Guanosine Cyclic 3′,5′-Phosphate Dependent Protein Kinase, a Chimeric Protein Homologous with Two Separate Protein Families†

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

ABSTRACT: The amino acid sequence of bovine lung cGMP-dependent protein kinase has been determined by degradation and alignment of two primary overlapping sets of peptides generated by cleavage at methionyl or arginyl residues. The protein contains 670 residues in a single N^{α}-acetylated chain corresponding to a molecular weight of 76 331. The function of the molecule is considered in six segments of sequence which may correspond to four folding domains. From the amino terminus, the first segment is related to the dimerizing property of the protein. The second and third segments appear to have evolved from an ancestral tandem internal gene duplication, generating twin cGMP-binding domains which are homologous to twin domains in the regulatory subunits of cAMP-dependent protein kinase and to the cAMP-binding domain of the catabolite gene activator of *Escherichia coli*. The fourth and fifth segments may comprise one domain which is homologous to the catalytic subunits of cAMP-dependent protein kinase, of calcium-dependent phosphorylase *b* kinase, and of certain oncogenic viral protein tyrosine kinases. The regulatory, amino-terminal half of cGMP-dependent protein kinase appears to be related to a family of smaller proteins that bind cAMP for diverse purposes, whereas the catalytic, carboxyl-terminal half is related to a family of protein kinases of varying specificity and varying sensitivity to regulators. These data suggest that ancestral gene splicing events may have been involved in the fusion of two families of proteins to generate the allosteric character of this chimeric enzyme.

Studies of protein kinases in many laboratories have made it clear that they serve important regulatory roles by activation or inactivation of key enzymes in various cellular processes [cf. reviews by Krebs & Beavo (1979) and Flockhart & Corbin (1982)]. Among the more thoroughly studied protein kinases are those that phosphorylate serine or threonine residues in response to either cyclic nucleotides or calcium/calmodulin. More recently it has been shown that the receptor of the epidermal growth factor (Ushiro & Cohen, 1980) and certain retroviral oncogene products phosphorylate specific tyrosyl residues in acceptor proteins [reviewed by Sefton & Hunter (1984)].

The various protein kinases respond to different signals and serve different regulatory roles in cells but share the analogous function of controlling substrate proteins by transferring the terminal phosphate of ATP to specific serine, threonine, or tyrosine residues. Several groups of investigators have speculated that some of the kinases might share an underlying structural basis for their analogies in function (e.g., Lincoln & Corbin, 1977; Gill et al., 1977). The first sequence data in support of this view showed homology between the catalytic subunit of cAMP-dependent protein kinase and $pp60^{src}$, the protein tyrosine kinase encoded by the src gene of Rous sarcoma virus (Barker & Dayhoff, 1982). At about the same time the amino acid sequence of an affinity-labeled fragment of cGMP-dependent protein kinase (cGK)[1] was shown to be homologous with a corresponding segment of cAK, the cAMP-dependent protein kinase (Hashimoto et al., 1982). Since then, Takio et al. (1983) have reported similarities in the hinge regions and autophosphorylation sites of these two enzymes, and Weber et al. (1982) have reported that the cAMP-binding catabolite gene activator protein of *Escherichia coli* (CAP) is homologous with the regulatory subunit of cAK. Our laboratory has also presented a preliminary report of extensive homology between cGK and cAK (Titani et al., 1983) and a complete report of the homology of cAK with the γ subunit of phosphorylase *b* kinase, a calmodulin-controlled enzyme (Reimann et al., 1984). Accompanying reports describe the homology of the two isozymes of cAK with each other (Takio et al., 1984; Titani et al., 1984). Finally, there is a growing series of reports of homology among a wide variety of retrovirally encoded transforming proteins including $pp60^{src}$ (e.g., Groffen et al., 1983; Privalsky et al., 1984), that in turn appear to be related to the epidermal growth factor receptor (Downward et al., 1984) that mediates mitogenic effects [cf. review by Sefton & Hunter (1984)].

The present report shows that the single polypeptide chain of cGK can be considered as a series of domains that are homologous not only with domains within the subunits of cAK but also with CAP, with $pp60^{src}$, and with the γ subunit of phosphorylase kinase. Of particular significance is the finding that all of the functional domains and homologous regions are within the *single* 670-residue chain of cGK rather than distributed among separate subunits as in the cAMP-dependent protein kinases. The experimental data in this report provide proof of the sequence of cGK. The derived sequence is then compared to portions of the structures of the six other proteins which either bind cAMP or catalyze the phosphorylation of proteins.

The holoenzyme cGK is composed of two identical polypeptide chains, each with a molecular weight of about 75 000. When two molecules of cGMP are bound to each chain (Corbin & Døskeland, 1983), the enzyme becomes activated in an allosteric manner (Lincoln et al., 1977; Gill et al., 1977; Takai et al., 1976). This is in contrast to the activation by

† 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: cGK and cAK, cGMP-dependent and cAMP-dependent protein kinase, respectively; R_I (R_{II}), regulatory subunit of type I (II) cAMP-dependent protein kinase; PbKγ, the γ subunit of rabbit skeletal muscle phosphorylase *b* kinase; CAP, catabolite gene activator protein of *E. coli*; $pp60^{src}$, transforming phosphoprotein of Rous sarcoma virus; HPLC, high-performance liquid chromatography; CM, *S*-carboxymethyl; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/84/0423-4207$01.50/0

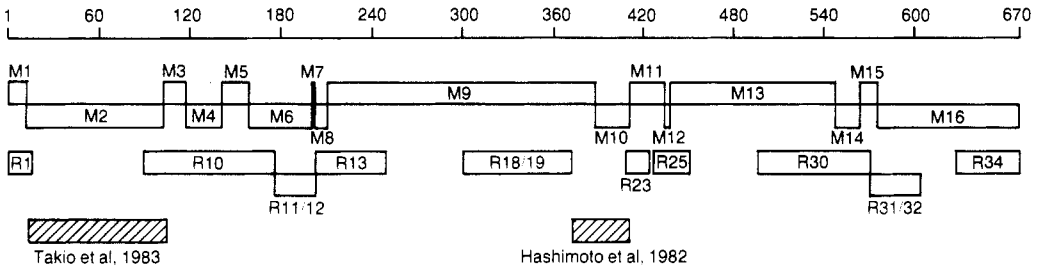

FIGURE 1: Relative orientiation of 26 of the major fragments of cGK: M1–M16, products of cleavage of methionyl residues; R1–R34, selected larger fragments from cleavage at arginyl residues; hashed bars, sequences published previously from these laboratories. Each peptide is designated by a bar of a length consistent with the 670-residue number line at the top. Details of the structural analyses of these and other peptides are illustrated in Figure 2.

cAMP of the two well-studied isozymes of cAK (type I and II), which involves in each case dissociation of an inactive tetrameric holoenzyme into two active catalytic subunits and a dimer of regulatory subunits [cf. review by Krebs & Beavo (1979)]. Despite this difference, the two protein kinases resemble each other in amino acid composition (Lincoln & Corbin, 1977), in the tendency of cGK and type II cAK to autophosphorylate (Erlichman et al., 1974; deJonge & Rosen, 1977), in their substrate specificities [reviewed in Glass & Krebs (1980)], and, as will now be demonstrated, in their amino acid sequences.

Materials and Methods

Bovine lung cGMP-dependent protein kinase was purified to homogeneity as described by Glass & Krebs (1979). Carboxypeptidase Y was a generous gift of Dr. M. Ottesen (Carlsberg Laboratories, Copenhagen). Pronase was purchased from Calbiochem. All other enzymes were obtained from the sources listed by Titani et al. (1984).

Citraconylation followed the procedure of Takio et al. (1984). Reduction and alkylation, fragmentation with CNBr, amino acid analysis, sequenator analyses, and phenylthiohydantoin identifications were performed as described by Takio et al. (1983). Mass spectrometric analysis followed the outline of Titani et al. (1984). Peptides were purified by reversed-phase HPLC with a Varian 5000 liquid chromatograph and, unless otherwise stated, a SynChropak RP-P column (SynChrom) with an acetonitrile gradient in aqueous trifluoroacetic acid (Mahoney & Hermodson, 1980). Gel filtration columns on Sephadex G-50 superfine or Sephacryl S-200 were monitored with fluorescamine by the method of Nakai et al. (1974).

[^{14}C]CM-cGK (30 mg) 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 trypsin (0.3 mg) for 1 h. The digestion mixture was then separated directly on a Sephadex G-50 superfine column at pH 8.8. Pooled fractions of eluate were lyophilized and decitraconylated in 9% formic acid at 37 °C for 2 h. Subdigestions of peptides with chymotrypsin, trypsin, *S. aureus* protease, endoproteinase Lys-C, and Pronase used 1:30–1:100 weight ratios of enzyme:substrate in 0.1 M NH_4HCO_3, pH 8, 37 °C, for 3–4 h, unless otherwise indicated. Thermolytic digestion was performed in 0.1 M Tris, pH 8, 37 °C, for 1 h.

Searches for homologous proteins were performed on a VAX/VMS computer using the SEARCH program and a mutation data matrix that scores both similarities and identities, as described by Dayhoff et al. (1983). Optimal alignment of related proteins was established by using the same data matrix, a penalty for gaps, and the ALIGN program of Dayhoff et al. (1983). Alignment scores were expressed in units of standard deviation from a mean of those from randomly generated sequences of the same composition. Previous applications of these programs in our laboratory are described in greater detail by Reimann et al. (1984). An alignment score of greater than 5 is taken as convincing evidence of a homologous relationship, whereas a score of 2 has a probability of 0.023 that a randomly generated sequence of the same composition would generate an equivalent score.

Results

General Strategy of Sequence Analysis. A complete set of 16 fragments was generated by cleavage at methionyl bonds; these fragments were largely sequenced, overlapped, and aligned by using peptides derived after cleavage at arginyl bonds. This general strategy is illustrated in Figure 1, which displays the relative orientation of many of the peptides. Figure 1 also indicates the location of the two sequences previously placed in the "hinge" region (Takio et al., 1983) and in the ATP-binding site (Hashimoto et al., 1982). More complete details of the proof of sequence are summarized in Figure 2.

Cleavage with Cyanogen Bromide. Figure 3A illustrates the primary separation of the products of cyanogen bromide cleavage. Pooled fractions 3 and 4 yielded virtually pure peptides M16 and M10, respectively, and each was finally purified by HPLC on an ODS-HC/Sil-X-1 column. Pooled fraction 2 was separated on a SynChropak RP-P column into a major peptide M2 and minor peptides M9 and M_A. The latter is an uncleaved overlap of M1 and M2, previously described by Takio et al. (1983), who referred to M2 as M_B. Complex mixtures in pooled fractions 1, 5, and 6 were resolved into 12 peptides as described in Figures 3B,C and 4. Pooled fraction 7 was similarly separated on a μBondapak column into homoserine (M7) and leucylhomoserine (M12).

Peptides M1 and M_A were completely blocked at the amino terminus and correspond to the amino terminus of the whole protein which was shown by Monken & Gill (1980) to be blocked. M11 was also largely blocked, but it was obtained in two major forms (Figure 4A) of identical composition, one of which had an amino-terminal glutamine which must have cyclized during the acidic cleavage and fractionation procedures to a pyroglutamyl residue in the other.

Amino acid analyses of peptides M1–M16 were reasonably consistent with the recovery of all 16 fragments of methionyl cleavage (Table I), and direct sequenator analyses of the unblocked peptides placed 310 residues in 15 segments of sequence (Figure 2).

Cleavage at Arginyl Residues. Thirty milligrams of [^{14}C]CM-cGK was citraconylated, digested with trypsin, and fractionated on Sephadex G-50 at pH 8.8 (Figure 5). Pooled fractions were decitraconylated in acid and resolved into 36 different peptides by HPLC. Peptides are numbered R1, R2, etc., from the amino terminus, except in six cases where un-

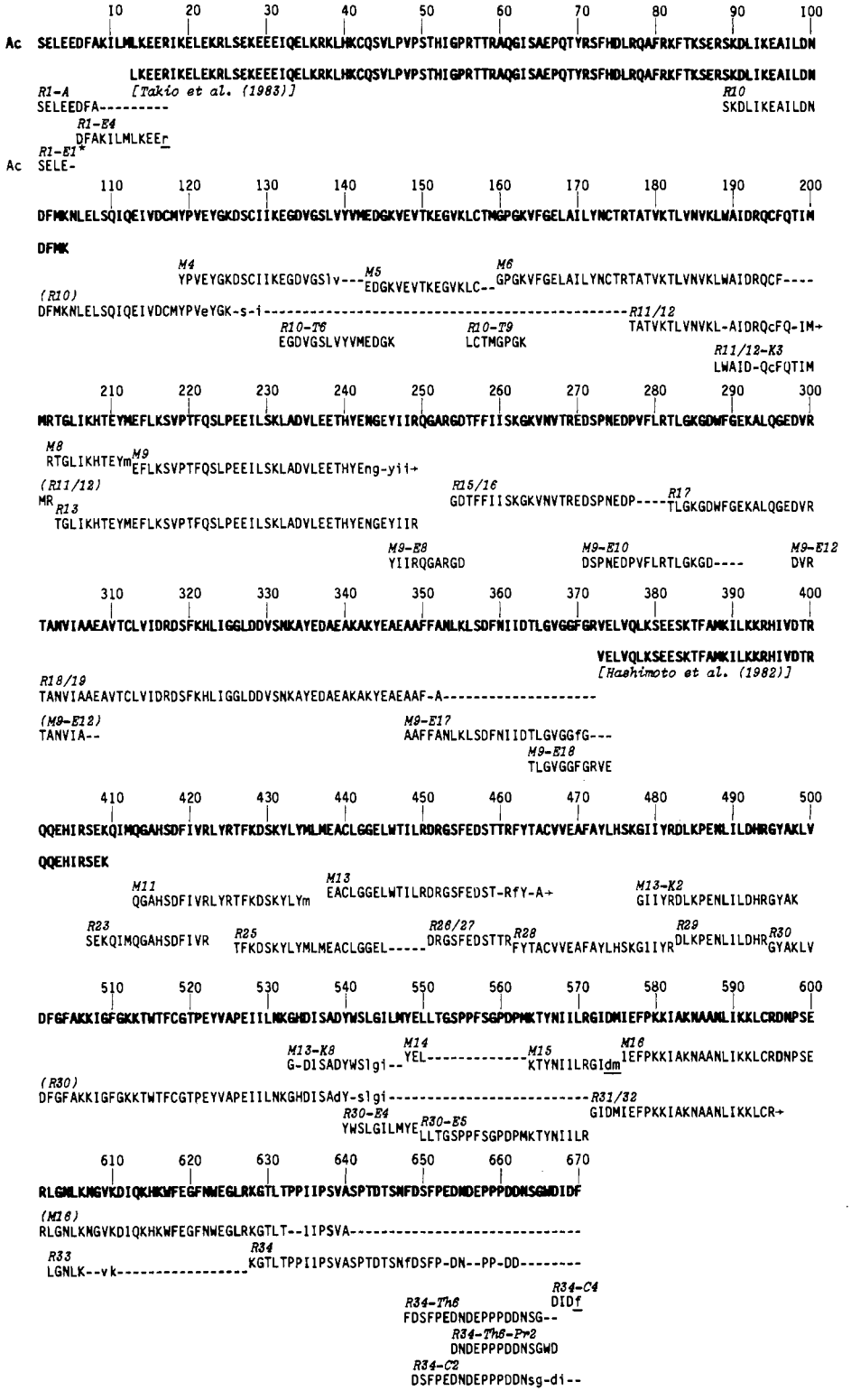

FIGURE 2: Detailed summary of the proof of sequence of cGK. The proven sequences of specific peptides (named in italics) are given in one-letter code below the summary sequence (bold type). Prefixes M and R denote peptides generated by cleavage of the CM protein at methionyl and arginyl bonds, respectively. The arginyl peptides are numbered in order from 1 to 34 except where an uncleaved arginyl residue gives an overlap, e.g., R11/12. Subpeptides are identified in hyphenated suffixes, with the following code indicating the subdigesting agent: A, acid-promoted N to O shift; C, E, K, Pr, T, or Th, enzymatic cleavage with chymotrypsin, *S. aureus* protease (at Glu), endoproteinase Lys-C, Pronase, trypsin, or thermolysin, respectively. Peptide sequences written in upper case letters were proven by Edman degradation (except R1-E1*, where mass spectrometry was used), those in lower case letters indicate tentative identification, and those underlined are deduced from amino acid compositions. Those not identified are shown by dashes or by an arrow, which indicates a long unidentified sequence. Ac denotes an acetyl group. The two sequences in boldface (residues 13–103 and 378–409) refer to those published previously by Takio et al. (1983) or Hashimoto et al. (1982), respectively.

cleaved arginyl bonds are designated by linked nomenclature such as R11/12 in Figure 2. Sequenator analyses of these peptides placed 431 residues in sequence and either provided overlaps for cyanogen bromide fragments or linked them to previously reported segments of sequence.

Overlaps of the Primary Peptides. Completion of the individual peptide sequences and provision of overlaps among the two primary sets of peptides are discussed below in order

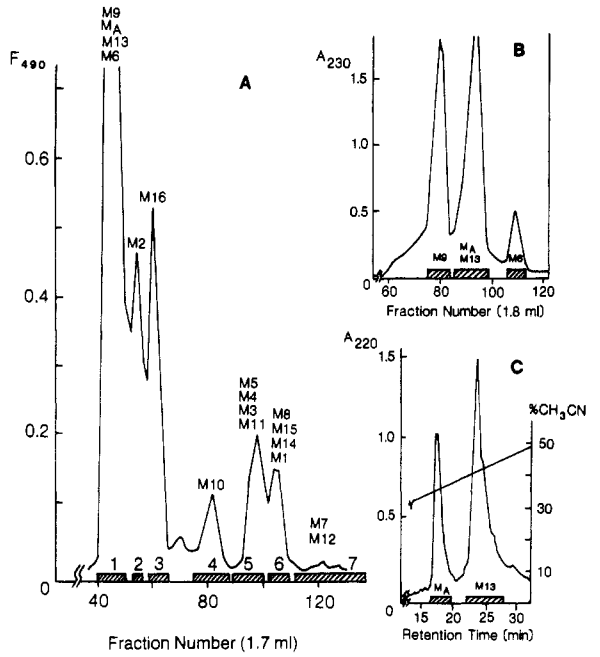

FIGURE 3: Separation of a cyanogen bromide digest of 26 mg of [^{14}C]CM-cGK. Horizontal bars indicate pooled fractions. (A) Primary fractionation on a column (1.5 × 146 cm) of Sephadex G-50 superfine in 9% formic acid at 8.5 mL/h into 1.7-mL fractions. Symbols M1–M16 identify the location of fragments in pooled fractions. M_A refers to a fragment that overlaps M1 and M2 (Takio et al., 1983). (B) Further fractionation of pooled fraction 1 (from A) on a column of Sephacryl S-200 (1.5 × 171 cm) in 7 M urea containing 10 mM HCl at 10.8 mL/h into 1.8-mL fractions. (C) Separation of pooled fraction 2 [from (B)] by HPLC on an ODS-HC/Sil-X-1 column at 2 mL/min with an acetonitrile gradient.

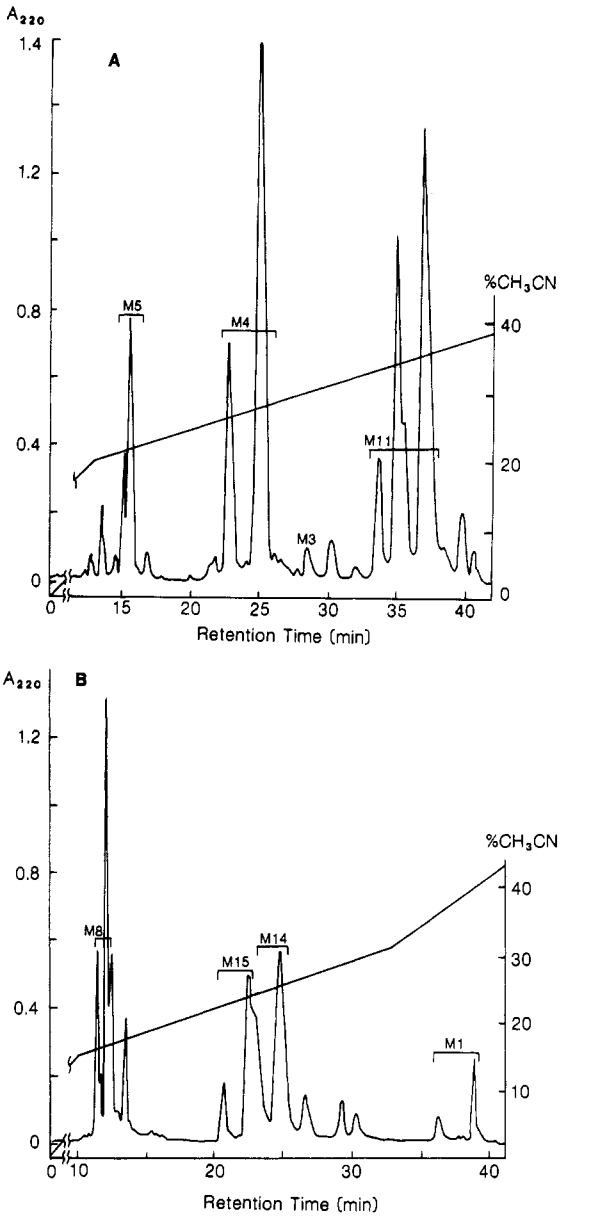

FIGURE 4: HPLC separations on a μBondapak C18 column (2 mL/min) of peptides in pooled fractions from Figure 3A. (A) Pooled fraction 5 and (B) pooled fraction 6. Purified peptides are identified by the prefix M, as in Figure 2.

from the amino terminus to the carboxyl terminus in those cases where the proof is not self-evident in Figure 2. As is usual, the most difficult problems are at the ends of the whole molecule.

The amino terminus is blocked by an acetyl group. This was shown by mass spectrometry of the permethylated subpeptide R1-E1 (Tables II and III) which yielded fragment ions of (methylated) acetyl-Ser, acetyl-Ser-Glu, acetyl-Ser-Glu-Leu, and acetyl-Ser-Glu-Leu-Glu. This sequence was confirmed and extended by first unblocking peptide R1 by an *N,O*-acyl shift (12 N HCL, 22 °C, 15 h) and then performing sequenator analysis (R1-A in Figure 2). This placed 8 residues, overlapping R1-E4, which in turn overlapped the 92-residue segment of Takio et al. (1983) to reach residue 104. Peptide R10 and its subpeptides (Table II) aligned M4–M6 to Phe-196. The poorly cleaved arginyl bond in the Asp-Arg-Gln sequence at residue 193 provided the overlap peptide R11/12 in 24% yield ending in Ile-Met-Met-Arg_{202}. Fortunately, only one cyanogen bromide fragment, M8, had an amino-terminal arginine, and it must follow Met-201. Peptide R13 clearly overlaps M8 and the largest CNBr fragment M9.

Fragment M9 contained 176 residues. Subdigestion with *S. aureus* protease provided several arginine-containing peptides (Figure 6A) that overlapped R13 through R18/19 (Figure 2) ending with the sequence FGRVE, which overlapped the peptide isolated from the labeled ATP-binding site by Hashimoto et al. (1982). The latter overlap was only two residues, but no other product of cleavage at arginine had an amino-terminal valine. Complementarity of the two sets of peptides extended the sequence to Arg-461, where M13 provided a marginal overlap of R28, leading to Arg-482.

Extension of the sequence required subdigestion of two large fragments, M13 by endoproteinase Lys-C for 22 h (Figure 6B) and R30 by *S. aureus* protease for 15 h in 50 mM phosphate, pH 7.8. The isolated subpeptides provided linkages from R28 to Leu-550 in M14. Edman degradation of M14 identified only three residues, and the stepwise degradation yield decreased rapidly as if the peptide was washing out of the cup. As a consequence, the overlap to R30-E5 consists of a single leucyl residue. This very weak overlap is, however, consistent with the composition of M14 (Table I).

Peptides M15 and M16 align R30–R34. The carboxyl terminus was established as follows: of the CNBr peptides, only M16 lacked homoserine, identifying it as the carboxyl-terminal peptide. Similarly R34, which overlapped M16, was placed by its unique lack of arginine. Sequenator analysis of R34 extended the sequence to Asp-662, but with some gaps. These gaps were filled by using a thermolytic subpeptide of R34 and its cleavage product with Pronase (Table II), which provided a sequence to Asp-667. Finally, two overlapping chymotryptic subpeptides of R34 (Table III) yielded evidence of the terminal sequence Asp-Ile-Asp-Phe. Digestion of 1–2 nmol of M16, R34, or R34-C2 with 0.6 μg of carboxypeptidase

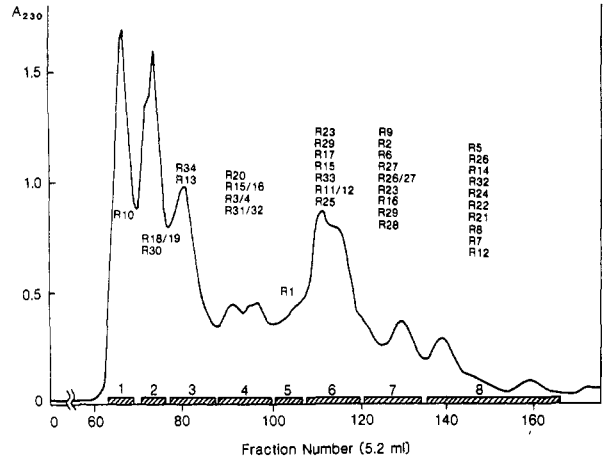

FIGURE 5: Primary separation of tryptic peptides of citraconylated [^{14}C]CM-cGK (30 mg) on a column (2.5 × 200 cm) of Sephadex G-50 superfine in 0.1 M NH_4HCO_3 (pH 8.8) at 30 mL/h into 5.2-mL fractions. Horizontal bars indicate pooled fractions. Symbols R1–R34 identify the location of fragments. The order of the symbols (top to bottom) indicates the order of elution during subsequent fractionation on a reversed-phase HPLC SynChropak RP-P column.

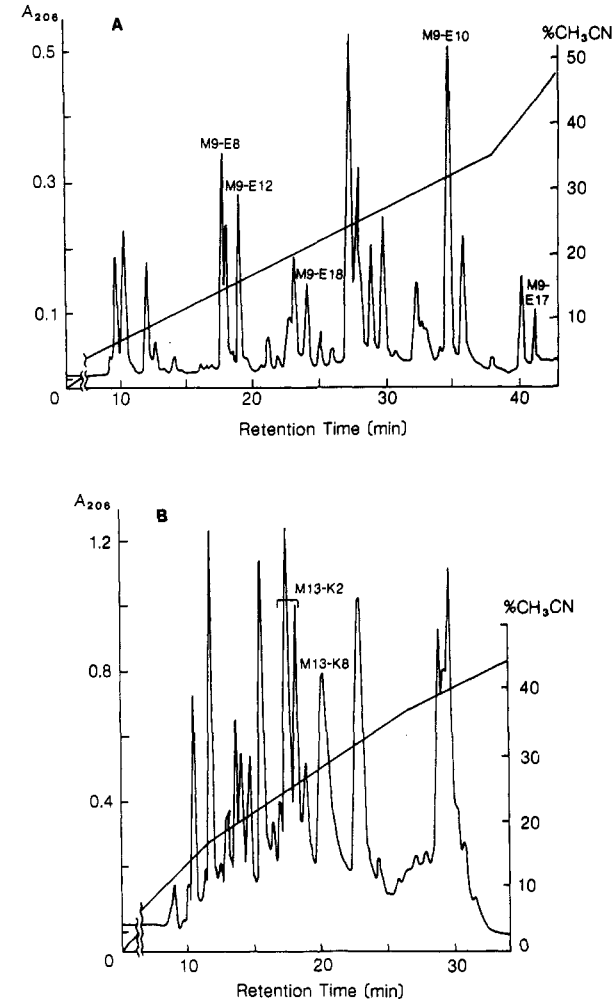

FIGURE 6: HPLC separations, as in Figure 4, of (A) a subdigest of fragment M9 (14 nmol) by the *S. aureus* protease and (B) fragment M13 (100 nmol) by endoproteinase Lys-C.

Y for 2–4 h at pH 5.4 yielded 0.8–1.0 equiv of Phe and Ile, 1.4–1.6 equiv of Asp, and less than 0.5 equiv of any other amino acid. Glycyl residues are resistant to this enzyme, so the data confirm the placement of (Asp, Ile, Asp, Phe) at the carboxyl termini of both M16 and R34.

Search for Homologous Proteins. The cGK sequence was arbitrarily divided into 45-residue segments and compared with the Jan 12, 1984, edition of the data bank of the National Biomedical Research Foundation (2511 proteins, supplemented with locally generated sequences) by using the SEARCH program of Dayhoff et al. (1983). A relationship was considered to be significant when more than two segments of cGK displayed high scores for relatedness with a corresponding set of segments of another protein. The four proteins that scored highest by this analysis were the regulatory subunits R_I and R_{II} of bovine cAMP-dependent protein kinase, the corresponding catalytic subunit, and the γ subunit of rabbit phosphorylase *b* kinase.

Internal Homology. In the course of the sequence analysis of cGK, it became obvious that there was evidence of partial duplication of sequences, as had been observed previously in R_{II} (Takio et al., 1982) and in R_I (Titani et al., 1984). Optimization of the internal alignment in cGK was performed with the ALIGN program (cf. Materials and Methods). This procedure was repeated to generate interprotein comparisons (Table IV). Specific residue alignments are illustrated in Figure 7. An alignment score of 8.5 established the internal homology of residues 102–219 with residues 220–340 in cGK. Furthermore, these regions, denoted as segments B and C in Table IV, were each homologous with both members of the previously characterized tandem pairs of segments in R_I and R_{II} (Titani et al., 1984; Takio et al., 1984) and with the cAMP-binding segment of the catabolite gene activator protein (CAP) of *E. coli* (Weber et al., 1982). The alignment score for the segments B/C comparison in cGK (8.5) was not as high as that for the corresponding comparison in R_I [11.7 (Titani et al., 1984)] or in R_{II} [12.7 (Takio et al., 1984)], but all of the internal comparisons had higher scores than those derived by comparisons of the six segments to the *E. coli* protein, CAP (Table IV). These data provide evidence that these seven segments of four different proteins are all homologous, implying their divergence from a common ancestral prototype.

The data in Table IV also show that CAP is somewhat more satisfactorily compared with segment C (scores 6.2–7.8) than segment B (scores 3.0–6.8), particularly in cGK, implying that CAP and segment C may have shared a more recent ancestor than has either with segment B.

Comparisons of segments B and C *within each* protein (alignment scores 6.2–8.9) reveal significantly less similarity than comparisons of segment B among the three proteins (15.5–20.9) or of segment C (10.3–17.8). These data imply that internal duplication of some common ancestor of B and C was followed by divergence to the archetypes of segments B and C *before* divergence to contemporary species of R_I, R_{II}, and cGK had occurred.

Discussion

Proof of the structure of cGMP-dependent protein kinase involved primarily two sets of peptides generated by cleavage at either arginyl or methionyl residues. In general the two sets proved to be quite complementary, although the larger fragments required subdigestion for completion of proof of their sequences. The proof relied largely on replicate identifications of phenylthiohydantoins rather than upon the integral numbers in amino acid compositions, particularly when suboptimal amounts of peptides were analyzed. Eighty-two percent of the amino acid residues were identified in more than one peptide, but the remainder were examined only once (e.g., residues 322–347). In all cases independent identifications were made in two complementary HPLC systems, adding strength to our conclusions. The weakest points in the proof of sequence are

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

Peptide		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16	Total	
Residues No.		1-12	13-103	104-118	119-142	143-159	160-200	201	202-212	213-388	389-412	413-435	436-437	438-547	548-563	564-575	576-670	from sequence	from amino acid analysis of cGK
Asp/Asn	(D/N)	1.1(1)	4.6(5)	2.1(2)	1.9(2)	1.1(1)	3.3(3)	0.4(0)	0.6(0)	18.8(20)	1.0(1)	1.9(2)		8.6(9)	1.2(1)	1.6(2)	14.7(19)	68	65.0
Thr	(T)		4.3(5)	0.5(0)	0.4(0)	1.7(2)	4.5(5)		1.4(2)	9.5(9)	0.9(1)	0.9(1)		6.1(7)	1.0(1)	0.9(1)	3.7(4)	38	34.2
Ser	(S)	1.1(1)	6.4(7)	2.2(1)	2.6(2)	1.1(0)		1.0(0)	1.8(0)	8.8(10)	1.0(1)	2.2(2)		5.0(5)	2.0(2)	1.0(0)	5.3(6)	37	32.4
Glu/Gln	(E/Q)	2.7(3)	17.0(17)	3.4(4)	2.0(2)	3.0(3)	3.4(3)		1.2(1)	22.3(23)	5.0(5)	1.3(1)		9.1(7)	1.2(1)	0.5(0)	8.1(7)	77	83.4
Pro	(P)		3.6(4)		0.7(1)		1.1(1)			5.1(4)				2.9(3)	3.9(4)		9.1(10)	27	26.0
Gly	(G)		2.3(2)	1.3(0)	3.0(3)	1.8(2)	3.1(3)	0.8(0)	1.6(1)	15.0(14)	0.6(0)	1.6(1)		10.5(11)	2.2(2)	1.4(1)	5.8(6)	46	44.2
Ala	(A)	1.3(1)	3.7(4)	0.5(0)	0.5(0)		3.1(3)		0.6(0)	16.7(16)		1.0(1)		7.1(8)			4.0(4)	37	37.6
CMCys	(C)		0.4(1)	0.3(1)	0.4(1)	0.6(1)	1.1(2)			0.5(1)				1.6(3)			0.8(1)	11	9.6
Val	(V)		3.5(2)	0.6(1)	2.4(4)	2.4(3)	4.0(4)			12.8(13)	1.0(1)	1.0(1)		4.0(4)			2.1(2)	35	32.6
Met[b]	(M)	0.6(1)	0.6(1)	0.9(1)	0.3(1)	0.5(1)	0.4(1)	1.0(1)	0.3(1)	0.6(1)	0.4(1)	0.6(1)	0.5(1)	0.4(1)	0.5(1)	0.4(1)		15	15.0
Ile	(I)	1.0(1)	6.0(6)	1.6(2)	1.1(2)		2.7(3)		0.8(1)	8.0(10)	3.3(4)	1.0(1)		7.9(9)		1.6(3)	6.0(7)	49	41.9
Leu	(L)	2.0(2)	9.3(9)	2.0(2)	1.1(1)	1.0(1)	4.0(4)		1.0(1)	16.0(16)	1.2(1)	2.0(2)	1.0(1)	11.0(11)	2.0(2)	1.0(1)	6.0(6)	60	60.0
Tyr	(Y)		1.0(1)		1.8(3)		1.5(1)		0.8(1)	2.9(4)		1.9(3)		4.7(6)	1.1(1)	0.7(1)	0.4(0)	21	20.6
Phe	(F)	1.0(1)	4.1(4)				2.1(2)			10.8(12)		1.7(2)		6.1(7)	1.0(1)		5.5(6)	35	33.4
His	(H)		2.9(3)						0.7(1)	2.2(2)	1.7(2)	0.6(1)		2.5(3)			1.0(1)	13	12.2
Lys	(K)	1.1(1)	10.3(11)	1.2(1)	1.1(2)	2.8(3)	2.7(3)		0.7(1)	13.2(13)	3.5(4)	1.7(2)		7.1(8)		0.9(1)	9.1(10)	60	54.9
Arg	(R)		8.6(9)				2.0(2)		0.8(1)	7.8(7)	2.7(3)	2.0(2)		3.6(5)		0.8(1)	3.3(3)	33	32.1
Trp[c]	(W)						(1)			(1)				(3)			(3)	8	
Total		12	91	15	24	17	41	1	11	176	24	23	2	110	16	12	95	670	
Yield %		25	28	34	60	47	25	18	33	36	31	58	46	17	20	27	36		

[a]Residues per peptide by amino acid analysis (6 N HCl, 110°C, 20 h). Values of less than 0.3 are not reported. Numbers in parentheses indicate those found in the sequence (Figure 2). Ratios are calculated to the integral values underlined. One-letter amino acid abbreviations are indicated in parentheses.

[b]Methionine measured as homoserine.

[c]Not determined.

Table II: Source of Subpeptides

subpeptide[a]	residues	digesting enzyme	acetonitrile in eluate[b] (%)
R1-E1	1–5	*S. aureus* protease	14[c]
R1-E4	6–17	*S. aureus* protease	22[c]
R10-T3	95–104	trypsin	18
R10-T4	105–125	trypsin	30
R10-T6	132–146	trypsin	25
R10-T9	156–163	trypsin	13
R11/12-K3	188–202	endoproteinase Lys-C	25[c]
R30-E4	540–549	*S. aureus* protease	29
R30-E5	550–571	*S. aureus* protease	25
R34-Th6	648–667	thermolysin	21
R34-Th6-Pr2	654–667	Pronase	15
R34-C4	667–670	chymotrypsin	19
R34-C2	649–670	chymotrypsin	26

[a] The prefix in the name of a subpeptide denotes the source peptide; the suffix includes the coded name (Figure 2) of the digesting enzyme. [b] HPLC with an acetonitrile gradient on a SynChropak RP-P column except as indicated by footnote *c*. [c] HPLC on a μBondapak C18 column.

Table III. Amino Acid Compositions[a] of Peptides[b] from Terminal Regions

Peptide	R1	R1-E1	R1-E4	R34	R34-C2	R34-C4
Residues	1-17	1-5	6-17	628-670	649-670	667-670
Asx	1.3 (1)		1.0 (1)	10.4 (11)	9.0 (9)	1.8 (2)
Thr				3.5 (4)	0.8 (0)	
Ser	0.9 (1)	1.0 (1)		4.5 (5)	2.5 (2)	
Glx	5.0 (5)	3.0 (3)	2.0 (2)	2.4 (2)	2.4 (2)	
Pro				7.3 (8)	5.2 (4)	
Gly	0.4 (0)			2.5 (2)	1.6 (1)	
Ala	1.2 (1)		1.0 (1)	1.7 (1)	0.7 (0)	
Val				0.9 (1)	1.2 (0)	
Met	0.6 (1)		0.8 (1)			
Ile	1.0 (1)		0.9 (1)	2.8 (3)	1.2 (1)	0.8 (1)
Leu	3.0 (3)	1.0 (1)	2.0 (2)	1.5 (1)		
Phe	1.0 (1)		1.0 (1)	3.0 (3)	1.7 (2)	1.0 (1)
Lys	2.3 (2)		1.9 (2)	1.0 (1)		
Arg	1.1 (1)		0.9 (1)			
Trp				N.D.(1)	N.D.(1)	
Total	17	5	12	43	22	4
Yield %	29	62	25	37	25	20

[a] See footnote [a] in Table I.

[b] The source of the peptides is given in Figure 5 or Table II.

adjacent to residues 202, 461, 550 and 667. In these cases the overlap was two residues or less, and it was necessary to support the alignments with compositional arguments somewhat lacking in rigor.

The protein contains 670 residues, corresponding to a molecular weight of 76331. This is in good agreement with prior reports of the molecular weight which ranged from 74000 to 81000 (Gill et al., 1977; Lincoln et al., 1977). The molecular weight in solution is approximately twice this value, and it is now clear that the two polypeptide chains in the dimer are identical.

Limited Proteolysis of cGK. Inoue et al. (1976) were the first to show that cGK from silkworm pupae was activated by limited proteolysis with trypsin, which cleaved the 140000-dalton dimer into a 34000-dalton catalytic fragment and a 36000-dalton cGMP-binding fragment. In contrast to the native molecule, the catalytic fragment was not stimulated by cGMP, indicating separation of the regulatory and catalytic functions. Later, Lincoln et al. (1978) reported that limited treatment of bovine cGK with trypsin reduces the molecular weight of the monomer from 80000 to 75000 in a fragment with analogous cGMP-independent activation of the enzyme and inhibition of its tendency to autophosphorylate. Monken & Gill (1980) demonstrated that chymotryptic treatment generated a 16000-dalton disulfide-linked dimer and a 65000-dalton monomeric species which was catalytically active and capable of binding cGMP. The smaller fragment was blocked at its amino terminus and presumed to include the amino terminus of the protein. Takio et al. (1983) observed similar fragmentation upon protease treatment but suggested the disulfide might be an artifact formed during limited proteolysis. Whatever its origin, our present structural data place the disulfide bond between Cys residues 42 in neighboring monomers and the chymotryptic cleavage sites at residues 70 and 80. Furthermore, it appears that the oligomeric interaction sites of cGK must be in the amino-terminal 70 residues of each chain.

Table IV: Internal Homology[a] in cGK in Relation to Other cAMP Binding Domains[b]

segments compared[c]	cGK, segment B (residues 102–219)	cGK, segment C (residues 220–340)	CAP[d] (residues 1–130)	CAP[d] (residues 131–209)
cGK segment B		8.5 (28%)	3.0 (21%)	
cGK segment C	8.5 (28%)		6.2 (25%)	0.9 (18%)
cGK segment D				1.1 (25%)
R_I segment B	18.7 (37%)	6.2 (31%)	6.8 (20%)	
R_I segment C	8.9 (25%)	13.8 (34%)	7.8 (22%)	0.5 (10%)
R_{II} segment B	15.5 (31%)	6.9 (26%)	6.4 (18%)	
R_{II} segment C	8.5 (31%)	10.3 (36%)	7.0 (26%)	0.6 (18%)

[a] Similarities in sequence are recorded as alignment scores (see Materials and Methods) and, in parentheses, as percentage identity. [b] Comparisons within and between R_I and R_{II} are separately reported (Titani et al., 1984; Takio et al., 1984). [c] Segments are defined diagramatically in Figure 8. For R_I segments B and C are residues 135–252 and 253–379; for R_{II} segments B and C are residues 135–256 and 257–400. [d] CAP denotes the catabolite gene activator protein of *E. coli* which binds cAMP in a domain within residues 1–130 (Weber et al., 1982).

The autophosphorylation site in cGK is in the amino-terminal region near the site of limited proteolysis, as it is in the regulatory subunit of type II cAK (Takio et al., 1983). The present data place this site on Thr-58 in cGK, as compared to Ser-95 in cAK. In each protein, limited proteolysis severs the autophosphorylated site from the molecule in a small amino-terminal fragment (~6000–10000 daltons), indicating that the sites of phosphorylation, of limited proteolysis, and of oligomeric association are in comparable amino-terminal regions of the two molecules. Several studies have indicated that additional sites are phosphorylated in the presence of cAMP (e.g., Foster et al., 1981; Hofmann & Flockerzi, 1983), but identification of these sites has not yet been reported.

ATP-Binding Site of cGK. The present sequence data indicate that it is Lys-389 that was labeled by Hashimoto et al. (1982) with an ATP analogue. A homologous sequence is seen both in the catalytic subunit of cAK (Zoller & Taylor, 1979; Shoji et al., 1983) and in the γ subunit of phosphorylase *b* kinase (Reimann et al., 1984), but in these proteins the corresponding lysine residues are found only 72 and 48 residues, respectively, from the amino termini of their catalytic subunits. These data imply that the ATP-binding site of cGK is dis-

FIGURE 7: Alignment of segments B and C of cGK with the corresponding segments of R_I, R_{II}, and CAP. The residue numbers at the left are within the segments defined in Figure 8; dashes indicate gaps; the alignment scores and percentage identities are listed in Table IV. Although the various segments B and C are homologous, CAP is aligned only with the segment C. Boxes enclose residues that are identical with cGK in at least one other protein. If three proteins other than cGK have identical residues, they are also boxed. Alignments near the ends of segments were not always unambiguous and are omitted. The sequences of R_I, R_{II}, and CAP are from Titani et al. (1984), Takio et al. (1984), and Weber et al. (1982), respectively.

placed more than 300 residues toward the carboxyl terminus as compared to the other kinases.

Homology of cGK with Other Proteins. It has been suggested on the basis of amino acid compositions, binding characteristics, and molecular asymmetry that cGK and cAK might be homologous proteins even though they are of very different size (Gill, 1977; Lincoln & Corbin, 1977). As mentioned above, the small segments of sequence determined by Hashimoto et al. (1982) and by Takio et al. (1983) around the ATP binding and autophosphorylation sites of cGK are similar to corresponding segments of cAK. Our present data provide complete substantiation of the idea and extend the evidence of homology to several other proteins, some of which appear to have only a remote functional relationship to cGK. We will develop the hypothesis that the amino-terminal half of the cGK molecule contains two domains that bind cGMP and that the carboxyl-terminal half serves as the catalytic domain. Other proteins (or subunits) which exercise only one of these two functions will be shown to be homologous with the appropriate half of the cGK molecule, raising questions of recombination and splicing events in the ancestral DNA that may have been involved in the primordial genesis of this complex 670-residue allosteric molecule. To illustrate the relationship of cGK to the other proteins, it is useful to consider first the molecule as if it were a composite of six somewhat arbitrarily divided tandem segments, denoted A–F in Figure 8, so that the functions and homologous relationships of each segment can be defined and in several regions related to possible substructural domains. We will then return to broader

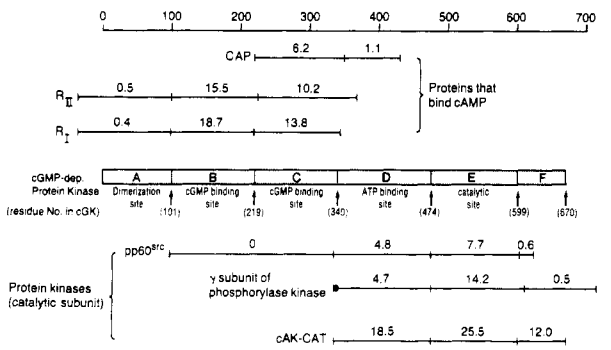

FIGURE 8: Six somewhat arbitrarily chosen segments (A–F) of cGK (central, segmented bar) are aligned and compared with corresponding segments of homologous proteins. Each protein is indicated as a segmented line, above each segment of which is its alignment score with the vertically aligned segment of cGK. Residues included in the boxed segments of cGK are indicated in parentheses below them. In the other proteins, successive segments from the amino terminus end at the following residues number: CAP, 130, 219; $pp60^{src}$, 236, 377, 488, 526; γ subunit of phosphorylase *b* kinase, 140, 276, 386; R_{II}, 134, 256, 400; R_I, 134, 252, 379; catalytic subunit of cAMP-dependent protein kinase (cAK-CAT), 157, 278, 350. Functions associated with each segment of cGK are discussed in the text and indicated below that segment. Segments B and C may correspond to folding domains in each protein; segments D and E probably correspond to a single (catalytic) domain in each protein and are arbitrarily divided into two to facilitate comparisons.

questions of the relationships of the whole molecules to each other and of mechanisms that may have generated this pattern of interrelated proteins encompassing of 6-fold range of mo-

Table V: Homology[a] of Segments of cGK to Segments of Other Protein Kinases[b]

segments compared[c]	residues	cGK, segment A (residues 1–101)	cGK, segment D (residues 330–474)	cGK, segment E (residues 475–599)	cGK, segment F (residues 600–670)
R_I, segment A	1–134	0.4 (27%)			
R_{II}, segment A	1–134	0.5 (10%)			
PbKγ, segment D	1–140		4.7 (18%)		
PbKγ, segment E	141–276			14.2 (39%)	
PbKγ, segment F	279–386				0.5 (17%)
cAK, segment D	1–157		18.5 (40%)		
cAK, segment E	158–279			25.5 (47%)	
cAK, segment F	280–350				12.0 (39%)
pp60^{src}, segment D	237–377		4.8 (21%)		
pp60^{src}, segment E	378–507			7.7 (29%)	
pp60^{src}, segment F	489–526				0.6 (16%)

[a] Same as footnote *a* in Table IV. [b] Comparisons of PbKγ with cAK and with pp60^{src} are reported separately (Reimann et al., 1984). [c] PbKγ refers to the γ subunit of rabbit muscle phosphorylase *b* kinase and pp60^{src} to the transforming phosphoprotein of Rous sarcoma virus (Barker & Dayhoff, 1982).

lecular weights from 12 000 to 76 000.

Segment A includes both the oligomeric interaction site and the autophosphorylation site in cGK. The latter site (Thr-58) is near one accessible to chymotrypsin (Tyr-70), and it has been suggested that an interdomain hinge region exists within segment A (Takio et al., 1983). Similar observations in the amino-terminal segments of R_I and R_{II} (Titani et al., 1984; Takio et al., 1984) led to analogous conclusions of possible domain substructure. Although these segments of R_I and R_{II} appear to be homologous with each other (Takio et al., 1984), an objective search for homology of each with segment A of cGK yields little evidence that these alignments have statistical signifcance (Figure 8).

In contrast, segments B and C of cGK are not only homologous with each other, but also with the corresponding segments of R_I and R_{II} (Table IV). This is illustrated diagramatically in Figure 8, and the specific residue alignments are detailed in Figure 7. Weber et al. (1982) have shown that the cAMP-binding protein, CAP, is also homologous to segments B and C of R_{II}, and the present analysis extends this homology to cGK and R_I (Table IV). Significance is added to this observation by knowledge of the three-dimensional structure of the cAMP-binding domain of CAP (McKay & Steitz, 1981) and by observations that cGK, R_I, and R_{II} each bind two molecules of cyclic nucleotide (Corbin & Døskeland, 1983; Weber & Hilz, 1979; Corbin et al., 1978). It is thus reasonable to propose that each of the three kinase proteins contains two such domains, each of which is folded much like the corresponding domain of CAP to provide a cyclic nucleotide binding site. In cAK these domains would comprise the bulk of each regulatory subunit, whereas in the longer chain cGK they would comprise regulatory domains within the allosteric chain. It will be interesting to see if the hypothesis of three-dimensional similarity of domains, first proposed for R_{II} by Weber et al. (1982), extends to yet other proteins which bind cyclic nucleotides. As discussed previously (Takio et al., 1984), the analogous regulatory subunit of cAMP-dependent protein kinase in *Dictyostelium discoideum* appears to be 10 000 daltons smaller than R_{II}, binds only one cAMP, and is capable of interaction with the bovine catalytic subunit (Majerfeld et al., 1984). It is possible that that regulatory subunit contains a single folding domain homologous with segments B and C, implying that a protein in a common ancestor of slime mold, cows, and *E. coli* had a single binding site for cAMP.

Segments D–F constitute three somewhat arbitrarily divided regions of the catalytic domain of cGK. The ATP-binding site has been identified at Lys-389 within segment D. The catalytic subunit of cAMP-dependent protein kinase is homologous through all three segments (Table V; Figure 8), as shown in detail in Figure 9. Recently other proteins have been reported to be homologous with cAK, the γ subunit of phosphorylase *b* kinase (Reimann et al., 1984), the epidermal growth factor receptor (Downward et al., 1984), and certain members of a family of transforming retroviral, tyrosine-specific protein kinases, typified by pp60^{src} [e.g., Barker & Dayhoff, 1982; Groffen et al., 1983; Privalsky et al., 1984; reviewed in Sefton & Hunter (1984)]. Each of these proteins contains segments of sequence that correspond to cGK segments D and E although the alignment scores indicate less similarity than between cGK and cAK (Table V; Figure 8). As mentioned previously by Reimann et al. (1984), the carboxyl-terminal 110-residue segment of the γ subunit of PbK lacks homology with the catalytic subunit of cAK, although portions corresponding to segments D and E have high alignment scores. Analogous dichotomies appear here in the comparison of cGK with PbKγ and with pp60^{src}. Although segments D and E of cGK are clearly homologous with both proteins, segment F shows no significant similarity to the corresponding region of either (Table V), thus providing a rationale for speculation that segment F might form a separate folding domain of undetermined function (Figure 8). There is no corresponding rationale to distinguish segment D from E. They are considered separately in Figure 8 and Table V simply to divide the protein in roughly equal segments. They could well form a single folding domain in each protein.

The retroviral products vary in length, depending upon the particular member of the family examined. In the case of pp60^{src}, alignment with segments D and E of cGK leaves a large amino-terminal segment (Figure 8) which shows no detectable similarity to segment C (or any other segment) of cGK. Thus, it appears that this protein contains in its interior about 250 residues of sequence homologous with cGK, cAK, and PbKγ, and an equal number of residues surrounding that region which are unrelated to segments C and F of the kinase sequences. This is consistent with the view that retroviral nucleic acids contain segments of "captured" eukaryotic genetic information within the viral message (Bishop, 1983).

Chimeric Nature of cGK. To understand the genesis of relationships among the seven proteins in Figure 8, it is necessary to consider not only the occurrence of gene duplication and divergent evolution but also the probability that whole segments of genes have been excised, shuffled, and spliced. How else can one explain that segment D is totally different in CAP (where it binds DNA) as compared to the kinases, that the carboxyl-terminal segments of pp60^{src} and PbKγ are totally unrelated to cAK and cGK, and, most importantly, that the sequence of cGK comprises elements of sequences which

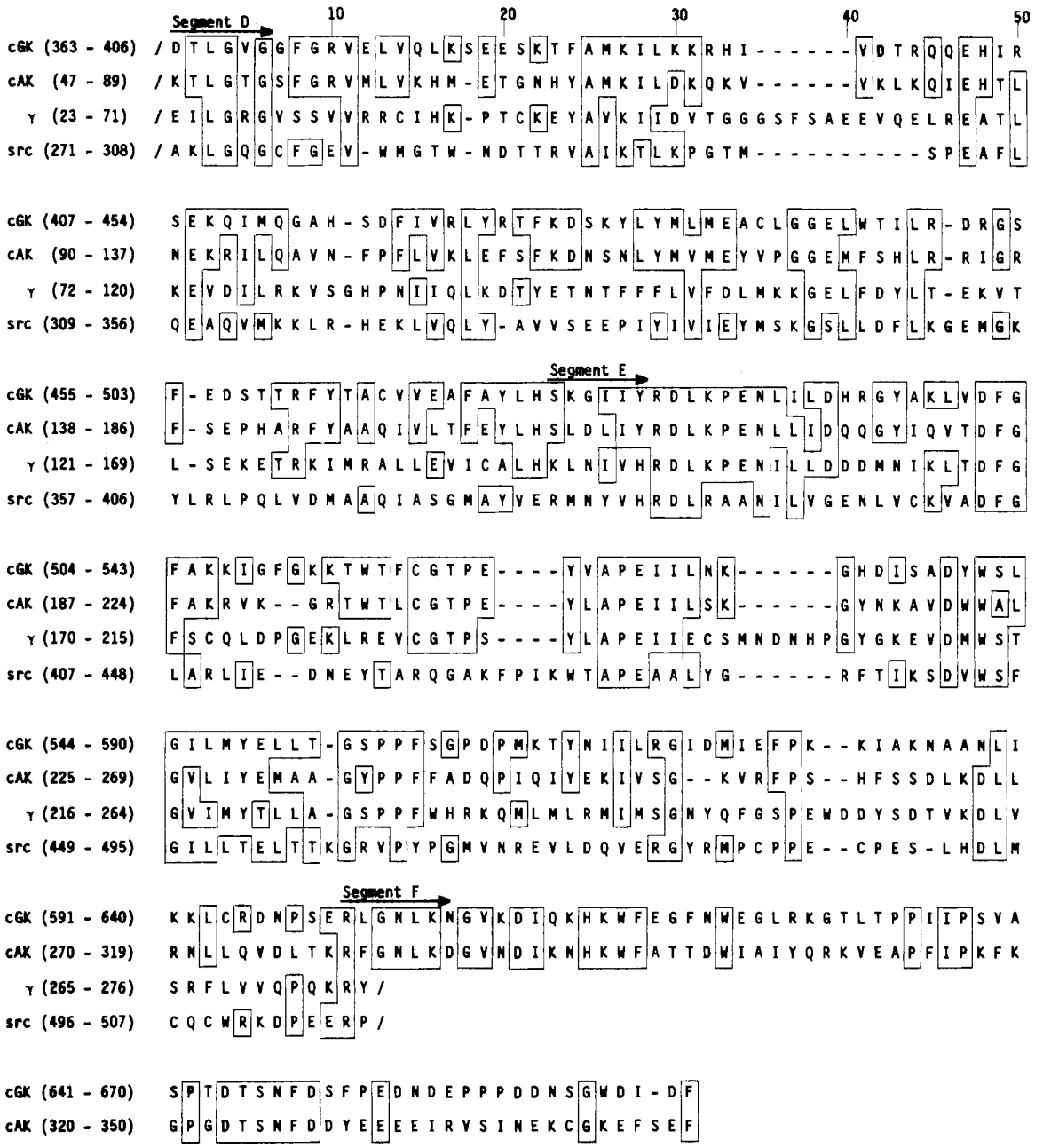

FIGURE 9: Alignment of segments D, E, and F of cGK (Figure 8) with the corresponding segments of cAK, pp60^{src} (src), and the γ subunit of phosphorylase *b* kinase (γ). Boxes enclose identities as in Figure 7. The terminal portions of segments are not necessarily included. The sequence of pp60^{src} is taken from Schwartz et al. (1983), that of γ from Reimann et al. (1984), and that of cAK from Shoji et al. (1983).

are homologous with two distinct and separate families of proteins, those binding cyclic nucleotides and those acting as kinases?

One could suggest that proteolytic processing has led to the two subunits of cAK from a longer precursor protein. However, the length of the cDNA of R_I (Lee et al., 1983) is in accord with our amino acid sequence, ruling out this possibility. Alternatively either an ancestral gene of cAK coded for both cAMP-binding and catalytic functions of a large protein (like cGK) and evolution separated these functions on separate genes, or separate genes were spliced to produce an ancestor of contemporary cGK.

First, let us consider the idea of a large ancestral gene acquiring both functions and evolving into separate genes. Presumably the genesis of the master gene included an internal gene duplication event giving rise to the tandem segments B and C. A fragment of CAP would have diverged at an early stage from primordial segment C and become fused to a DNA-binding segment, and the D/E segments found in the retroviral family would have resulted from capture and insertion into other coding sequences. Even this scenario requires several gene fusion events, but its overall direction is from a multifunctional protein to several simpler proteins. It seems improbable that a complex allosteric protein would have evolved before its simpler counterparts.

Alternatively one could imagine the separate evolution of a catalytic entity in a primordial protein kinase that, by gene duplication and divergent evolution, gave rise to the progenitors of the catalytic subunits of cAK, PbKγ, and the cellular versions of the retrovirally coded proteins. Separately, a family of proteins that bound cyclic nucleotides, for whatever purpose, could have evolved and undergone internal duplication. Obvious regulatory advantage would accrue to an organism from association of these two proteins either as heterologous oligomers (as in the isozymes of cAK) or by gene fusion (as in cGK). Meanwhile various versions of the catalytic entity could

associate with calcium/calmodulin-regulated proteins (e.g., in PbKγ), evolve tyrosyl specificity (e.g., in pp60^{src} or its cellular counterpart), or appear in yet unrecognized forms.

A precedent for such rearrangements of segments of genomes has been reported in recent years, although the underlying mechanisms are by no means understood [cf. review by Doolittle (1981)]. Gilbert (1978) has suggested a role for introns in the fusion events. Spliced species of haptoglobin (Dixon, 1966), Lepore hemoglobin (Baglioni, 1962), and γ-globulin (Tonegawa, 1983) are well-known but may speak more to mechanisms of internal gene duplication than to fusion of dissimilar segments. Single-chain, multienzyme conjugates represent better analogies, and many have now been reported (Kirshner & Bisswanger, 1976). Gaertner (1978) and Welch & De Moss (1978) have discussed the functional advantages of multienzyme clusters, whether associated as noncovalent complexes or fused in polypeptide conjugates. Recently artifically generated chimeric proteins have been produced in the laboratory by splicing coding regions (e.g., Bedouelle et al., 1980). Perhaps the most notable examples of chimeric enzymes are found in the family of fatty acid synthases from various organisms [cf. reviews by Wakil et al. (1983) and McCarthy & Hardie (1984)], where multienzyme conjugates of seven different proteins are found to be fused in either one or two chains in sequential arrangements which differ from one organism to another.

Another general example which demands analogous underlying splicing mechanisms is found among the zymogens of the serine proteases (Neurath, 1984), which vary in length from the 229 residues of trypsinogen (Titani et al., 1975) to the 790 residues of plasminogen (Sottrup-Jensen et al., 1978) and occur in both one-chain and two-chain versions (Katayama et al., 1979), with and without so-called "kringle" domains or domains rich in γ-carboxyglutamyl residues. This family of proteins is characterized by their homologous active site domains, but one can only speculate about the splicing mechanisms that must have led to the variations in their content of other domains. A recent report concerning this family suggests that tissue plasminogen activator may contain elements of sequences of both serine proteases and fibronectin (Banyai et al., 1983). In a similar vein, one can cite the occurrence of collagen-like sequences within complement factor C1q (Reid & Porter, 1976).

Finally, several recent reports indicate that segments resembling domains of immunoglobulins are substructural elements within the polypeptide chains of the major histocompatibility antigen complexes, T-lymphocyte antigen receptors [cf. review by Williams (1984)], and receptors for transcellular transport of polymeric immunoglobulins into external secretions (Mostov et al., 1984). In each case, sequence data suggest that immunoglobulin-like domains have been spliced into chimeric chains serving complex functions which probably rely in part on the recognition capability dictated by the immunoglobulin-like domains.

Returning to the protein kinase family (Figure 8), the simplest explanation of this interrelated series of proteins involves both specific gene fusion and gene duplication events in the distant history of their development. The divergence of their sequences during evolution has been so extensive that it is difficult to deduce a probable order of events but has not been enough to disguise the segments of homology which underlie the present speculation about the source of these relationships. Natural extensions of the homologies shown in Figure 8 lead to predictions that other protein kinases, whatever their mode of regulation, may share a variant of the active site domain and that other proteins which interact with cyclic nucleotides may represent spliced or mutated variants of the regulatory domains of cGK. In a more general sense, the patterned relationships among these proteins suggest that other unrelated proteins may also be assembled by analogous splicing of yet other domains to generate complex proteins with interactive catalytic, binding, and/or regulatory functions.

Acknowledgments

We are indebted to Drs. Hans Neurath and Edmond Fischer for suggestions and encouragement, to Santosh Kumar and Lowell Ericsson for help in many phases of the investigation, to Richard Granberg, Brad McMullen, Barbara Flug, and Floyd Kennedy for excellent technical assistance, and to Mary Woods for her patience and skill in typing the manuscript.

References

Baglioni, C. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 1880.
Banyai, L., Varadi, A., & Patthy, L. (1983) *FEBS Lett. 163*, 37.
Barker, W. C., & Dayhoff, M. O. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 2836.
Bedouelle, H., Bassford, P. J., Jr., Fowler, A. V., Zabin, I., Beckwith, J., & Hofnung, M. (1980) *Nature (London) 285*, 78.
Bishop, J. M. (1983) *Annu. Rev. Biochem. 52*, 301.
Corbin, J. D., & Døskeland, S. O. (1983) *J. Biol. Chem. 258*, 11391.
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.
deJonge, H. R., & Rosen, O. M. (1977) *J. Biol. Chem. 252*, 2780.
Dixon, G. H. (1966) in *Essays in Biochemistry* (Campbell, P. N., & Preville, G. D., Eds.) Vol. 2, p 147, Academic Press, New York.
Doolittle, R. F. (1981) *Science (Washington, D.C.) 214*, 149.
Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J., & Waterfield, M. D. (1984) *Nature (London) 307*, 521.
Erlichman, J., Rosenfeld, R., & Rosen, O. M. (1974) *J. Biol. Chem. 249*, 5000.
Flockhart, D. A., & Corbin, J. D. (1982) *CRC Crit. Rev. Biochem. 12*, 133.
Foster, J. L., Guttmann, J., & Rosen, O. M. (1981) *J. Biol. Chem. 256*, 5029.
Gaertner, F. H. (1978) *Trends Biochem. Sci. (Pers. Ed.) 3*, 63.
Gilbert, W. (1978) *Nature (London) 271*, 501.
Gill, G. N. (1977) *J. Cyclic Nucleotide Res. 3*, 153.
Gill, G. N., Walton, G. M., & Sperry, P. J. (1977) *J. Biol. Chem. 252*, 6443.
Glass, D. B., & Krebs, E. G. (1979) *J. Biol. Chem. 254*, 9728.
Glass, D. B., & Krebs, E. G. (1980) *Annu. Rev. Pharmacol. Toxicol. 20*, 363.
Groffen, J., Heisterkamp, N., Reynolds, F. H., Jr., & Stephenson, J. R. (1983) *Nature (London) 304*, 167.
Hashimoto, E., Takio, K., & Krebs, E. G. (1982) *J. Biol. Chem. 257*, 727.
Hofmann, F., & Flockerzi, V. (1983) *Eur. J. Biochem. 130*, 599.
Hunter, T., & Sefton, B. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1311.

Inoue, M., Kishimoto, A., Takai, Y., & Nishizuka, Y. (1976) *J. Biol. Chem. 251*, 4476.
Katayama, K., Ericsson, L. H., Enfield, D. L., Walsh, K. A., Neurath, H., Davie, E. W., & Titani, K. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4990.
Kirshner, K., & Bisswanger, H. (1976) *Annu. Rev. Biochem. 45*, 143.
Krebs, E. G., & Beavo, J. A. (1979) *Annu. Rev. Biochem. 48*, 923.
Lee, D. C., Carmichael, D. F., Krebs, E. G., & McKnight, G. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3608.
Lincoln, T. M., & Corbin, J. D. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3239.
Lincoln, T. M., Dills, W. L., & Corbin, J. D. (1977) *J. Biol. Chem. 252*, 4269.
Lincoln, T. M., Flockhart, D. A., & Corbin, J. D. (1978) *J. Biol. Chem. 253*, 6002.
Mahoney, W. C., & Hermodson, M. A. (1980) *J. Biol. Chem. 255*, 11199.
Majerfeld, I. H., Leichtling, B. H., Meligeni, J. A., Spitz, E., & Rickenberg, H. V. (1984) *J. Biol. Chem. 259*, 654.
McCarthy, A. D., & Hardie, D. G. (1984) *Trends Biochem. Sci. (Pers. Ed.) 9*, 60.
McKay, D. B., & Steitz, T. A. (1981) *Nature (London) 290*, 744.
Monken, C. E., & Gill, G. N. (1980) *J. Biol. Chem. 255*, 7067.
Mostov, K. E., Friedlander, M., & Blobel, G. (1984) *Nature (London) 308*, 37.
Nakai, N., Lai, C. H., & Horecker, B. L. (1974) *Anal. Biochem. 58*, 563.
Neurath, H. (1984) *Science (Washington, D.C.) 224*, 350.
Privalsky, M. L., Ralston, R., & Bishop, J. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 704.
Reid, K. B. M., & Porter, R. R. (1976) *Biochem. J. 155*, 19.
Reimann, E. M., Titani, K., Ericsson, L. H., Wade, R. D., Fischer, E. H., & Walsh, K. A. (1984) *Biochemistry* (accompanying paper in this issue).
Schwartz, D. E., Tizard, R., & Gilbert, W. (1983) *Cell (Cambridge, Mass.) 32*, 853.
Sefton, B. M., & Hunter, T. (1984) *Adv. Cyclic Nucleotide Protein Phosphorylation Res. 18*, 195.
Shoji, S., Ericsson, L. H., Walsh, K. A., Fischer, E. H., & Titani, K. (1983) *Biochemistry 22*, 3702.
Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T. E., & Magnusson, S. (1978) in *Progress in Chemical Fibrinolysis and Thrombolysis* (Davidson, J. F., Rowan, R. M., Samama, M. M., & Desnoyers, P. C., Eds.) Vol. 3, pp 191–209, Raven Press, New York.
Takai, Y., Nakaya, S., Inoue, M., Kishimoto, A., Nishiyama, K., Yamamura, H., & Nishizuka, Y. (1976) *J. Biol. Chem. 251*, 1481.
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. (1984) *Biochemistry* (second paper of three in this issue).
Titani, K., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1975) *Biochemistry 14*, 1358.
Titani, K., Takio, T., 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* (first paper of three in this issue).
Tonegawa, S. (1983) *Nature (London) 302*, 575.
Ushiro, H., & Cohen, S. (1980) *J. Biol. Chem. 255*, 8363.
Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983) *Annu. Rev. Biochem. 52*, 537.
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.
Welch, G. R., & De Moss, J. A. (1978) in *Microenvironments and Metabolic Compartmentation* (Srere, P. A., & Estabrook, R. W., Eds.) p 323, Academic Press, New York.
Williams, A. F. (1984) *Nature (London) 308*, 108.
Zoller, M. J., & Taylor, S. S. (1979) *J. Biol. Chem. 254*, 8363.