition, several tyrosine-specific protein kinases coded by retroviral oncogenes appear to be members of the same family (Groffen et al., 1983). Thus, several variants of a single primordial gene appear to find expression in diverse regulatory roles in contemporary cellular functions.

There is currently limited chemical evidence to identify the residues that comprise the active sites of protein kinases (Taylor et al., 1981). The affinity label [(fluorosulfonyl)-benzoyl]adenosine has identified lysyl residues that appear to be in the vicinity of the normal binding sites of the γ-phosphate of ATP in two protein kinases, porcine cAMP-dependent protein kinase (Zoller et al., 1981) and bovine cGMP-dependent protein kinase (Hashimoto et al., 1982). From these studies and from the homology shown by the alignment in Figure 4, one component of the active site of phosphorylase *b* kinase can be predicted. The homologous regions around the labeled lysine are illustrated in Figure 5 for the cyclic nucleotide dependent kinases, for the γ subunit, and for two retroviral gene products. It should be noted that the lysine labeled in both cyclic nucleotide dependent kinases is found in a homologous locus in each of the other proteins and that the neighboring peptide structures retain their general chemical character, although not the identical sequences. On this basis, it appears that lysyl residue 48 in phosphorylase *b* kinase may interact with the terminal phosphate of the ATP substrate.

A more quantitative examination of the relationship between the γ subunit and the catalytic subunit of cAMP-dependent protein kinase reveals that homology is only evident in the midportion of the molecule (Figure 6). Statistical analysis of alignment scores for the N-terminal 20 residues of the γ subunit revealed no significant relationship to cAMP-dependent protein kinase above alignment scores generated by first randomizing the sequences. Strikingly, the homologous relationship in the middle 250 residues of the molecule ends in the vicinity of residue 276. No combination of gaps gen-

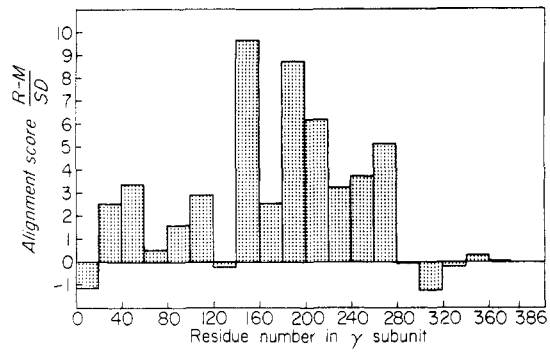

FIGURE 6: Alignment scores for 20-residue segments of the γ subunit with their corresponding regions (Figure 4) of bovine cAMP-dependent protein kinase. The scores were determined with the computer program (ALIGN) of Dayhoff (1979), using appropriate 40–60-residue segments of the bovine kinase and a gap penalty of 30 to prevent misalignment.

erated any significant relationship between the C-terminal 110 residues of the γ subunit and any part of cAMP-dependent protein kinase or pp60^{src}. It is of course not clear whether one is observing the product or pervasive divergence in the evolution of a portion of the chains of whether the apparent discontinuity is the product of a splice in an ancestral recombinant event.

The idea of a splice opens the door to speculation that the C-terminal 110 residues could be identified with a key difference between phosphorylase *b* kinase and cAMP-dependent protein kinase. However, there are several differences, namely, in substrate specificity, in the nature of the regulators, and in the interactions with other subunits. Because there is some overlap in the specificity of the two kinases (e.g., toward glycogen synthase and troponin I), it is unlikely that this region of the molecule is a key determinant in that specificity. The γ subunit binds both the α and the δ subunits of phosphorylase kinase (Chan & Graves, 1982), and it is possible that these oligomeric interactions involve this domain of the γ subunit. The role of the α subunit is unclear, but the δ subunit (calmodulin) mediates the calcium-induced activation of several enzymes. Thus, it is possible that the regulatory role of calcium ions involves interaction of calmodulin with the C-terminal portion of the γ subunit. To test this possibility, we are currently seeking homologous domains among the sequences of other calcium-activated proteins.

Acknowledgments

We are grateful to Dr. Edwin G. Krebs for providing some of the phosphorylase kinase, to Dr. Koji Takio for many valuable suggestions, and to Andy Karplus for help with the computer programs. We thank Christy Smith, Santosh Kumar, and Sandy Ferguson for their technical assistance and Mary Woods for the typescript.

References

Adelstein, R. S., & Eisenberg, E. (1980) *Annu. Rev. Biochem. 49*, 921.

Allen, G. (1981) in *Laboratory Techniques in Biochemistry and Molecular Biology 9* (Work, T. S., & Burdon, P. H., Eds.) p 53, Elsevier/North-Holland Biomedical Press, New York.

Atassi, M. Z., & Habeeb, A. F. S. A. (1972) *Methods Enzymol. 25*, 546.

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

Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry 14*, 3029.

Bridgen, P. J., Cross, G. A. M., & Bridgen, J. (1976) *Nature (London) 263*, 613.

Carlson, G. M., Bechtel, P. J., & Graves, D. J. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol. 50*, 41.

Carpenter, F. H. (1967) *Methods Enzymol. 11*, 237.

Chan, K.-F. J., & Graves, D. J. (1982) *J. Biol. Chem. 257*, 5939.

Cohen, P. (1973) *Eur. J. Biochem. 34*, 1.

Cohen, P., Burchell, A. Foulkes, J. G., Cohen, P. T. W., Vanaman, T. C., & Nairn, A. C. (1978) *FEBS Lett. 92*, 287.

Dayhoff, M. O. (1979) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, pp 1–8, National Biomedical Research Foundation, Washington, DC.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80.

Embi, N., Rylatt, D. B., & Cohen, P. (1979) *Eur. J. Biochem. 100*, 339.

Embi, N., Parker, P. J., & Cohen, P. (1981) *Eur. J. Biochem. 115*, 405.

Ericsson, L. H., Wade, R. D., Gagnon, J., MacDonald, R. M., Granberg, R. R., & Walsh, K. A. (1977) in *Solid Phase Method in Protein Sequence Analysis* (Previero, A., & Coletti-Previero, M. A., Eds.) p 137, Elsevier/North-Holland Press, New York.

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

Fontana, A., Vito, C., & Toniolo, C. (1973) *FEBS Lett. 32*, 139.

Grand, R. J. A., Shenolikar, S., & Cohen, P. (1981) *Eur. J. Biochem. 113*, 359.

Groffen, J., Heisterkamp, N., Reynolds, F. H., & Stephenson, J. R. (1983) *Nature (London) 304*, 167.

Gross, E. (1967) *Methods Enzymol. 11*, 238.

Hashimoto, E., Takio, K., & Krebs, E. G. (1982) *J. Biol. Chem. 257*, 727.

Hayakawa, T., Perkins, J. P., Walsh, D. A., & Krebs, E. G. (1973a) *Biochemistry 12*, 567.

Hayakawa, T., Perkins, J. P., & Krebs, E. G. (1973b) *Biochemistry 12*, 574.

Hunter, T., & Sefton, B. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1311.

Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y., & Yoshida, M. (1982) *Nature (London) 297*, 205.

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

Laemmli, U. K. (1970) *Nature (London) 227*, 680.

Landon, M. (1977) *Methods Enzymol. 47*, 145.

Malencik, D. A., & Fischer, E. H. (1982) in *Calcium and Cell Function*, Vol. III, pp 161–188, Academic Press, New York.

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

Skuster, J. R., Chan, K.-F. J., & Graves, D. J. (1980) *J. Biol. Chem. 255*, 2203.

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., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A., & Titani, K. (1984) *Biochemistry* (accompanying 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.

Uphaus, R. A., Grossweiner, L. I., Katz, J. J., & Koppler, K. D. (1959) *Science (Washington, D.C.) 129*, 641.

Yaoi, Y., Titani, K., & Narita, K. (1964) *J. Biochem. (Tokyo) 56*, 222.

Zoller, M. J., Nelson, N. C., & Taylor, S. S. (1981) *J. Biol. Chem. 256*, 10837.