Smillie, L. B., & Hartley, B. S. (1966) *Biochem. J. 101*, 232.
Tarr, G. E., Beecher, J. F., Bell, M., & McKean, D. J. (1978) *Anal. Biochem. 84*, 622.
Titani, K., Cohen, P., Walsh, K. A., & Neurath, H. (1975) *FEBS Lett. 55*, 120.
Titani, K., Koide, A., Hermann, J., Ericsson, L. H., Kumar, S., Wade, R. D., Walsh, K. A., Neurath, H., & Fischer, E. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 4762.
Titani, K., Koide, A., Ericsson, L. H., Kumar, S., Hermann, J., Wade, R. D., Walsh, K. A., Neurath, H., & Fischer, E. H. (1978) *Biochemistry 17* (third of three papers in a series in this issue).
Walsh, K. A., Ericsson, L. H., & Titani, K. (1978) *Int. Symp. Proteins, Taipei* (in press).
Weber, K., & Osborn, M. (1969) *J. Biol. Chem. 244*, 4406.
Wikler, M., Titani, K., & Putnam, F. W. (1970) *J. Biol. Chem. 245*, 2158.
Yaoi, Y., Titani, K., & Narita, K. (1964) *J. Biochem. (Tokyo) 56*, 222.
Zarkadas, C. G., Smillie, L. B., & Madsen, N. B. (1968) *J. Mol. Biol. 38*, 245.

Amino Acid Sequence of Two Cyanogen Bromide Fragments of Glycogen Phosphorylase†

Jacques Hermann,‡ Koiti Titani,§ Lowell H. Ericsson, Roger D. Wade, Hans Neurath, and Kenneth A. Walsh*

ABSTRACT: This communication presents the strategy and experimental details to prove the amino acid sequence of two large fragments of rabbit muscle glycogen phosphorylase generated by cleavage with cyanogen bromide. These fragments, CB18 and CB15, represent 241 of the 841 residues in the whole molecule. In addition to applying methods of automated liquid phase Edman degradation, techniques of selective immobilization and solid phase Edman degradation are used. One of the two cyanogen bromide fragments (CB15) contains two of the sites of cleavage with hydroxylamine which have proved to be important in the overall strategy of determining the complete sequence of this molecule. Together with the accompanying reports by Koide, A., et al., and Titani, K., et al. ((1978) *Biochemistry 17* (first and third papers, respectively, in a series in this issue)), the present communication completes the proof of the amino acid sequence of phosphorylase and provides the basis for examining the relationship between its structure and function.

The amino acid sequence of the 841 residues of rabbit muscle glycogen phosphorylase has been presented in a preliminary report (Titani et al., 1977). The basic strategy included the generation and sequence determination of large fragments obtained by enzymatic or chemical cleavage; i.e., limited proteolysis by subtilisin which generated two large segments, L_S and H_S (30 000 and 70 000 daltons, respectively), cleavage by cyanogen bromide of 21 methionyl bonds, cleavage of 4 Asn-Gly bonds with hydroxylamine, and acidic cleavage of 3 Asp-Pro bonds. The present communication describes the proof of the amino acid sequence of the two largest fragments generated from segment H_S by cyanogen bromide. One, CB15, contains 163 residues including two of the Asn-Gly loci; the other, CB18, contains 78 residues. The analysis is based largely on specific fragmentation by chemical and enzymatic means and automated Edman degradation of the resulting peptides. Together with the accompanying reports by Koide et al. (1978) and Titani et al. (1978), the present work completes the proof of structure of phosphorylase. In conjunction with recent X-ray crystallographic data (Johnson et al., 1974; Sygush et al., 1977), this sequence analysis promises to provide the basis for a detailed understanding of the relationship between the structure and function of this enzyme which controls the first step of the degradation of glycogen.

Materials and Methods

Reagents for solid phase sequencing (Sequenal grade) were purchased from Pierce Chemical Co. Anhydrous reagent grade methanol was from Mallinckrodt and dichloroethane from Burdick and Jackson. *tert*-Butyl-*S*-4,6-dimethylpyrimid-2-yl thiocarbonate (*S*-Boc reagent[1]) was obtained from Beckman Co. Staphylococcal protease was a gift from Dr. G. Drapeau (Houmard & Drapeau, 1972). The same enzyme from Miles Laboratories was also used. Carboxypeptidase Y was a gift from Dr. M. Ottesen (Carlsberg Laboratory, Copenhagen, Denmark).

Liquid phase sequencing was performed as described in the preceding paper (Koide et al., 1978). High performance liquid chromatography was used to confirm identifications of the Pth-amino acids when the need arose.

High Performance Liquid Chromatography. Dry Pth-amino acids were redissolved in 25 μL of methanol. Suitable aliquots were injected into a Waters C-18 μ-Bondapack col-

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received July 6, 1978.* This work has been supported by research grants from the National Institutes of Health (GM15731 and AM7902).

‡ On leave from INSERM, France, recipient of NATO Scholarship. Present address: Institut Pasteur, 28 Rue du Docteur Roux, Paris 75015, France.

§ Investigator of the Howard Hughes Medical Institute.

[1] Abbreviations used: MITC, methyl isothiocyanate; PITC, phenyl isothiocyanate; *S*-Boc, *tert*-butyl-*S*-4,6-dimethylpyrimid-2-yl thiocarbonate; Pth, phenylthiohydantoin; TETA, triethylenetetramine; BNPS-skatole, bromine adduct of 2-(2-nitrophenylsulfenyl)-3-methylindole; SCM-Cys, *S*-carboxymethylcysteine; Hse, homoserine.

0006-2960/78/0417-5672$01.00/0

TABLE I: Amino Acid Composition of CB15 and CB18.[a]

	CB15	CB18
Asp	17.2 (17)	8.5 (8)
Thr	4.6 (5)	4.6 (5)
Ser	4.5 (5)	0.5 (0)
Glu	17.7 (18)	9.2 (8)
Pro	5.8 (6)	4.4 (4)
Gly	5.0 (5)	1.1 (1)
Ala	10.0 (10)	6.0 (6)
Cys[b]	2.0 (3)	0.6 (1)
Val	12.3 (13)	7.4 (8)
Met[c]	0.9 (1)	0.5 (1)
Ile	13.1 (14)	2.0 (3)
Leu	16.3 (17)	11.0 (12)
Tyr	6.7 (7)	1.6 (2)
Phe	7.0 (7)	2.0 (2)
His	5.8 (7)	2.6 (3)
Lys	15.9 (16)	2.2 (2)
Arg	10.5 (11)	8.3 (9)
Trp[d]	0.6 (1)	2.9 (3)
total no. of amino acid	163	78
mol wt	19 129	9322

[a] The numbers in parentheses indicate those found by sequence analysis. [b] Determined as SCM-Cys. [c] Determined as Hse. [d] Determined after hydrolysis with mercaptoethanesulfonic acid (Penke et al., 1974).

umn and separated by a 26-min linear gradient of methanol (14–55%) in an aqueous buffer (0.01 M sodium acetate, pH 4.13) by an adaptation of the system described by Bridgen et al. (1976). The flow rate was controlled at 2.2 mL/min by a Waters Associates system with a Model 660 solvent programmer.

Of the 18 amino acids in the organic phase, Pth-methionine and Pth-valine elute together as do Pth-phenylalanine and Pth-isoleucine. A serine derivative elutes in low yield with tryptophan, presumably because Pth-serine is converted to Pth-dehydroalanine. Artifacts of threonine appear in low yield as two peaks—one just before Pth-isoleucine and the other before Pth-leucine. The remainder appear in unique positions on the chromatogram. Pth-ε-succinyllysine elutes between Pth-alanine and Pth-tyrosine. Pth-arginine and Pth-histidine were separated by isocratic elution with 27% methanol.

Solid Phase Sequencing. Specific attachments of carboxyl-terminal homoserine peptides (100–300 nmol of peptide) to triethylenetetramine resin (TETA resin) were performed according to Horn & Laursen (1973). Carboxyl-terminal coupling of peptides to the TETA resin has occasionally been used. Before coupling, the amino group was blocked with *tert*-butyl-*S*-4,6-dimethylpyrimid-2-yl thiocarbonate (*S*-Boc reagent) (Nagasawa et al., 1973). The peptides were dissolved in 300 μL of water containing 10% triethylamine. *S*-Boc reagent (1.25 mg) in 200 μL of dimethylformamide was then added. The mixture was kept at room temperature for 2 h. Excess reagent was removed by ethyl acetate extraction and the aqueous phase was lyophilized. Coupling was accomplished with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (Schellenberger et al., 1972; Previero et al., 1973).

The Sequemat Model 12 was employed for solid phase Edman degradation according to the method of Laursen (1971). The thiazolinones of amino acids were converted to Pth-amino acids and divided into two equal aliquots. One of these was identified by gas chromatography after silylation, the other provided confirmation by high performance liquid chromatography.

Preparation of Fragments CB18 and CB15. Reduced and carboxymethylated phosphorylase *b* was prepared from rabbit muscle as described by Koide et al. (1978). The product was then treated with cyanogen bromide and fragments CB18 and CB15 were isolated on a Sephadex G-50 SF column (160 × 5 cm) in 9% formic acid using the procedures of Saari & Fischer (1973). Five pools were recovered [A through E in Figure 1 of Saari & Fischer (1973)]. By rechromatography of pool B on the same column, fragment CB15 was obtained in pure form with an average yield of 80%. Pool C was fractionated on a SP-Sephadex C-25 column (60 × 4 cm) equilibrated with 7 M urea in 0.1 M sodium formate at pH 2.9. Fragment CB18 was eluted (40% yield) with a linear gradient from the equilibration buffer (800 mL) to 7 M urea in 0.6 M sodium formate at pH 2.9 (800 mL).

Other materials and methods which are not mentioned in this section are described in a companion paper (Koide et al., 1978).

Results

The purities of fragments CB18 and CB15 were established by sodium dodecyl sulfate–polyacrylamide gel electrophoresis, amino acid analysis, and N-terminal sequence determination. The fragments migrated as single bands, and the amino acid compositions are presented in Table I. Sequenator analysis indicated a single sequence in each case beginning with Arg-Val-Leu- and with Ala-His-Leu- for CB18 and CB15, respectively.

The amino acid sequences of the two fragments were determined largely by Edman degradation using the strategy of fragmentation illustrated in Figure 1. The residue numbers used in the description of each analysis refer to the linear sequence of that fragment. However, in the complete sequence of the phosphorylase molecule, CB18 corresponds to residues 350–427 and CB15 to residues 441–603 (Titani et al., 1977).

Amino Acid Sequence of CB18. This fragment contained 78 residues including 3 tryptophan, 2 lysine, 9 arginine, and 1 homoserine residue (Table I).

Amino-terminal sequence analysis of the fragment identified 34 residues with a 92.3% stepwise degradation yield beginning with the sequence Arg-Val-Leu- and ending with Glu-Ala-Leu[34] (Figure 1).

The carboxyl-terminal sequence was determined by solid-phase sequencing. Fragment CB18 (220 nmol) was dissolved in 3 mL of water and digested for 2 h in a pH-stat with 0.04 mg of thermolysin at pH 8 (37 °C). A homoserine-containing peptide (Th-1) was selectively attached to the TETA resin and the amino sequence was established by the solid phase method as Leu-Arg-Arg-Hse. In order to identify the amino-terminal leucine as Pth-Leu (10% yield), the remaining amino groups were blocked with PITC instead of the MITC which is usually employed in solid phase sequencing. In a similar experiment, 200 nmol of CB18 was dissolved in 1 mL of 5% formic acid and digested with 0.05 mg of pepsin for 24 h. Two homoserine-containing peptides were selectively attached to the TETA resin. The sequence of a major peptide (P-1, 42% attachment yield) was Arg-Arg-Hse which corresponded to that within Th-1. However, a minor peptide (P-2) was also observed during the same analysis. It could be distinguished from P-1 by the lower attachment yield (24%) and greater length, i.e., Gly-Asp-Val-Asp-Arg-Leu-Arg-Arg-Hse (stepwise degradation yield: 90%).

Cleavage at Tryptophan. A sample of CB18 (4.8 μmol) was dissolved in 5 mL of 80% acetic acid and treated with 140 mg (385 μmol) of BNPS-skatole for 6 h at room temperature.

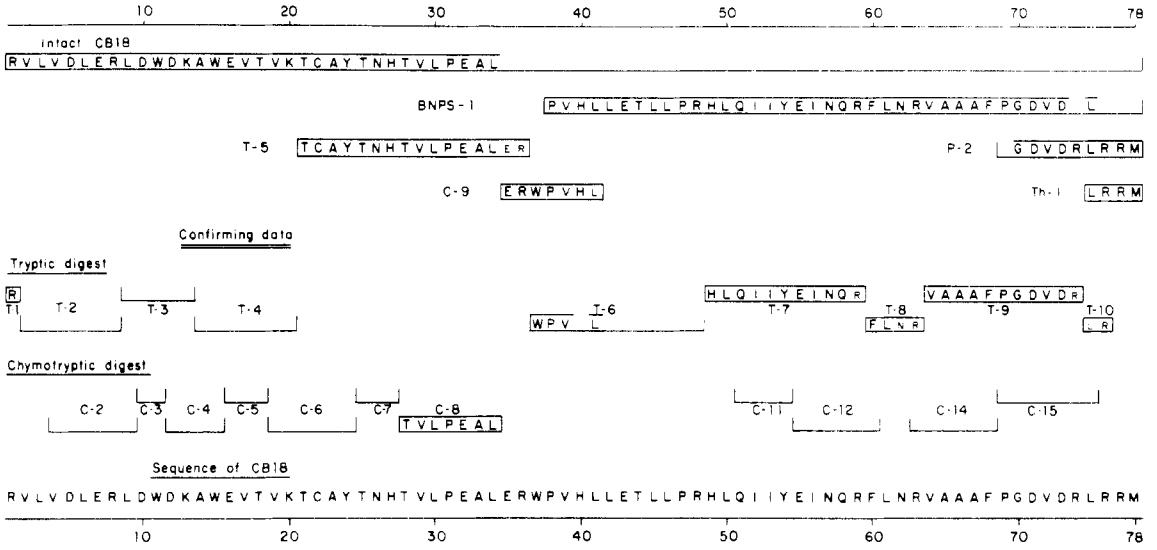

FIGURE 1: Summary proof of the amino acid sequence of CB18. The one-letter code within the bars designates amino acid residues in the peptide. Large capital letters indicate amino acids identified after Edman degradation or by carboxypeptidase digestion. Small capital letters indicate placement by composition and cleavage specificity. The length of each bar indicates the length of the peptide analyzed; enclosure of the top of the bar indicates the proven portion of the sequence; gaps in the upper enclosure indicate portions of sequence not identified. Methionine was identified as homoserine. The upper portion of the diagram summarizes the basic proof of structure. The lower portion illustrates confirmatory data and summarizes the structure. One letter amino acid abbreviations used are: A (alanine), C (cysteine), D (aspartic acid), E (glutamic acid), F (phenylalanine), G (glycine), H (histidine), I (isoleucine), K (lysine), L (leucine), M (methionine), N (asparagine), P (proline), Q (glutamine), R (arginine), S (serine), T (threonine), V (valine), W (tryptophan), and Y (tyrosine).

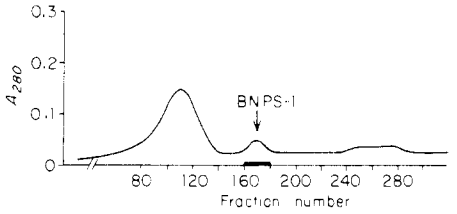

FIGURE 2: Separation of fragments of CB18 generated by cleavage at tryptophanyl residues. The column of Sephadex G-50 SF (140 × 5 cm) was equilibrated and eluted with 9% formic acid at 25 mL/h. Fractions of 5 mL were collected.

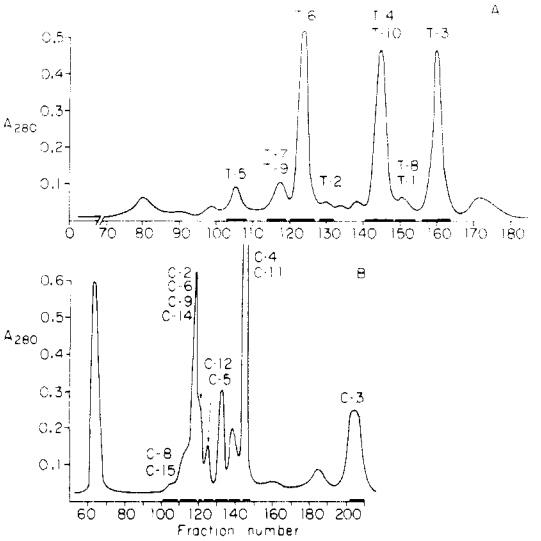

FIGURE 3: Separation of (A) tryptic and (B) chymotryptic peptides from CB18 (7.5 μmol) on a Sephadex G-25 SF column (200 × 2.5 cm) with 9% formic acid at 20 mL/h. Fractions were pooled as indicated by the horizontal bars. In most cases, pooled fractions were subjected to further purification as designated in Tables II and III.

Excess reagent was removed by extraction with chlorobutane. The resulting peptides were separated on a Sephadex G-50 SF column (Figure 2). Since cleavage at tryptophan with BNPS-skatole is reported to occur with a yield of about 60%, cleavage at the 3 tryptophan residues of CB18 was expected to yield a set of 9 overlapping peptides. Nonetheless, one pure peptide fraction (BNPS-1), containing 41 amino acid residues, could be isolated (Figure 2). Its amino acid composition is included in Table II. Sequenator analysis identified 39 residues (yield: 46%; stepwise yield: 94.6%) beginning with Pro-Val-His- and ending with Phe-Pro-Gly-Asp-Val-Asp-?-Leu- which overlaps the C-terminal peptide P-2 (see Figure 1).

Subfragmentation of CB18. Fragment CB18 was only slightly soluble at pH 8 but treatment of 50 mg (5.3 μmol) of the peptide with 2 mg of trypsin in 0.1 M ammonium bicarbonate (pH 8) with vigorous stirring at 37 °C for 24 h yielded a soluble digest. After removal of insoluble material by centrifugation, the soluble peptide fraction was lyophilized. The peptides were separated on a Sephadex G-25 SF column (Figure 3A) and further purified on Dowex 1-X2 or Dowex 50-X2 columns (50 × 0.9 cm). Table II summarizes their amino acid compositions and the purification procedures. Peptides T-5, T-6, T-7, and T-9 were sequenced using the liquid phase method. Manual Edman degradation was used for T-8.

A chymotryptic digest was prepared and fractionated in a similar manner (Figure 3B). Table III shows the amino acid compositions and the purification procedures of the chymotryptic peptides. Peptide C-9 was sequenced using the liquid phase method; it overlaps T-5 and T-6. Peptide C-8 was attached with a 17% yield to the TETA resin by coupling with water-soluble carbodiimide and sequenced by the solid-phase method. No homoserine containing peptides were recovered in either the tryptic or chymotryptic digest.

Figure 1 summarizes the sequence data and indicates the primary structure of CB18. In essence, residues 1–34 were identified by sequenator analysis of CB18 and residues 28–34 were confirmed in peptide C-8. The sequences of residues 35–40 were provided by C-9 and the overlap between Ala-Leu[34] and Glu-Arg[36] was provided by the amino acid sequence

TABLE II: Amino Acid Composition[a] of Peptides Obtained from CB18 by Tryptic Digestion (T) and by Cleavage with BPNS-Skatole (BPNS-1).

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	BNPS-1
Residue Number (Fig. 1)	1	2-8	9-13	14-20	21-36	37-48	49-59	60-63	64-74	75-76	38-78
Asp		1.1 (1)	1.8 (2)		1.0 (1)		1.0 (1)	0.9 (1)	2.0 (2)		3.9 (4)
Thr				0.9 (1)	2.6 (3)	1.0 (1)					1.1 (1)
Ser											0.3 (0)
Glu		1.2 (1)		1.0 (1)	1.9 (2)	1.2 (1)	3.0 (3)				4.0 (4)
Pro					1.1 (1)	1.9 (2)			1.0 (1)		2.7 (3)
Gly									1.0 (1)		1.1 (1)
Ala				0.9 (1)	1.2 (2)				2.8 (3)		3.0 (3)
Cys[b]					0.4 (1)						
Val		1.6 (2)		1.9 (2)	1.0 (1)	0.8 (1)			1.9 (2)		2.7 (3)
Met[c]											0.64 (1)
Ile							2.0 (3)				2.0 (3)
Leu		1.8 (2)	1.0 (1)		2.0 (2)	3.9 (4)	1.0 (1)	1.0 (1)		0.9 (1)	6.5 (7)
Tyr					0.6 (1)		0.7 (1)				0.0 (1)
Phe								0.8 (1)	0.9 (1)		1.8 (2)
His					0.9 (1)	1.0 (1)	0.9 (1)				1.6 (2)
Lys			1.0 (1)	0.9 (1)							0.6 (0)
Arg	1 (1)	1.0 (1)			0.9 (1)	1.1 (1)	0.9 (1)	1.0 (1)	1.0 (1)	1.0 (1)	5.5 (6)
Trp			n.d. (1)	n.d. (1)		n.d. (1)					
Purific. Procedure[d]	(e)		(e)	(e)		(f)	(e)	(e)	(e)	(e)	(g)
Yield %	64	42	36	50	72	11	22	14	30	30	30

[a] The numbers in parentheses indicate those found by sequence analysis. [b] Determined as SCM-Cys. [c] Determined as Hse. [d] All tryptic peptides were first separated on Sephadex G-25 SF (Figure 3A), then purified as indicated. [e] Dowex 50-X2. [f] Dowex 1-X2. [g] Sephadex G-50 SF (Figure 2).

TABLE III: Amino Acid Composition[a] of Chymotryptic Peptides from CB18.

	C2	C3	C4	C5	C6	C7	C8	C9	C11	C12	C14	C15
Residue number (Fig. 1)	4-9	10-11	12-15	16-18	19-24	25-27	28-34	35-41	51-54	55-60	63-68	69-75
Asp	1.1 (1)	(1)	1.0 (1)			1.0 (1)				0.9 (1)		2.0 (2)
Thr				0.8 (1)	0.8 (1)	0.9 (1)	0.9 (1)					
Ser												
Glu	1.2 (1)			1.0 (1)			1.2 (1)	0.9 (1)	1.1 (1)	1.9 (2)		
Pro							0.8 (1)	0.9 (1)				0.9 (1)
Gly												1.0 (1)
Ala			1.1 (1)		1.0 (1)		0.9 (1)				3.0 (3)	
Cys[b]					0.1 (1)							
Val	1.0 (1)			0.8 (1)	0.9 (1)		1.0 (1)	0.8 (1)			1.2 (1)	1.0 (1)
Met[c]												
Ile									0.9 (2)	1.0 (1)		
Leu	1.7 (2)						1.7 (2)	1.0 (1)				1.0 (1)
Tyr					0.7 (1)				0.9 (1)			
Phe										1.1 (1)	0.8 (1)	
His						0.9 (1)		0.8 (1)				
Lys			1.0 (1)		1.0 (1)							
Arg	1.0 (1)							1.0 (1)		1.0 (1)	0.9 (1)	1.0 (1)
Trp		n.d. (1)	n.d. (1)					n.d. (1)				
Purific. Procedure[d]	(e)		(e)	(f)	(e)	(g)	(f)	(g)	(e)	(h)	(g)	(i)
Yield %	13	40	27	10	11	3	8	7	26	20	3	13

[a] The numbers in parentheses indicate those found by sequence analysis. [b] Determined as SCM-Cys. [c] Determined as Hse. [d] All peptides were first separated on Sephadex G-25 SF (Figure 3B), then purified as indicated. [e] Dowex 1-X2. [f] Dowex 50-X2. [g] Dowex 1-X2, then electrophoresis at pH 3.7. [h] Electrophoresis at pH 3.7. [i] Dowex 50-X2, then electrophoresis at pH 3.7.

of T-5. This argument is strengthened by the placement of all other arginine residues in unique sequences. Residues 38–73 were identified by sequenator analysis of BNPS-1 and the carboxyl-terminal sequence (residues 70–78) by solid phase analysis of peptides P-2 and Th-1. These data were confirmed by the remaining sequences of tryptic and chymotryptic peptides.

Amino Acid Sequence of CB15. This fragment contained 163 amino acids, including 1 tryptophan, 16 lysine, 11 arginine, and 1 homoserine residue (Table I). A summary of the strategy

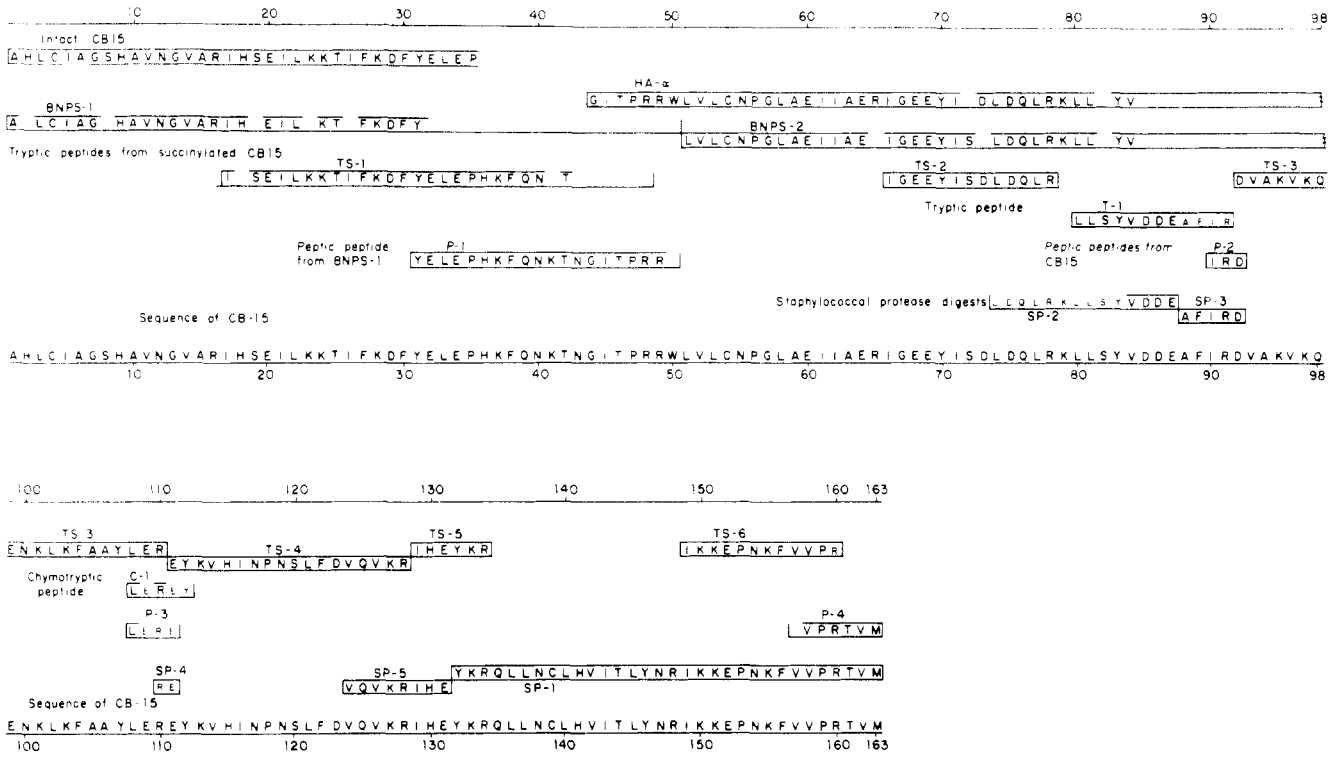

FIGURE 4: Summary proof of the sequence of CB15, using the same annotations as in Figure 1. For simplicity, only those peptides crucial to the proof are shown. The upper half summarizes the data for residues 1–98, the lower half of residues 99–163. The intact fragment and peptides HA-α and BNPS-2 extend to the carboxyl terminus, but are omitted from the lower half. The continuous sequence of TS-3 (residues 92–110) bridges the two portions of the diagram. SP-1 is derived from a staphylococcal protease digest of CB15. SP-2 through SP-5 are derived from a staphylococcal protease digest of BNPS-2.

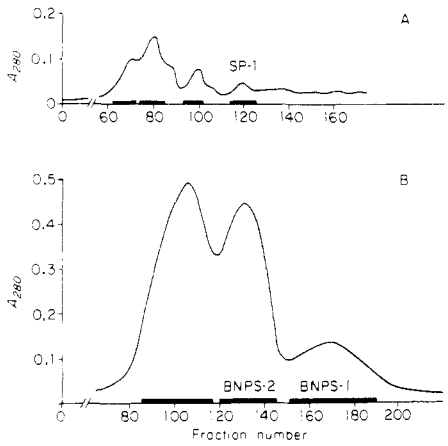

FIGURE 5: (A) Separation of a staphylococcal protease digest of CB15 (2 μmol) on a Sephadex G-50 SF column (112 × 2.5 cm) equilibrated and eluted with 0.1 M ammonium bicarbonate. Flow-rate, 15 mL/h; volume of the fractions, 5 mL. Fractions indicated by bars contained homoserine. (B) Fractionation of peptides obtained by reaction of CB15 (5.5 μmol) with BNPS-skatole. The column (110 × 2.5 cm) containing Sephadex G-75 was eluted with 9% formic acid. Flow rate, 24 mL/h; volume of the fractions, 2 mL. Fractions were pooled as indicated by the horizontal bars.

used in the sequence determination of CB15 is shown in Figure 4.

Amino-Terminal Sequence. Thirty-five residues were determined by sequenator analysis of 1.47 μmol of CB15 with a stepwise degradation yield of 92%.

Carboxyl-Terminal Sequence. Fragment CB15 (200 nmol) was digested with pepsin in 5% formic acid or with chymotrypsin in 0.1 M ammonium bicarbonate (pH 8) for 24 h at 37 °C (enzyme/substrate, 2% w/w). In each case, the homoserine-containing peptide was selectively attached to the TETA resin and sequenced by the solid-phase method. Both digests gave the same result, i.e., Val-Pro-Arg-Thr-Val-Hse. The yield of attachment (estimated by the recovery of Val in the second cycle) was 20% and the stepwise degradation yield was 82%. In an analogous experiment with staphylococcal protease, 2 μmol of CB15 was digested for 24 h in 0.1 M pyridine-acetate, pH 4 (enzyme/substrate = 3% w/w). The digest was separated on a Sephadex G-50 SF column (Figure 5A). The homoseryl peptide of lowest molecular weight (SP-1) was attached to TETA resin and sequenced without further purification in the same manner as above. A single sequence was identified, beginning with Tyr-Lys-Arg- (132–134) and proceeding for 32 cycles through Thr-Val-Hse (Figure 4), thus confirming and extending the sequence of the smaller peptic and chymotryptic homoserine peptides.

Chemical Cleavage of CB15. Cleavage at Tryptophan. Fragment CB15 (5.5 μmol) was reacted with 240 μmol of BNPS-skatole for 6 h in 80% acetic acid at room temperature. The resulting peptides were dissolved in 9% formic acid and separated on a Sephadex G-75 column (Figure 5B). Since CB15 contains only one tryptophan residue, two fragments (BNPS-1 and -2) were recovered and rechromatographed on the same column with a final yield of 27–30% (Table IV). Twenty-six residues were placed within the 50 amino acids of fragment BNPS-1 using the liquid phase method. This analysis showed that BNPS-1 represents the amino-terminal portion of CB15 (residues 1–50). Similar analysis of BNPS-2 identified 31 residues out of 113 (residues 51–163). In the latter experiment, 1.3 μmol of peptide was analyzed in the spinning cup, recovering Pth-Leu at the first turn in 25% yield. The stepwise yield was 94%.

Cleavage at Asn-Gly. Fragment CB15 (2.7 μmol) was dissolved in 10 mL of 6 M guanidine hydrochloride and treated

TABLE IV: Amino Acid Composition[a] of Peptides Obtained from CB15.

	BNPS-Skatole Digest		Tryptic Peptides from Succinylated CB15						Tryptic Peptides	Peptic Peptides			Chymo-tryptic Peptides	Peptides from Staphylococcal Protease Digestion				
	BNPS-1	BNPS-2	TS-1	TS-2	TS-3	TS-4	TS-5	TS-6	T-1	P-1	P-2	P-3	C-1	SP-1	SP-2	SP-3	SP-4	SP-5
Residue Number (Fig. 4)	1-50	51-163	17-48	66-78	92-110	111-128	129-134	149-160	80-91	31-50	90-92	108-111	108-112	132-163	74-87	88-92	110-111	124-131
Asp	3.7(4)	13.0(13)	3.5(3)	2.2(2)	2.8(2)	2.9(3)		1.1(1)	2.2(2)	2.0(2)	0.9(1)			3.8(3)	3.3(3)	1.1(1)		
Thr	2.6(3)	2.4(2)	2.5(3)							1.7(2)				1.9(2)				
Ser	1.7(2)	3.2(3)	1.1(1)	0.9(1)		1.0(1)			0.9(1)						0.9(1)			
Glu	4.3(4)	13.6(14)	4.6(4)	3.1(3)	4.0(3)	2.0(2)	1.3(1)	1.0(1)	0.9(1)	2.9(3)		2.0(2)	2.1(2)	2.4(2)	2.4(2)		1.3(1)	2.4(2)
Pro	1.9(2)	4.3(4)	2.1(2)			1.0(1)		1.9(2)		1.8(2)				1.7(2)				
Gly	2.8(3)	2.8(2)	1.2(1)	1.1(1)						1.0(1)								
Ala	3.9(4)	7.0(6)	0.8(0)		2.5(3)				1.1(1)							1.0(1)		
Cys[b]	0.3(1)	0.2(2)												0.4(1)				
Val	2.3(2)	9.7(11)	0.4(0)		2.0(2)	2.1(3)		1.4(2)	1.0(1)					2.3(4)	1.1(1)			2.1(2)
Met[c]	(0)	0.5(1)												0.8(1)				
Ile	4.0(5)	7.6(9)	3.2(4)	1.9(2)		1.1(1)	0.9(1)	0.8(1)	0.7(1)	1.0(1)	0.9(1)			2.0(2)		1.0(1)		1.0(1)
Leu	3.6(3)	13.7(14)	2.4(2)	1.9(2)	2.3(2)	1.2(1)			1.7(2)	1.0(1)		0.9(1)	0.9(1)	4.4(4)	4.4(4)			
Tyr	0.4(1)	2.4(6)	0.5(1)	0.5(1)	0.2(1)	0.6(1)	0.5(1)		0.7(1)	0.8(1)			0.9(1)	2.0(2)	0(1)			
Phe	2.9(3)	4.3(4)	2.6(3)		1.4(1)	1.0(1)		0.8(1)	0.7(1)	1.0(1)				0.9(1)		1.0(1)		
His	4.0(4)	3.1(3)	1.5(2)			0.9(1)	1.1(1)			1.0(1)				1.0(1)				0.8(1)
Lys	5.0(5)	9.8(11)	5.0(5)		4.4(4)	1.8(2)	1.3(1)	2.1(3)		2.1(2)				4.1(4)	0.8(1)			0.8(1)
Arg	2.8(3)	8.0(8)	0.8(1)	1.0(1)	0.8(1)	1.0(1)	0.8(1)	0.9(1)	0.8(1)	1.8(2)	1.0(1)	1.0(1)	1.0(1)	3.0(3)	0.9(1)	0.8(1)	0.7(1)	0.8(1)
Trp	n.d.(1)									n.d.(1)								
Purific. Proced.	(d)	(d)	(e)	(f)	(e)	(f)	(g)	(g)	(h)	(e)	(i)	(i)	(j)	(e)	(k)	(l)	(l)	(k)
Yield %	30	27	9	19	12	5	6	35	71	45	33	30	26	9	31	21	20	14

[a] The numbers in parentheses indicate those found by sequence analysis. [b] Determined as SCM-Cys. [c] Determined as Hse. [d] Sephadex G-50 SF (Figure 5B). [e] Sephadex G-50 SF (Figure 5A or 6B). [f] Sephadex G-50 SF, then Dowex 50-X2, then Dowex 1-X2. [g] Sephadex G-50 SF, then Dowex 50-X2. [h] Dowex 50-X8. [i] Sephadex G-25 SF plus electrophoresis at pH 3.7. [j] Dowex 50-X8 plus electrophoresis at pH 3.7. [k] Sephadex G-50 SF, then DEAE Sephadex A-25. [l] Sephadex G-50 SF, then electrophoresis at pH 3.7.

at 45 °C (4 h) with 2 M hydroxylamine hydrochloride maintained at pH 9 by the addition of 4.5 M LiOH in a pH-stat. The mixture was then desalted and fractionated on a Sephadex G-75 column (Figure 6A). The peak fraction which eluted first contained a large peptide (HA-α) which was found to be pure by sequenator analysis. Thirty residues were placed in sequence beginning with Gly-44 (Figure 4) and overlapping the sequence of BNPS-2.

Enzymatic Digests of CB15. Cleavage at Arginine. Succinylated CB15 (2.7 μmol) was digested with trypsin at pH 8 in a pH-stat for 2 h at 37 °C (enzyme/substrate, 3% w/w). Tryptic peptides were separated on a Sephadex G-50 SF column (Figure 6B). Six peptides were purified as described in Table IV and sequenced by the liquid phase method (Figure 4). Of these, four could be aligned with peptides already examined. TS-1 extends the N-terminal sequence of intact CB15. TS-2 can be placed within the partial sequences of both HA-α and BNPS-2 providing confirmation of residues 72 and 73. TS-5 and TS-6 overlap the sequence of the carboxyl-terminal fragment SP-1. Two other fragments (TS-3 and TS-4) are discussed below.

Cleavage with Pepsin. Peptide BNPS-1 (3.5 μmol) was digested with pepsin for 4 h in 5% formic acid (enzyme/substrate, 2% w/w). The digest was separated on a Sephadex G-25 SF column (112 × 2.5 cm) in 9% formic acid, yielding peptide P-1 (Table IV) which was sequenced by the liquid phase method. The two arginine residues located in the C-terminal part of the peptide were identified as Pth derivatives by high pressure liquid chromatography. This peptide overlaps the amino-terminal sequence of intact CB15, TS-1 and HA-α (Figure 4), extending the amino-terminal sequence to residue 81.

Cleavage with Trypsin. The amino-terminal sequence was further extended by peptide T-1, isolated from a tryptic digest of 3.8 μmol of CB15 (pH 8, 24 h, 37 °C) separated on a Dowex 50-X8 column (17 × 1.8 cm). This dodecapeptide (Table IV)

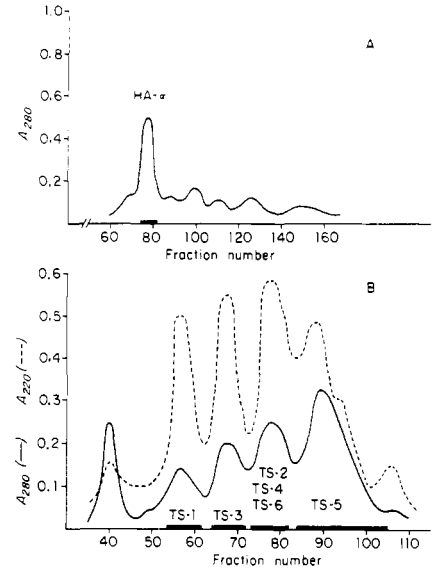

FIGURE 6: (A) Purification of a peptide (HA-α) obtained by reaction of hydroxylamine with CB15 (2.7 μmol). Sephadex G-75 (110 × 2.5 cm) was equilibrated and eluted with 9% formic acid. Flow rate, 10 mL/h; fraction volume, 3.5 mL. (B) Separation of tryptic peptides from succinylated CB15 (2.7 μmol) on a Sephadex G-50 SF column (100 × 2.5 cm) equilibrated and eluted with 0.1 M ammonium bicarbonate. Flow rate, 30 mL/h; volume of fractions, 5 mL. Fractions were pooled as indicated by the bars, and further purified as designated in Table IV.

yielded by manual Edman degradation the sequence Leu-Leu-Ser-Tyr-Val-Asx-Asx-Glx. Chymotryptic subdigestion and paper electrophoresis (pH 3.7) yielded four peptides containing, respectively, leucine; (1.0 Ser, 0.6 Tyr); (1.1 Val, 1.9 Asx, 0.8 Glx, 0.8 Ala, 0.8 Phe); and (1.0 Ile, 1.0 Arg). Digestion of the peptide (Val, Asx, Asx, Glx, Ala, Phe) by carboxypeptidase Y (30 min in 0.02 M sodium acetate, pH 4.9) yielded mole ratios of 0.9 Phe, 0.9 Ala, and 0.38 Glu. Thus, the

sequence of T-1 is Leu-Leu-Ser-Tyr-Val-Asx-Asx-Glu-Ala-Phe-Ile-Arg, the sequence Ala-Phe-Ile-Arg being based on the specificity of chymotrypsin and trypsin. This peptide overlaps the previously established sequence of HA-α and extends the amino-terminal sequence to residue 91, except for residues 85 and 86 which are either aspartic acid or asparagine. These two positions have been identified in peptide SP-2 as aspartyl residues (see below).

At this point, the sequence of the carboxyl-terminal half of CB-15 was only partly solved. Residues 129–163 (Figure 4) had been established by analysis of the overlapping fragments TS-5 and SP-1. Clearly, TS-3 and TS-4 possessed the length and composition to bridge the gap between residues 91 and 129 and it remained to find overlapping peptides to connect these two peptides to T-1 at one end and TS-5 at the other.

Cleavage with Staphylococcal Protease, Chymotrypsin and Pepsin. Towards this end, succinylated BNPS-2 (2 μmol) was dissolved in 8 mL of water adjusted to pH 9 with trimethylamine and digested in a pH stat at pH 8 for 24 h at 37 °C with staphylococcal protease (enzyme/substrate, 4% w/w). The digest was separated on a column of Sephadex G-50 SF (180 × 2.5 cm) equilibrated with 0.1 M ammonium bicarbonate. Four peptides were purified (SP-2 through SP-5) as indicated in Table IV. SP-3 and SP-5 were sequenced by the liquid phase method. The peptide SP-5 provided a key overlap tying together TS-4 and TS-5 and completing the sequence from residue 111 to the carboxyl terminus. The peptide SP-3 overlaps TS-3 and T-1, but by only a single residue (Asp-92). Apparently, the staphylococcal protease cleaved at this aspartyl residue under our experimental conditions. Peptide SP-4 was too small to provide a meaningful overlap of TS-3 and TS-4, but it is consistent with that alignment, and there is no other Arg-Glu sequence in CB15. SP-2 contained 14 amino acid residues and its amino acid composition is in agreement with an overlap of Leu-Asp-Gln-Leu-Arg-Lys-Leu-Leu- in BNPS-2 and Leu-Leu-Ser-Tyr-Val-Asx-Asx-Glu in T-1. Carboxypeptidase Y digestion of SP-2 (30 min, 37 °C, 0.02 M sodium acetate, pH 4.9) yielded 0.61 Glu, 0.84 Asp, 0.4 Val mol/mol of peptide. The latter experiment shows that the two Asx residues in T-1 are aspartyl residues in the sequence Val-Asp-Asp-Glu[87].

A better overlap of TS-3 and TS-4 was found in peptide C-1 (Table IV) obtained from a chymotryptic digest of 2.7 μmol of succinylated CB15 (2 h, pH 8, enzyme/substrate, 5% w/w) in the presence of 1 mM *p*-aminobenzamidine. This pentapeptide was attached to TETA resin with water-soluble carbodiimide and examined by solid phase Edman degradation. Phenylthiohydantoins of Leu and Arg were identified at cycles 1 and 3, respectively, but cycles 2 and 4 did not remove the glutamyl residues which remained covalently attached to the resin. The composition and partial sequence of C-1 are consistent with the overlap at residues 110–111 in Figure 4.

Finally, a peptic digest of CB15 (as above) yielded two peptides P-2 and P-3 which can only be placed to overlap T-1/TS-3 on the one hand and TS-3/TS-4 on the other.

The sequence of CB15 is presented in Figure 4, which includes only peptides essential for the proof of structure. Sequenator analysis of the amino-terminal region of CB15 and of peptide TS-1 identified residues 1 to 40. A peptic peptide, P-1 (residues 31–50), provided the overlap with the partially sequenced peptide HA-α extending the sequence to residue 84 (unidentified residues at positions 72 and 82). Residue 72 was provided by both BNPS-2 and TS-2. The tryptic peptide T-1 (residues 80–91), sequenced by manual techniques, extended the sequence to residue 91. Peptides TS-3 (residues 92–110), TS-4 (residues 111–128), and TS-5 (residues 120–134) were then aligned by overlapping sequences in SP-3, C-1, and SP-5. TS-5 overlapped the sequence of a large homoserine-containing peptide (SP-1) obtained from a staphylococcal protease digest and completed the sequence determination of CB15. Several other short peptides consistent with this sequence were obtained from enzymatic digests, but are not described here since they are not necessary for the proof.

Discussion

Since fragments CB18 and CB15 correspond in size to many naturally occurring proteins, the determination of the amino acid sequence of each presented a major analytical problem, aggravated by the lack of internal cleavage points at methionine residues. These analyses have spanned the period since the initial isolation of these fragments by Saari & Fischer (1973) and the experimental approach was revised as new tactics and methods of sequence analysis were developed. Initially, a series of conventional enzymatic digests using trypsin, chymotrypsin, and pepsin was prepared from each fragment, and a large number of peptides were isolated and their partial structures derived. Later, as more specific methods of limited cleavage at tryptophanyl, arginyl, glutamyl, and Asn-Gly residues were applied, sequence analysis of large fragments proved more effective. These analyses revealed the location of many of the smaller peptides within the overall structure and resulted in redundancies which are included only in part of the proof of structure of CB18 (Figure 1) to document our experimental approach. In the case of fragment CB15, however, the detail of minimal proof is so complex (Figure 4) that redundant information was omitted from this report.

Although structural analyses of large subfragments usually do not provide sufficient evidence to complete a sequence project, it becomes much easier to generate and identify necessary overlapping fragments from conventional enzymatic digests as the project nears completion. For example, to complete the sequence of CB15, it was only necessary to isolate one arginine-containing peptide (P-1) from a peptic digest to overlap and align the three major fragments BNPS-1, TS-1, and HA-2.

It is generally accepted that cleavage techniques, to be productive, must function with a high yield to minimize the problem of purification of fragments from uncleaved parent fragments. In the case of tryptophanyl cleavage where the yield is low (50–60%), the method is usually applied only in instances where a single cleavage point is present. However, this method was used productively with fragment CB18 which contains three tryptophanyl residues. These three residues were so situated that one large fragment (BNPS-1) of unique size was easily separated from the much smaller (or larger) products of incomplete cleavage.

The principal method of elucidating the amino acid sequences of the two fragments was the liquid phase mode of Edman degradation using the method of Edman & Begg (1967) in conjunction with the DMBA buffer of Hermodson et al. (1972). Modification of the protocol for extractions improved the retention of small peptides and extended their stepwise degradation (Koide et al., 1978). In addition, solid phase Edman degradation has proved useful when applied to peptides containing a carboxyl-terminal residue of homoserine. The method of Horn & Laursen (1973) provided a reliable procedure to select and couple to column matrices homoseryl peptides resulting from enzymatic subdigestion of CB15 and CB18. This was usually achieved without prior purification. In one case, however (SP-1 from CB15), incomplete enzymatic digestion required a preliminary partial purification by gel

filtration. Since it is clearly advantageous to perform extended structural analyses of both ends of a large fragment as initial steps, the homoseryl coupling method fulfills a need by providing the carboxyl-terminal sequence of cyanogen bromide fragments. In contrast, carboxylic attachment of peptides to TETA resin using soluble carbodiimide has not been as successful. The yields of attachment were often low and in these cases the results were difficult to interpret.

The analyses of CB18 and CB15 detailed in this communication provide the sequences of two major portions of the total sequence of segment H_s of rabbit muscle glycogen phosphorylase. Their alignment in the whole molecule at residues 350–427 and 441–603 and the interposition of CB9 at residues 428–440 are indicated by Titani et al. (1977). Detailed proof of their placement within the whole molecule and their alignment to the other 19 fragments generated by cyanogen bromide cleavage is presented in the succeeding report by Titani et al. (1978).

Acknowledgments

We are indebted to Dr. David L. Enfield, Dr. Mark A. Hermodson, Mrs. Sheila Balian, and Mrs. Brita Moody for contributions to the initial phases of this work and to Mrs. S. Kumar and Richard R. Granberg for excellent technical assistance. We are also indebted to Dr. E. H. Fischer for his encouragement and for supplying the phosphorylase, and to Dr. A. Koide for his continuing interest in this work.

References

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

Drapeau, G. R., Boily, Y., & Houmard, J. (1972) *J. Biol. Chem. 247,* 6720.

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

Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry 11,* 4493.

Horn, M. J., & Laursen, R. A. (1973) *FEBS Lett. 36,* 285.

Houmard, J., & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A 69,* 3506.

Johnson, L. H., Madsen, N. B., Mosley, J., & Wilson, K. S. (1974) *J. Mol. Biol. 90,* 703.

Koide, A., Titani, K., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1978) *Biochemistry 17* (first of three papers in a series in this issue).

Laursen, R. A. (1971) *Eur. J. Biochem. 20,* 89.

Nagasawa, T., Kuroiwa, K., Narita, K., & Isowa, Y. (1973) *Bull. Chem. Soc. Jpn. 46,* 1269.

Penke, B., Ferenczi, R., & Kowacs, K. (1974) *Anal. Biochem. 60,* 45.

Previero, A., Derancourt, J., Coletti-Previero, M. A., & Laursen, R. A. (1973) *FEBS Lett. 33,* 135.

Saari, J. C., & Fischer, E. H (1973) *Biochemistry 12,* 5225.

Schellenberger, A., Graubaum, H., Meck, C., & Sternkop, G. (1972) *Z. Chem. 12,* 62.

Sygush, J., Madsen, N. B., Kasvinsky, P. J., & Fletterick, R. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74,* 4757.

Titani, K., Koide, A., Hermann, J., Ericsson, L. H., Kumar, S., Wade, R. D., Walsh, K. A., Neurath, H., & Fischer, E. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74,* 4762.

Titani, K., Koide, A., Ericsson, L. H., Kumar, S., Hermann, J., Wade, R. D., Walsh, K. A., Neurath, H., & Fischer, E. H. (1978) *Biochemistry 17* (third of three papers in a series in this issue).