Plant Physiol. (1984) 75, 788–795
0032-0889/84/75/0788/08/$01.00/0

Structural Characterization of a Higher Plant Calmodulin[1]

SPINACIA OLERACEA

Received for publication January 10, 1984 and in revised form March 21, 1984

THOMAS J. LUKAS*, DAVID B. IVERSON[2], MICHAEL SCHLEICHER[3], AND D. MARTIN WATTERSON
Laboratory of Cellular and Molecular Physiology, Howard Hughes Medical Institute, and Department of Pharmacology, Vanderbilt University, Nashville, Tennessee 37232

ABSTRACT

Calmodulin is a eukaryotic calcium binding protein which has several calcium-dependent *in vitro* activities. Presented in this report is a structural characterization of calmodulin from spinach leaves (*Spinacia oleracea*). Spinach calmodulin may be representative of higher plant calmodulins in general since calmodulin from the monocotyledon barley (*Hordeum vulgare*) is indistinguishable by a variety of physical, chemical, and functional criteria (Schleicher, Lukas, Watterson 1983 Plant Physiol 73: 666–670). Spinach calmodulin is homologous to bovine brain calmodulin with only 13 identified amino acid sequence differences, excluding a blocked NH_2-terminal tripeptide whose sequence has not been elucidated. Two extended regions of sequence identity are in the NH_2-terminal half of the molecule, while nine of the 13 identified differences are in the COOH-terminal half of the molecule. Two of the changes, a cysteine at residue 26 and a glutamine at residue 96, require a minimum of two base changes in the nucleotide codons. Both of these changes occur in the proposed calcium binding loops of the molecule. Five additional amino acid differences found in spinach calmodulin had not been observed previously in a calmodulin. As described in an accompanying report (Roberts, Burgess, Watterson 1984 Plant Physiol 75: 796–798), these limited number of amino acid sequence variations appear to result in differential effects on the activation of calmodulin-dependent enzymes by plant and vertebrate calmodulins.

Calmodulin is a eukaryotic Ca binding protein which modulates the activity of a number of enzymes *in vitro* (13). Complete amino acid sequences have been determined for bovine brain (12, 27, 29) and human brain calmodulin (21). Nearly complete sequences have been published for rabbit skeletal muscle (6), bovine uterus (6), rat testis (4), sea anenome (24), *Renilla* (10), *Tetrahymena* (30), and scallop (25) calmodulins. In addition, protein sequences of the electric eel (15) and chicken (19) calmodulins have been derived from cDNA structures. These data reveal that, except for amidation states, calmodulins may be identical among mammals and avian species and that most of the known sequence differences in other species are conservative substitutions which could be explained by single nucleotide changes in the DNA codons.

A Ca-dependent protein activator in plants was discovered several years ago (1, 18) and subsequently calmodulins from spinach (28), barley (22), and peanut (1) have been chemically characterized. The amino acid compositions of plant calmodulins indicate some differences in tyrosine, threonine, leucine, arginine, and lysine content compared to bovine calmodulin. Initial characterization of spinach calmodulin illustrated that it contained a single trimethyllysine residue like bovine and most other calmodulins. The presence of cysteine has been reported in spinach and peanut calmodulins (1, 28). Recently, we reported that barley calmodulin is indistinguishable from spinach calmodulin by the criteria of peptide mapping, amino acid composition, and phosphodiesterase activator activity. Thus, spinach calmodulin may be a prototypical higher plant calmodulin. Further chemical characterization of a plant calmodulin was undertaken to determine what differences might exist in the primary structure of the plant protein compared to other known calmodulins. Any functional differences must be related to the covalent structure which, in turn, determines the three-dimensional structure. Presented here is amino acid sequence data and homologous alignments which provide a proposed primary structure of spinach calmodulin. Certain portions of the amino acid sequence of spinach calmodulin have appeared elsewhere (3, 27).

MATERIALS AND METHODS

General. Deionized H_2O was from a Darco or Milli-Q water purification system. Buffer reagents were analytical grade and acetonitrile for liquid chromatography was from Fisher or Burdick and Jackson. HPLC columns were from the sources indicated for the separations described below. Cyanogen bromide was from Eastman.

Spinach leaves were obtained from a local supplier. Spinach calmodulin was isolated as described previously (28), with the following changes. Buffer H was 50 mM Tris-HCl, 5 mM EGTA[4], 1 mM β-mercaptoethanol, and 1% (w/v) PVPP (pH 7.5). Buffer B was 20 mM Tris-HCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, 0.2 M NaCl (pH 7.5). After affinity-based chromatography over phenothiazine-Sepharose (17), the final step in the purification of spinach calmodulin was semipreparative HPLC on a Waters μBondapak Phenyl reverse phase column (0.78 × 30 cm). The proteins were eluted from the column at a flow rate of 2.0 ml/min in a 10 mM K-phosphate buffer, 0.5 mM EGTA, at pH 6.1, and acetonitrile. The sample was loaded onto the column equilibrated in 20% (v/v) acetonitrile and step eluted with 29% (v/v) acetonitrile after 1 column volume. The protein pools were then dialyzed against 10 mM ammonium bicarbonate and lyophilized. The yield of protein from this step of purification was 68 to 75%.

[1] Supported in part by a grant from the National Science Foundation, PCM 8242875.

[2] Current address: New Jersey Department of Health, Trenton, NJ 08625.

[3] Current address: Max Plank Institute for Psychiatry, Am. Klopferspitz 18a, D-8033 Martensreid, Munich, West Germany.

[4] Abbreviations: EGTA, ethyleneglycol bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid; ODS, octadecylsilanyl; Pth, phenylthiohydantoin; Hse, homoserine; PVPP, polyvinylpolypyrollidone.

Spinach calmodulin was stored as a lyophilized powder at −70°C. Stock solutions and peptides were stored at −20°C.

Automated Edman degradations were performed with three different sequenators. A Beckman 890 spinning cup sequencer was used to sequence peptides from trypsin digestion of citraconylated spinach calmodulin. A Sequemat solid-phase sequencer was used for some peptides from cyanogen bromide digests of spinach calmodulin. Homoserine peptides were coupled to triethylenetetraamine resin (TETA, Sequemat) after cyclization with anhydrous TFA, as described (16). An Applied Biosystems model 470A sequencer was also used following the general protocol described (9), and the standard program supplied with the instrument. Glass fiber filters were pretreated with 29 μl of a 60 mg/ml solution of polybrene (Pierce) and run for six cycles. The peptide sample was then loaded onto the filter in 30-μl aliquots.

Amino acid analyses were determined after acid hydrolysis in 6 N HCl (Ultrex, Baker) on Durrum D500 or LKB4400 instruments. Trimethyllysine was determined as described previously (26).

Phenylthiohydantoin (Pth) derivatives of amino acids were identified by HPLC by a modification of a method (31) using a Zorbax (Dupont) octadecyl silanyl (ODS) silica column (4.1 × 250 mm), and mobile phases of 50 mM sodium acetate (pH 5.0) and acetonitrile. Separation parameters were optimized from column to column on a Hewlett Packard 1084B liquid chromatograph, operated at a flow rate of 1.5 ml/min and a column temperature of 60°C. Data collection and reduction were on a Hewlett Packard 3356/3357 Laboratory Automation system.

Proteolytic Digests. Spinach calmodulin was citraconylated in 60% yield following the general method of Atassi *et al.* (2). Citraconyl spinach calmodulin (6.5 mg, 390 nmol) was suspended in 2.0 ml of 0.1 M NH_4HCO_3, 1 mM EGTA, and the pH immediately adjusted to 8.5. Trypsin (TPCK, Worthington), 300 μg, was added and the solution incubated with stirring for 3 h at 37°C. The reaction was terminated by adding 600 μg of soybean trypsin inhibitor (Sigma). After lyophilization, the citraconyl groups were removed by suspending the peptides in 10% (v/v) HCOOH at 37°C for 2.5 h. The mixture was then evaporated to dryness with a rotary evaporator and the residue suspended in 50% (v/v) acetic acid. The soluble supernatant was lyophilized and the acid-insoluble pellet dissolved in water adjusted to pH 7.5 with 1 M NaOH. The acid-soluble pool of peptides was fractionated on a reverse phase column (Whatman ODS-2, 9.4 × 250 mm) using a mobile phase of 0.01 N HCl and acetonitrile. Peptides were eluted with a multistep gradient program described below. Peptide pools not completely resolved in the first preparative fractionation were rechromatographed in the same solvent system.

Trypsin digests of alkylated or native spinach calmodulin and peptide subdigests were done in a buffer containing 0.1 M ammonium bicarbonate, 1 mM EGTA (pH 7.9). Substrate concentrations were 0.4 to 2.5 mg/ml and the final enzyme concentration was 1:50 (w/w) of the substrate. The enzyme was added in two equal aliquots over a 12- to 16-h incubation period at 37°C. Reactions were terminated by lyophilization. Cleavage at arginine residues was mediated by Arg Proteinase C (Boehringer-Mannheim), a mixture of proteases from mouse submaxillary glands. The substrate concentrations and buffer were the same as that used for trypsin digests except that the final enzyme concentration was 1:25 (w/w) of substrate. In these digests, two to three additions of enzyme were used over the time course of the reaction which was followed by intermittent sampling and analytical HPLC. A 200-nmol preparation of spinach calmodulin required 48 h of digestion to complete the reaction. Cyanogen bromide cleavage was done as described (29). Thermolysin (Calbiochem) digests were done for 36 h in 0.1 M ammonium bicarbonate in the absence of chelator.

Peptide Separations. Separations of Arg protease, trypsin, and cyanogen bromide cleavage peptides of spinach calmodulin were all done with a reverse phase column (Whatman ODS-3 column 10 mm × 250 mm). The solvent system was A: 0.01 N HCl, and solvent B: acetonitrile (Fisher). Flow rate was 2.0 ml/min. The effluent was monitored at 215 nm. The column was equilibrated initially with 5% acetonitrile. After 5 min, the %B was increased linearly between isocratic steps in the following order: 5.5 min, %B = 10; 10 min, %B = 10; 20 min, %B = 25; 30 min, %B = 25; 40 min, %B = 35; 50 min, %B = 35; 60 min, %B = 45; 65 min, %B = 45; 70 min, %B = 60; 80 min, %B = 60. The characteristics of this program control can be found in Hewlett Packard 1084B technical literature. This program was also used on the ODS-2 column for the separation of peptides from the citraconylated spinach calmodulin digest. Separation of peptides from other trypsin digests used the same program except that the 20- and 30-min steps were to 22% B rather than 25% B. For separation of the cyanogen bromide cleavage digest, a different program was used: 10 min, %B = 5; 20 min, %B = 20; 60 min, %B = 35; 65 min, %B = 35; and 70 min %B = 45. Peptides from the thermolysin digest were separated with a solvent system of A: 0.01 N HCl and B:40% (v/v) acetonitrile in water. The peptide digest was eluted with the following program: 0 min, %B = 5; 10 min, %B = 5; 20 min, %B = 25; 25 min, %B = 25; 35 min, %B = 62.5; 45 min, %B = 62.5.

Reduction and Alkylation of Spinach Calmodulin. Alkylation of the cysteine residue in spinach calmodulin was carried out under a variety of conditions. The most reproducible incorporation of ^{14}C iodoacetic acid (Amersham, 54 mCi/mmol) was obtained by incubation of the protein (90 μM) in a nitrogen-purged buffer of 100 mM Tris-HCl, 1.0 mM EDTA, 6 M guanidine hydrochloride, at pH 8.3. A four-equivalent excess of either DTT or β-mercaptoethanol was added and the solution incubated at 37°C for 1 h. Undiluted [^{14}C]iodoacetic acid was added to a concentration of 900 μM, and the reaction allowed to proceed for 3 h in the dark at room temperature. Incorporation of label was determined after dialysis against 10 mM ammonium bicarbonate and lyophilization. Generally 35 to 45% incorporation was obtained based on radioactivity. Amino acid analysis of the labeled protein gave 0.4 to 0.6 mol carboxymethyl cysteine/mol protein. [^{14}C]Carboxymethyl spinach calmodulin was then digested with trypsin as described above.

Peptide Nomenclature. Tc, R, CB, T, and Th refer to citraconyl modified protein digest, submaxillary protease digest, cyanogen bromide cleavage, trypsin, and thermolysin digest, respectively. RnTn refers to a trypsin subdigested peptide from the arginine protease digest. Peptides are numbered from their position in the amino acid sequence ascending to the COOH terminus.

RESULTS

When native spinach calmodulin was subjected to five cycles of automated Edman degradation, no sequential release of amino acids was observed indicating that the amino terminus is blocked as in other calmodulins. Since the composition of spinach calmodulin (Table I) suggested that there are five arginines in the molecule, a suitable number of fragments would be obtained from specific cleavage at arginine. Two procedures were used to generate specific cleavage at arginine residues: digestion of citraconylated calmodulin with trypsin, and treatment of calmodulin with a protease from mouse submaxillary glands. Citraconylated spinach calmodulin was prepared, subjected to digestion with trypsin, and the resulting peptide mixture was fractionated by HPLC as described in "Materials and Methods." The amino acid compositions of purified peptides obtained by this procedure are given in Table II. In addition, the use of the relatively arginine-specific protease from mouse submaxillary glands provided a

Table I. *Compositions of Peptides Isolated from Digestion of Spinach Calmodulin with Mouse Submaxillary Protease*

Isolation of these peptides is described in "Materialsa and Methods." Sepration is shown in Figure 1. Numbers in parentheses indicate residues calculated from sequence analysis.

Amino Acid	R1	R2	R3	R4	R5	Total	Protein Composition[a]
Asp	7.1 (4)	8.1 (8)	2.2 (2)	3.5 (3)	7.2 (7)	24	27.0
Thr	2.8 (4)	1.8 (2)	1.0 (1)		1.9 (2)	9	8.1
Ser	2.6 (1)	0.9 (1)	1.8 (1)	1.3 (1)		4	5.2
Glu	6.0 (7)	6.0 (6)	4.7 (4)	2.5 (2)	8.0 (8)	27	26.8
Pro		1.9 (2)				2	2.1
Gly	3.5 (3)	3.0 (3)	1.4 (0)	1.7 (1)	3.0 (3)	10	10.2
Ala	2.2 (3)	3.0 (3)	1.3 (1)	1.2 (2)	2.0 (2)	11	11.0
Val	1.9 (1)	1.0 (1)		1.0 (1)	4.5 (5)	8	7.8
Met	0.7 (1)	1.6 (2)	0.6 (1)		3.0 (4)	8	7.9
Ile	2.0 (2)	1.8 (2)		1.0 (1)	1.8 (2)	7	7.1
Leu	3.0 (3)	3.7 (4)	1.2 (1)	1.1 (1)	1.9 (2)	11	11.3
Tyr					1.0 (1)	1	1.1
Phe	2.6 (3)	1.9 (2)	1.2 (1)	1.8 (2)	1.0 (1)	9	9.1
His					0.8 (1)	1	0.9
TmLys[b]					1.0 (1)	1	1.0
Lys	3.0 (3)		2.6 (3)	1.2 (1)	2.0 (2)	9	9.1
Arg	1.0 (1)	0.9 (1)	0.8 (1)	1.0 (1)	0.7 (1)	5	5.2
CmCys[c]	0.6 (1)					1	
Total residues	37	37	16	16	42	148	
% Recovery	50	48	47	71	65		

[a] Determined from a 24-h hydrolysis of unmodified protein. Values are uncorrected for losses or incomplete digestion.
[b] TmLys-trimethyllysine.
[c] CmCys-carboxymethylcysteine.

Table II. *Compositions of Peptides Isolated from Trypsin Digestion of Citraconylated Spinach Calmodulin*

Numbers in parentheses indicate residues inferred from sequence analysis. Peptide nomenclatures and isolation are described in "Materials and Methods."

Amino Acid	TC1	TC2	TC3	TC4	TC5	TC6
Asp	5.9 (4)	7.9 (8)	2.4 (2)	3.3 (3)	3.4 (3)	3.8 (4)
Thr	2.7 (4)	1.8 (2)	1.1 (1)	0.5 (0)	1.9 (2)	
Ser	2.1 (1)	1.4 (1)	1.5 (1)	1.2 (1)	0.4 (0)	
Glu	7.2 (7)	7.6 (6)	4.6 (4)	3.0 (2)	4.2 (4)	4.2 (4)
Pro		1.7 (2)				
Gly	3.2 (3)	3.4 (3)	0.8 (0)	1.4 (1)	1.2 (1)	2.3 (2)
Ala	2.4 (3)	3.1 (3)	1.5 (1)	2.0 (2)	0.5 (0)	2.0 (2)
Val	0.9 (1)	0.7 (1)		1.2 (1)	1.5 (2)	3.0 (3)
Met	1.2 (1)	1.8 (2)	1.1 (1)		2.1 (2)	1.7 (2)
Ile	1.5 (2)	1.5 (2)		0.9 (1)	0.7 (1)	1.1 (1)
Leu	3.0 (3)	3.8 (4)	1.1 (1)	1.1 (1)	2.0 (2)	
Tyr						1.0 (1)
Phe	2.5 (3)	1.9 (2)	0.7 (1)	1.7 (2)		1.0 (1)
His					0.8 (1)	
TmLys[a]					0.7 (1)	
Lys	2.3 (3)		2.8 (3)	1.0 (1)		2.0 (2)
Arg	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	0.8 (1)	
Cys	ND[b]					
Total residues	37	37	16	16	20	22

[a] Trimethyllysine.
[b] Not determined.

more limited number of fragments which can be separated by HPLC as in Figure 1. The amino acid compositions of the peptides obtained by this procedure are shown in Table I. The peptides obtained from this digest encompass the entire protein and were used as the primary targets of sequence analysis. Peptides from arginine-specific protease digests were subdigested when appropriate to obtain more sequence information on a particular fragment. The amino acid sequences were determined and aligned by homology to bovine brain calmodulin following the recommendations of Goldstone and Needleman (7). Frag-

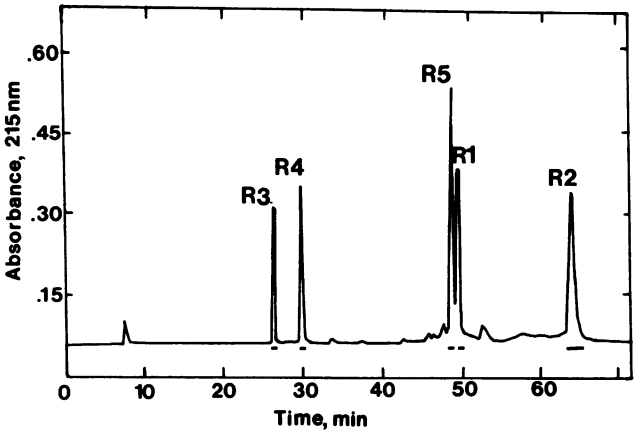

FIG. 1. Chromatogram of peptides separated from digestion of 30 nmol of spinach calmodulin with mouse submaxillary protease. Separation parameters are described in "Materials and Methods." Compositions of peptides are in Table II.

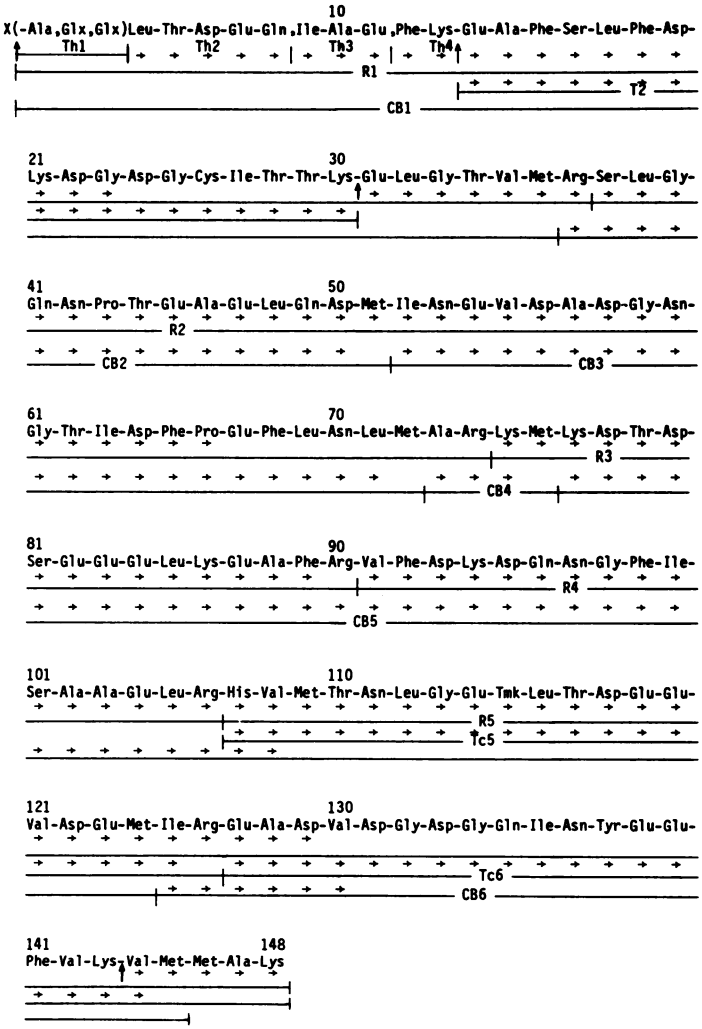

FIG. 2. Summary of the spinach calmodulin covalent structure. Major peptides begin where noted with the nomenclature described in "Materials and Methods." Solid lines indicate peptides identified by composition. Horizontal arrows indicate residues with peptides actually identified by automated Edman degradation. Vertical arrows point towards amino termini of trypsin peptides obtained from subdigests of arginine peptides R1 and R5.

ments were also generated from a cyanogen bromide digest of spinach calmodulin to provide overlaps between arginine cleavage peptides and confirm certain internal regions of arginine cleavage peptides.

A summary of the minimal number of peptides necessary for the elucidation of the protein structure is given in Figure 2. The results are summarized below in sections that correspond to peptides purified from the arginine-specific protease digest and start with the amino terminus of the protein.

Residues 1 to 37. Peptide R1 was resistant to Edman degradation so it was placed as the NH_2 terminus of the protein. Determination of the amino acid sequence of R1 was accomplished by means of trypsin and thermolysin digests. When subdigested with trypsin, peptide R1 yielded three peptides that were isolated by HPLC (Fig. 3) and designated R1T1, R1T2, and R1T3. R1T1 was tentatively placed as the NH_2 terminus of the protein since it was resistant to Edman degradation. A pool of the amino terminal peptide R1T1 was digested with thermolysin since the amino acid composition of this peptide (Table III) suggested that the thermolysin cleavage pattern should be similar to that found with bovine brain calmodulin (29). Four major peptides were isolated from this digest (Table IV). One peptide, Th4, gave the sequence Phe-Lys, which placed it at the COOH terminus of R1T1. Two other peptides, Th2, and Th3, provided the sequences: Leu-Thr-Asp-Glu-Gln and Ile-Ala-Glu, respectively. By homologous alignment to bovine brain calmodulin, the last nine residues of R1T1 are proposed to be Leu-Thr-Asp-Gln-Ile-Ala-Glu-Phe-Lys. Based upon the remaining residues from the R1T1 peptide, and the composition of the Th1 peptide, the remaining sequence probably contains two Glx residues, an alanine and possibly a serine. The serine content of the R1 and R1T1 peptides was variable among several preparations and extraneous serine, glycine, and some acidic residues were found in amino acid analysis of some peptides isolated from multiple digestion and chromatographic steps. The nature of the blocking group and overlap of peptides within R1T1 were not determined

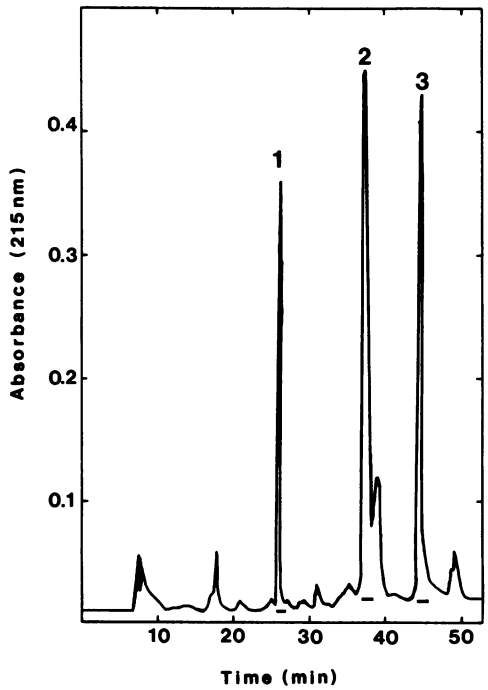

FIG. 3. Separation of peptides obtained from subdigestion of R1 (residues 1–37) with trypsin. Chromatographic parameters are described in "Materials and Methods." Peak 1 is R1T3 (residues 31–37), peak 2 is R1T2 (residues 14–30), and peak 3 is R1T1 (residues 1–13). Compositions of these peptides are given in Table III.

Table III. *Compositions of Peptides Isolated from Trypsin Digests and Subdigests*

Peptide T2 is from a trypsin digest of [^{14}C]carboxymethyl spinach calmodulin (Fig. 4); all other peptides were purified from a trypsin subdigestion of peptides isolated from digestion of spinach calmodulin with mouse submaxillary protease (Table I). Peptides were isolated as described in "Materials and Methods."

Amino Acid	R1T1	T2	R1T3	R5T1	R5T2	R5T3
Cm-Cys[a]		1.0 (1)				
Asp	1.4 (1)[b]	3.0 (3)		3.4 (3)	3.8 (4)	
Thr	1.0 (1)	1.8 (2)	0.8 (1)	2.0 (2)		
Ser		1.5 (1)		0.3 (0)	0.6 (0)	
Glu	5.0 (5)	1.1 (1)	1.2 (1)	4.4 (4)	3.9 (4)	
Pro						
Gly	0.4 (0)	2.8 (2)	1.2 (1)	1.5 (1)	2.2 (2)	0.9 (0)
Ala	2.1 (2)	1.1 (1)		0	1.7 (1)	1.1 (1)
Val			0.9 (1)	1.9 (2)	1.9 (2)	.9 (1)
Met			0.6 (1)	1.9 (2)		1.0 (2)
Ile	1.0 (1)	1.0 (1)		1.0 (1)	0.8 (1)	
Leu	1.2 (1)	1.2 (1)	0.9 (1)	2.1 (2)		
Tyr					0.7 (1)	
Phe	0.9 (1)	1.3 (2)			1.0 (1)	
His				0.7 (1)		
TmLys[c]				1.2 (1)		
Lys	1.0 (1)	1.8 (2)			1.3 (1)	1.0 (1)
Arg			1.0 (1)	1.0 (1)		
Total residues	13	17	7	20	17	5

[a] Carboxymethyl cysteine.

[b] Numbers in parentheses are the number of residues found on sequence analysis.

[c] Trimethyllysine.

Table IV. *Compositions of Peptides Obtained from Thermolysin Digest of Peptide R1T1*

Amino Acid	Th1[a]	Th2[b]	Th3[c]	Th4
Asp			1.0 (1)	
Thr			0.9 (1)	
Ser	0.76 (0–1)		0.4 (0)	
Glu	1.94 (2)	1.2 (1)	2.2 (2)	
Gly	0.60 (0)		0.3 (0)	
Ala	1.00 (1)	1.0 (1)	0.4 (0)	
Val				
Met				
Ile		0.8 (1)	0.3 (0)	
Leu			0.9 (1)	
Tyr				
Phe				0.8 (1)
His				
Lys				1.0 (1)
Arg				
Total residues	3–4	3	5	2

[a] Numbers in parentheses are residues assumed present.

[b] Numbers in parentheses are residues found by sequence analysis.

[c] Sequence analysis of this peptide demonstrated that it contained 25% of Th2. This accounts for the Ile and Ala in the composition. Numbers in parentheses indicate residues calculated from sequence analysis.

as part of this study.

Peptide R1T2 was degraded 10 cycles to give a partial sequence: Glu-Ala-Phe-Ser-Leu-Phe-Asp-Lys-Asp which is identical to residues 14 to 23 of bovine calmodulin. After alkylation of spinach calmodulin with [^{14}C]iodoacetic acid, one labeled peptide was isolated from a complete trypsin digest of the modified protein shown in Figure 4. The composition of this peptide (T2, Table III) showed the presence of carboxymethylcysteine, and the amino acid sequence was: Glu-Ala-Phe-Ser-Leu-Phe-Asp-Lys-Asp-Gly-Asp-Gly-Cys(CH_2-COOH)-Ile-Thr-Thr-Lys. Analysis of the radioactivity found in the Pth amino acid fractions (Fig. 4) demonstrated that most of the isotope is in cycle 13 which contained Pth-carboxymethyl cysteine. Spinach calmodulin, therefore, has a cysteine at residue 26.

Peptide R1T3 was the only arginine-containing peptide so it was tentatively placed at the COOH terminus of R1. Its composition (Table III) and amino acid sequence are identical to residues 31 to 37 of bovine brain calmodulin, Glu-Leu-Gly-Thr-Val-Met-Arg.

Residues 38 to 74. Peptide R2 was subjected to 37 cycles of Edman degradation and assignments up to the 30th cycle tentatively aligned with residues 38 to 67 of bovine calmodulin. Part of the remaining sequence was obtained from the cyanogen bromide digest peptide CB3 (see below). CB3 was covalently linked to a polymeric support and subjected to automated sequence analysis. Sequence assignments up to the terminal homoserine, residue 72, were obtained. The amino terminus of CB3, Ile-Asn-Glu-Val, overlapped peptide R2 beginning at cycle 15, residue 52 of the tentative sequence. The residues remaining from the composition of R2 are Ala-Arg. These were placed after the methionine residue based upon the specificity of cleavage of the mouse submaxillary enzyme. A unique cyanogen bromide peptide CB4, Ala-Arg-Lys-(Hse), confirmed this placement as discussed below. Sequence analysis of the citraconyl digest peptide Tc2 gave the same NH_2 terminal residues as R2 in the first 17 cycles of automated Edman degradation. Thus, Tc2 and R2 are equivalent. Two differences between bovine and spinach calmodulins were found in this region at residues 70 and 71 where Asn-Leu replaced Thr-Met.

Residues 75 to 90. Automated sequence analysis of R3 and Tc3 gave identical assignments up to the terminal arginine residue and by alignment with the bovine brain calmodulin sequence is apparently residues 75 to 90 with substitutions at residues 85 and 86, where Leu-Lys replaced Ile-Arg.

Residues 91 to 106. Peptide R4 was subjected to 16 cycles of automated Edman degradation. Unambiguous sequence assignments were made in 16 cycles to align this peptide with residues 91 to 106 of bovine calmodulin with substitutions of Gln/Gly at 96 and Phe/Tyr at 99. These substitutions were confirmed in sequence analysis of Tc4.

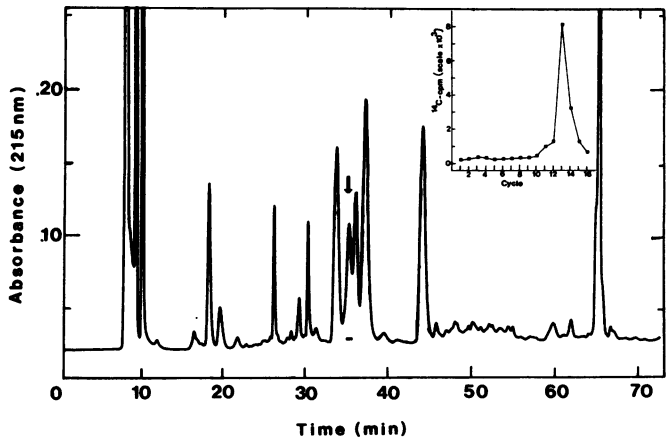

FIG. 4. Separation of peptides from a complete trypsin digest of [^{14}C] carboxymethyl spinach calmodulin (20 nmol). The arrow indicates the labeled peptide T2. Inset gives the radioactivity pattern from automated Edman degradation of T2.

Residues 107 to 148. Peptide R5 was degraded 24 cycles, the first 20 of which were identical to peptide Tc5. The last four residues were the same as the NH_2 terminus of Tc6. When automated sequence analysis was performed on peptide Tc5, its amino acid sequence was identical to bovine calmodulin residues 107 to 126. No Pth-trimethyllysine derivative was quantitatively identified by HPLC in cycle 9 of the automated sequencer run, but acid hydrolysis of the Pth-amino acid product yielded trimethyllysine by amino acid analysis. Thus, Tc5 and R5 contain the single trimethyllysine residue found in the native protein. Peptide Tc6 contained no arginine and was therefore placed at the COOH terminus of spinach calmodulin. Tc6 was subjected to automated Edman degradation and sequence assignments made in the initial 18 cycles. The first 16 residues were aligned with residues 127 to 142 of the bovine brain sequence with substitutions of Val for Ile at residue 130 and Ile for Val at 136. The last two cycles identified from Tc6 were Lys-Val. Thus, at this point, the amino acid sequence of residues 127 to 144 was Glu-Ala-Asp-Val-Asp-Gly-Asp-Gly-Gln-Ile-Asn-Tyr-Glu-Glu-Phe-Val-Lys-Val leaving two Met, one Ala, and one Lys to be assigned. Three peptides were then isolated from a trypsin subdigest of R5 (Fig. 5; Table III) and designated R5T1, R5T2, and R5T3. Peptide R5T1 was identical in composition to Tc5, and R5T2 was identical to the NH_2-terminal 17 residues of Tc6. Upon automated sequence analysis, peptide R5T3 gave: Val-Met-Met-Ala-Lys. Thus, the COOH terminus of spinach calmodulin ends with this peptide since the three subdigest peptides R5T1, R5T2, and R5T3 are contiguous in amino acid sequence and encompass the composition of R5. Furthermore, Lys-Val, assigned to residues 143 to 144 is the only Lys-Val peptide in spinach calmodulin.

Cyanogen Bromide Cleavage Peptides. The tentative order and amino acid sequence of the arginine cleavage peptides were confirmed by isolation of peptides from a cyanogen bromide digest (Fig. 6) of spinach calmodulin. These peptides were identified by amino acid composition (Table V) and amino acid

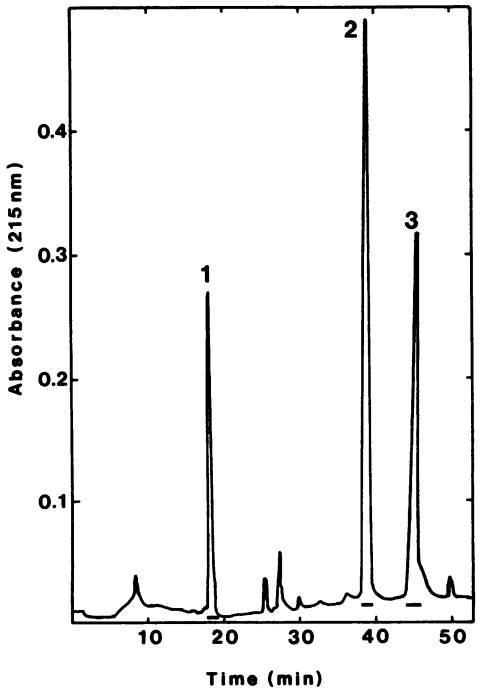

FIG. 5. Chromatogram of peptides obtained from trypsin digestion of peptide R5. Separation parameters are described in "Materials and Methods." Peak 1 is R5T3 (residues 144–148), peak 2 is R5T2 (residues 127–143), and peak 3 is R5T1 (residues 107–126). Compositions are in Table III.

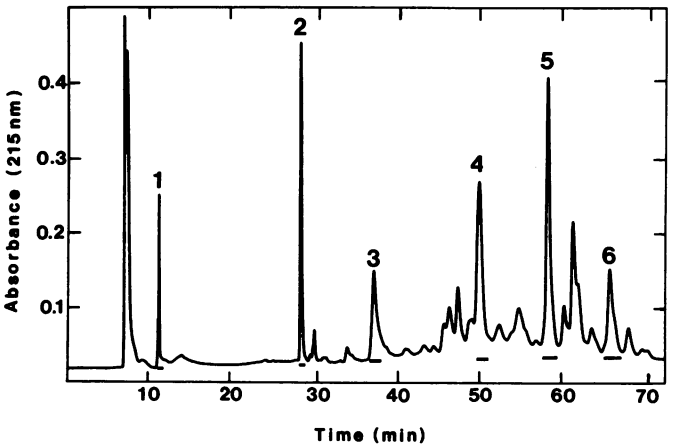

FIG. 6. Separation of peptides from a cyanogen bromide digest of spinach calmodulin (36 nmol). Chromatographic parameters are given in "Materials and Methods." Peak 1 is peptide CB4 (residues 73–76), peak 2 is CB2 (residues 37–51), peak 3 is CB6 (residues 125–145), peak 4 is CB5 (residues 77–124), peak 5 is CB3 (residues 52–72), peak 6 is CB1 (residues 1–36).

Table V. *Compositions of Peptides from Cyanogen Bromide Digests of Spinach Calmodulin*

Amino Acid	CB1[a]	CB2[b]	CB3[b]	CB4[b]	CB5[c]	CB6[c]
Asp	5.3 (4)	2.1 (2)	5.7 (6)		8.2 (8)	3.7 (4)
Thr	3.3 (4)	1.0 (1)	1.0 (1)		2.6 (3)	
Ser	1.9 (1)	1.0 (1)	0.6 (0)	0.2 (0)	1.7 (2)	0.4 (0)
Glu	7.1 (7)	4.0 (4)	2.6 (2)		9.4 (10)	3.5 (4)
Gly	3.5 (3)	1.5 (1)	2.3 (2)	0.2 (0)	2.4 (2)	2.1 (2)
Ala	3.0 (3)	1.1 (1)	1.3 (1)	1.0 (1)	3.2 (3)	1.0 (1)
Val	1.2 (1)		1.0 (1)		3.0 (3)	2.4 (3)
Ile	1.9 (2)		1.6 (2)		1.2 (1)	1.5 (2)
Leu	3.1 (3)	1.8 (2)	1.9 (2)		3.5 (4)	
Tyr						0.8 (1)
Phe	2.9 (3)		1.8 (2)		2.6 (3)	0.9 (1)
His					0.9 (1)	
TmLys[d]					0.9 (1)	
Lys	3.0 (3)			0.9 (1)	2.6 (3)	0.6 (1)
Arg		0.8 (1)		0.8 (1)	1.8 (2)	0.6 (1)
Hse[e]	+ (1)	+ (1)	+ (1)	+ (1)	+ (2)	+ (1)
Pro		1.3 (1)	0.8 (1)			
Cys	ND[f] (1)					
Total residues	36	15	21	4	48	21

[a] Numbers in parentheses are based on the sequence of previously determined arginine protease generated fragments.

[b] Numbers in parentheses are residues calculated from sequence analysis of the whole peptide.

[c] Residues based upon amino terminal sequence analysis and previously determined amino acid sequence of arginine protease generated fragments.

[d] Trimethyllysine.

[e] Homoserine was qualitatively observed in amino acid analysis. A '+' indicates a positive identification.

[f] Not determined.

sequence. The composition of peptide CB1 was consistent with residues 1 to 36. Peptide CB2 had the NH_2 terminal sequence: Arg-Ser-Leu-Gly- which overlaps R1 and R2 since only one Met-Arg sequence was found in the protein. As discussed above, the amino acid sequence of CB3 placed it within peptide R2 (residues 52–72). CB4 had the sequence: Ala-Arg-Lys-(Hse) and provides

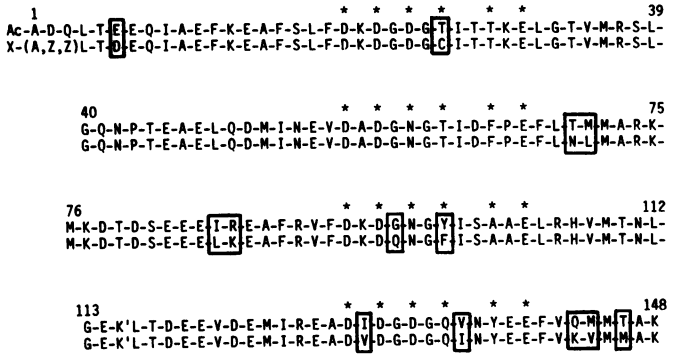

FIG. 7. Sequence comparison of bovine brain and spinach calmodulin. Asterisks indicate residues which are potential Ca binding ligands (14). Sequence differences are enclosed in boxes. Standard IUPAC single letter coding is used for amino acids.

the last two residues of R2 and the first two of R3. No other Lys-Met peptide was found in the protein. The composition of CB5 was consistent with that expected from prior studies of cyanogen bromide cleavages of calmodulin (12, 29) and includes most of peptides R3 and R5T1 as well as all of peptide R4 and the NH_2 terminus of R5 (Fig. 2). Sequence analysis of the first 32 residues of CB5 confirmed the placement of R3, R4, and R5. CB6 gave the NH_2 terminal sequence Ile-Arg-Glu-Ala-Asp to confirm its placement within R5.

As is evident from Figure 2 and Table I, the amino acid composition of spinach calmodulin is consistent with the sum of the amino acid sequences of arginine peptides determined in this study. An alignment of the resultant amino acid sequence of spinach calmodulin with the complete amino acid sequence of bovine brain calmodulin is given in Figure 7.

DISCUSSION

A structure of spinach calmodulin was established using homologous alignments of arginine cleavage peptides to bovine brain calmodulin. Peptides from a cyanogen bromide digest confirmed the order of these peptides and provided key sequence assignments. While these data are not a complete proof of structure, sufficient documentation is provided to justify the structure presented.

An alignment of spinach and bovine brain calmodulin revealed that exclusive of the amino terminal tripeptide, there are 13 differences in the amino acid sequence. There may be additional differences in the amino terminal tripeptide. Current studies using mass spectrometry of blocked peptides, and DNA sequence analysis of calmodulin genes should provide the amino terminal amino acid sequence and the blocking group. Figure 7 gives a comparison of bovine and spinach calmodulins. The most extensive sequence identity between bovine and spinach calmodulins occurs in three extended regions: residues 7 to 25, 27 to 69, and 100 to 116. These same regions appear to have similar amino acid sequences in all other calmodulins examined to date (3, 13). In overall comparison to other calmodulins, however, more divergence of the calmodulin covalent structure is evident in spinach calmodulin. Some of the changes seen in spinach calmodulin may be unique to higher plant calmodulins. Not seen in calmodulins other than higher plant are Asp at residue 6, Cys at 26, Asn at 70, Lys at 86, Gln at 96, Val at 144, and a Met at residue 146. Previous structural studies of barley calmodulin demonstrated that it also has Gln 96 (22).

The Cys/Thr substitution at residue 26 and the Gln/Gly at 96 require a minimum of two base changes in the nucleotide codons. Both of these differences are in the putative Ca binding loops (14). The positions of the Ca ligands are illustrated in Figure 7. A 12-residue segment containing six coordination ligands is flanked by helical regions. A variation in loop flexibility might result from the Gln substitution since this residue is a Gly in other calmodulins. In a search of other related Ca binding proteins, a Gln residue occupied this position in homologous regions of only two proteins (frog parvalbumin-alpha and parvalbumin from ray) (14). The crystal structure of a parvalbumin from carp has been determined (14), thus the relative positions of residues within the Ca binding loops can be extrapolated to other related Ca binding proteins such as calmodulin and troponin C. The presence of Cys-26 at the fourth ligand site (see Fig. 7) is apparently a rare occurrence. Only troponin C from cardiac muscle and slow skeletal muscle have Cys in the same relative position. These two proteins, however, have other replacements in this Ca binding loop which are consistent with a loss in Ca binding capacity (14). Based upon the parvalbumin structure, the fourth ligand site binds Ca^{2+} through the carbonyl oxygen of the peptide bond rather than the amino acid side chain. Therefore, the effect of Cys-26 and Gln-96 on Ca binding cannot be evaluated until the quantitative Ca binding properties of spinach calmodulin have been determined and related to a three-dimensional structure. While the specific function of cysteine-26 in plant calmodulins is unknown, its presence could be exploited for the specific introduction of affinity reagents and chromogenic probes into the first structural domain of calmodulin, and provides a cleavage site for the generation of a modified calmodulin from which 25 residues can be removed.

Besides Cys-26 and Gln-96 five other amino acid substitutions are unique in spinach calmodulin. These differences are not in Ca binding loops, but are in the flanking helical regions. Of these, Asn-70 and Met-146 are nonconservative with respect to amino acid side chain since they replace threonines in bovine calmodulin. Asp-6, Lys-86, and Val-144 are more conservative changes of charged or hydrophobic side chains with respect to equivalent positions of bovine calmodulin. Amino acid sequence differences exhibited by spinach calmodulin at residues 71, 85, 99, 130, 136, and 143 have been observed in other calmodulins. For example, *Tetrahymena* calmodulin also has Leu at residues 71 and 85, and Ile at 136, while *Metridium* and *Renilla* calmodulins contain Phe at residue 99 and Lys at 143. Amino acid sequence data obtained from a fragment of calmodulin from the green alga *Chlamydomonas reinhardtii* indicates that it contains a Val at residue 130 (23). Spinach calmodulin, therefore, shares six of its amino acid differences, with respect to bovine calmodulin, among other diverse calmodulins. The composite structural consequences of these shared substitutions and the seven unique replacements discussed above may be reflected in the functional activity of spinach calmodulin.

Spinach calmodulin and calmodulins from divergent sources stimulate the activity of a brain 3′,5′ cyclic-AMP phosphodiesterase (3, 13). This particular function of calmodulin is apparently independent of the structural differences discussed above. The evolution of calmodulin suggested by Goodman (8) illustrates that it underwent rapid change in its early stage and then decelerated to a very slow rate of less than one nucleotide replacement per 100 million years. Assuming that plant and other calmodulins evolved from the same ancestral precursor, the primary functional characteristics of Ca binding and protein-protein interaction were established early and then conserved. A corollary to this hypothesis is that divergence from this ancestral precursor was perhaps the result of a more specialized function or optimization dictated by the host environment. Thus, a more detailed study of the ability of different calmodulins to activate enzymes other than phosphodiesterase is necessary to address this corollary. Three Ca^{2+}-calmodulin-dependent plant enzymes, NAD kinase (18), microsomal ATPase (5), and a specific protein kinase (20) have been reported. These and other calmodulin-

dependent enzymes have not been studied in detail with respect to activation by various calmodulins.

Presented in the accompanying paper are comparisons of three calmodulins in their ability to activate two sensitive enzymes, NAD-kinase and myosin light chain kinase. The differences in activation of these enzymes add functional significance to our elucidation of the covalent structure of spinach calmodulin as a representative higher plant calmodulin. These studies also demonstrate the necessity to isolate additional calmodulin-dependent enzymes, especially from higher plants and algae, and examine their interactions with various calmodulins of known covalent structure. In this regard, the covalent structure of an algal calmodulin is now under investigation. The determination of its amino acid sequence should allow correlation of a unique chemical structure with a specific biochemical activity profile. These analyses will, in turn, provide a firm foundation for physiological studies and may provide insight into the role of calmodulin in plant cell function.

LITERATURE CITED

1. ANDERSON JM, H CHARBONNEAU, HP JONES, RD MCCANN, MJ CORMIER 1980 Characterization of the plant nicotinamide adenine dinucleotide kinase activator protein and its identification as calmodulin. Biochemistry 19: 3113–3120
2. ATASSI MZ, AFSA HABEEB 1972 Reaction of proteins with citraconic anhydride. Methods Enzymol 25: 546–553
3. BURGESS WH, M SCHLEICHER, LJ VAN ELDIK, DM WATTERSON 1983 Comparative studies of calmodulin. *In* WY Cheung, ed, Calcium and Cell Function, Vol 4. Academic Press, NY, pp 209–261
4. DEDMAN JR, RL JACKSON, WE SCHREIBER, AR MEANS 1978 Sequence homology of the Ca^{2+} dependent regulator of cyclic nucleotide phosphodiesterase from rat testis with other Ca^{2+} binding proteins. J Biol Chem 253: 343–346
5. DIETER P, D MARME 1981 A calmodulin-dependent, microsomal ATPase from corn (*Zea mays* L). FEBS Lett 125: 245–248
6. GRAND RJA, S SHENOLIKAR, P COHEN 1981 The amino acid sequence of the δ subunit of rabbit skeletal muscle phosphorylation kinase. Eur J Biochem 113: 359–367
7. GOLDSTONE AD, SB NEEDLEMAN 1975 Reconstruction of the primary sequence of a protein from peptides of known sequence. *In* SB Needleman, ed, Protein Sequence Determination. Springer-Verlag, NY, pp 298–321
8. GOODMAN M 1980 Molecular evolution of the calmodulin family. *In* FL Siegel, E Carafoli, RH Kretsinger, DH MacLennan, RH Wasserman, eds, Calcium-Binding Proteins: Structure and Function. Elsevier, NY, pp 347–354
9. HEWICK RM, WM HUNKAPILLER, LE HOOD, WJ DREYER 1981 A gas-liquid solid phase peptide and protein sequenator. J Biol Chem 256: 7990–7997
10. JAMIESON GA, DD BRONSON, FH SCHACHAT, TC VANAMAN 1980 Structure and function relationships among calmodulins and troponin C-like proteins from divergent eucaryotic organisms. Ann NY Acad Sci 356: 1–13
11. JARRETT HW, T DASILVA, MJ CORMIER 1982 Calmodulin activation of NAD kinase and its role in the metabolic regulation of plants. *In* DA Robb, WS Pierpoint, eds, The Uptake and Utilization of Metals in Plants. Academic Press, London, pp 205–218
12. KASAI H, Y KATO, T ISOBE, H KAWASAKI, T OKUYAMA 1980 Determination of the complete amino acid sequence of calmodulin (phenylalanine-rich acidic protein II) from bovine brain. Biomed Res 1: 248–264
13. KLEE CB, TC VANAMAN 1982 Calmodulin. Adv Protein Chem 35: 213–321
14. KRETSINGER RH 1980 Structure and evolution of calcium modulated proteins. CRC Crit Rev Biochem 8: 119–174
15. LAGACE L, T CHANDRA, SLC WOO, AR MEANS 1983 Identification of multiple species of calmodulin messenger RNA using a full length complementary DNA. J Biol Chem 258: 1684–1688
16. LAURSEN RA, MJ HORN 1977 Coupling methods and strategies in solid phase sequencing. *In* SB Needleman, ed, Molecular Biology, Biochemistry and Biophysics, Vol 25, Springer-Verlag, NY, pp 21–38
17. MARSHAK DR, DM WATTERSON, LJ VAN ELDIK 1981 Calcium-dependent interaction of S100B, troponin C, and calmodulin with an immobilized phenothiazine. Proc Natl Acad Sci USA 78: 6793–6797
18. MUTO S, S MIYACHI 1977 Properties of a protein activator of NAD kinase from plants. Plant Physiol 59: 55–60
19. PUTKEY JA, KF TSUI, T TANAKA, L LAGACE, JP STEIN, EC LAI, AR MEANS 1983 Chicken calmodulin genes. J Biol Chem 258: 11864–11870
20. RENJEVA R, G REFENO, AM BOCEDET, D MARME 1983 Activation of plant quinate: NAD^+ 3-oxidoreductase by CA^{2+} and calmodulin. Proc Natl Acad Sci USA 80: 5222–5224
21. SASAGAWA T, LH ERICSSON, KA WALSH, WE SCHREIBER, EH FISHER, K TITANI 1982 Complete amino acid sequence of human brain calmodulin. Biochemistry 21: 2565–2569
22. SCHLEICHER M, TJ LUKAS, DM WATTERSON 1983 Further characterization of calmodulin from barley (*Hordeum vulgare*). Plant Physiol 73: 666–670
23. SCHLEICHER M, TJ LUKAS, DM WATTERSON 1984 Isolation and characterization of calmodulin from the motile green alga *Chlamydomonas reinhardtii*. Arch Biochem Biophys 229: 33–42
24. TAKAGI T, T NEMOTO, KK KONISHI 1980 The amino acid sequence of the calmodulin obtained from sea anemone (*Metridium senile*) muscle. Biochem Biophys Res Commun 96: 377–381
25. TODA H, M YAZAWA, K KONDO, T HOMURA, K NARITA, K YAGI 1981 Amino acid sequence of calmodulin from sea scallop. J Biochem (Tokyo) 90: 1493–1505
26. VAN ELDIK LJ, AR GROSSMAN, DB IVERSON, DM WATTERSON 1980 Isolation and characterization of calmodulin from spinach leaves and *in vitro* translation mixtures. Proc Natl Acad Sci USA 77: 4779–4783
27. WATTERSON DM, WH BURGESS, TJ LUKAS, D IVERSON, DR MARSHAK, M SCHLEICHER, BW ERICKSON, K-F FOK, LJ VAN ELDIK 1984 Towards a molecular and atomic anatomy of calmodulin and calmodulin binding proteins. Adv Cycl Nucl Prot Phos Res 16: 205–226
28. WATTERSON DM, DB IVERSON, LJ VAN ELDIK 1980 Spinach calmodulin: Isolation, characterization and comparison with vertebrate calmodulins. Biochemistry 19: 5762–5768
29. WATTERSON DM, FS SHARIEF, TC VANAMAN 1980 The complete amino acid sequence of the Ca^{2+}-dependent modulator protein (calmodulin) of bovine brain. J Biol Chem 255: 962–975
30. YAZAWA M, K YAGI, H TODA, K KONDO, K NARITA, R YAMAZAKI, K SOBUE, S KAKIUCHI, S NAGAO, Y NOZAWA 1981 The amino acid sequence of the *Tetrahymena* calmodulin that specifically interacts with guanylate cyclase. Biochem Biophys Res Commun 99: 1051–1057
31. ZIMMERMAN CL, E APPELLA, JJ PISANO 1977 Rapid analysis of amino acid phenylthiohydantoins by high performance liquid chromatograhy. Anal Biochem 77: 569–573