Sequence of the Amino-Terminal 349 Residues of Rabbit Muscle Glycogen Phosphorylase Including the Sites of Covalent and Allosteric Control†

Atsushi Koide,‡ Koiti Titani,§ Lowell H. Ericsson, Santosh Kumar, Hans Neurath, and Kenneth A. Walsh*

ABSTRACT: The sequence of the amino-terminal 349 residues of rabbit muscle glycogen phosphorylase (EC 2.4.1.1) has been determined. Limited proteolysis of native phosphorylase *b* (841 residues, subunit molecular weight 97 412) by subtilisin BPN′, *Streptomyces* alkaline protease, or elastase yielded two large segments (light and heavy). The light segment isolated from the subtilisin digest was cleaved at methionyl bonds with cyanogen bromide to yield eight major fragments and two minor overlapping fragments. The alignment of the major fragments was obtained by analysis of the two minor fragments, of five tryptic peptides containing methionine and of one large fragment generated by cleavage of an aspartylproline bond. Analysis of two cyanogen bromide fragments (CB14 and CB17) isolated from the intact molecule identified the sites susceptible to limited proteolysis and the overlap between the light and the heavy segments. Serine-14 and tyrosine-155 were identified as the residues involved in the covalent and allosteric controls of the enzyme, respectively. Residues 108 and 142 were identified as the cysteine residues reported to be involved in the aggregation of subunits.

Muscle glycogen phosphorylase (1,4-α-D-glucan:orthophosphate α-glucosyltransferase, EC 2.4.1.1) is one of the key enzymes involved in the metabolism of glycogen. Upon activation during muscle contraction, the enzyme catalyzes the conversion of stored glycogen to glucose 1-phosphate, initiating glycogenolysis. Since its isolation and crystallization (Green & Cori, 1943), phosphorylase has attracted the interest of numerous investigators because of its important physiological role, its wide distribution in eukaryotic and prokaryotic cells, and the complex mechanisms by which its enzymatic activity is controlled (Fischer et al., 1971; Graves & Wang, 1972; Fischer et al., 1970).

It is apparent that a full understanding of the regulation and mechanism of action of phosphorylase requires a detailed analysis of the structure of the two identical subunits of the protein and their molecular assembly. As an initial approach, Saari & Fischer (1973) isolated 18 fragments generated by cleavage with cyanogen bromide. Some of these were related to function by specific labeling with phosphate or pryidoxal 5′-phosphate prior to fragmentation. Titani et al. (1975) showed that one of these fragments (CB14), labeled by [^{32}P]ATP in the presence of phosphorylase kinase, contains an acetyl group at its amino terminus, as does the whole molecule. From these observations serine-14 was identified as the site phosphorylated in the conversion of phosphorylase *b* to *a*.

The strategy of Saari & Fischer (1973) is extended herein and in the accompanying papers (Hermann et al., 1978; Titani et al., 1978) by taking advantage of a finding by Raibaud & Goldberg (1973) that subtilisin cleaves native phosphorylase into two reasonably homogeneous and complementary segments of ca. 30 000 and 70 000 daltons. Further chemical cleavage of each isolated segment at methionyl-, aspartylproline, and asparaginylglycine bonds facilitated the determination of the complete amino acid sequence of this molecule. The general strategy followed for establishing the total amino acid sequence has already been presented (Titani et al., 1977); the experimental details are presented in the present papers. This report describes procedures for generating segments of the molecule by subtilisin digestion, and sequence analyses of fragments comprising the amino-terminal 349 residues, including the entire 30 000 dalton segment.

Materials and Methods

Crystalline rabbit muscle glycogen phosphorylase *b* was generously supplied by Dr. Edmond H. Fischer and used throughout the present study without further purification. [^{14}C]-8-[*m*-(*m*-Fluorosulfonylbenzamido)benzylthio]adenine-modified rabbit muscle glycogen phosphorylase *b* (Anderson et al., 1973) was donated by Dr. Donald J. Graves (Iowa State University). Nagarse (subtilisin BPN′) and alkaline protease from *Streptomyces* were gifts of Teikoku Chemical Industry and Seikagaku Kogyo (Japan), respectively. Porcine elastase was prepared by the method of Lewis et al. (1956) and further purified on a column of DEAE-Sephadex A-25 (Smillie & Hartley, 1966). Bovine plasmin, factor X_{1a} (the activated form of blood coagulation factor X_1), and a protease from Russell's viper venom were donated by Dr. Kazuo Fujikawa (University of Washington). Protease from *Staphylococcus aureus* was a gift of Dr. G. R. Drapeau (University of Montreal) or a product of Pierce Chemicals. Carboxypeptidase Y was a gift from Dr. M. Ottesen (Carlsberg Laboratory, Copenhagen). Thermolysin and Pronase were purchased from Calbiochem, and pepsin, α-chymotrypsin (subsequently treated with Tos-LysCH_2Cl[1]), Tos-PheCH_2Cl-trypsin, and carboxypeptidases A and B were Worthington products.

† 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 (GM-15731 and AM-7902).

‡ Present address: Eisai Research Laboratories, Eisai Co., Ltd., Koishikawa 4, Bunkyo-ku, Tokyo, Japan.

§ Investigator of the Howard Hughes Medical Institute.

[1] Abbreviations used: AMP, adenosine 5′-phosphate; BNPS-skatole, bromine adduct of 2-(2-nitrophenylsulfenyl)-3-methylindole (Omenn et al., 1970); CM-Cys, *S*-carboxymethylcysteine; DITC-glass, isothiocyanatophenylthiocarbamoylaminopropyl glass; $NaDodSO_4$, sodium dodecyl sulfate; TETA, triethylenetetramine; Tos-LysCH_2Cl, L-1-chloro-3-tosylamido-7-amino-2-heptanone; Tos-Phe-CH_2Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone.

0006-2960/78/0417-5657$01.00/0

Sephadex and ion-exchange Sephadex of various grades were obtained from Pharmacia Fine Chemicals, Bio-Gels and ion-exchange resins (AG 50W-X2, AG 50W-X8 and AG 1-X2) were from Bio-Rad Laboratories, and guanidine hydrochloride was from Heico. Reagents used with the Sequencer were of Sequenal grade from Pierce; solvents were from Burdick and Jackson.

Cyanogen bromide, BNPS-skatole, and dithioerythritol were products of Pierce Chemicals. 4-Vinylpyridine (Aldrich) was clarified by distillation at reduced pressure; iodoacetic acid (Sigma) was recrystallized from hot chloroform and stored in a dark bottle. All of these reagents were stored in a freezer. Amino-terminal sequences of large fragments were analyzed with Beckman Sequencers (Model 890B) according to the method of Edman & Begg (1967) with a "protein program" as described by Hermodson et al. (1972). Repetitive yields of 93–96% were routinely observed. Small peptides were retained in the spinning cup by using a "peptide program" adapted from the techniques of Crewther & Inglis (1975), who precipitated the coupled peptide with benzene before reagent extraction, and removed the last traces of the heptafluorobutyric acid as an azeotropic mixture with chlorobutane before product extraction (cf. Hermodson et al., 1977). The products of degradations were identified primarily by gas–liquid chromatography after silylation (Pisano & Bronzert, 1972) and spot tests (Hermodson et al., 1972) and in a few experiments by high pressure liquid chromatography (Bridgen et al., 1976) as described in the accompanying paper (Hermann et al., 1978). The phenylthiohydantoin of ε-succinyllysine eluted between those of tyrosine and tryptophan during gas–liquid chromatography.

Some amino-terminal sequence analyses were performed with a Sequemat (Model 12) after specific attachment of the carboxyl terminus of the peptide to aminated resins or glass beads. The details of these solid phase techniques are described in the accompanying paper (Hermann et al., 1978). With small peptides subtractive Edman degradations were occasionally used (Konigsberg & Hill, 1962). Carboxyl-terminal sequences were determined using carboxypeptidases A and B (Ambler, 1967) or carboxypeptidase Y (Hayashi, 1977).

Miscellaneous Methods. Amino acid analysis was performed with a Durrum D-500 amino acid analyzer following manufacturer's instructions. Cysteine was determined as cysteic acid (Hirs, 1967), *S*-carboxymethylcysteine (Crestfield et al., 1963) or *S*-pyridylethylcysteine (Cavins & Friedman, 1970), and methionine as methionine sulfone (Hirs, 1967). Tryptophan was determined after base hydrolysis (Hugli & Moore, 1972) as a tryptophan/histidine ratio.

Molecular weights were determined by $NaDodSO_4$ gel electrophoresis (Weber & Osborn, 1969) using as standards proteins or peptide fragments previously isolated in our laboratory.

Carboxymethylation was accomplished by a modification of the method of Crestfield et al. (1963) using 7 M guanidine hydrochloride and dithioerythritol in place of urea and β-mercaptoethanol. The alkylated protein was separated from excess reagents and salts by dialysis against water in the dark.

Generation of Fragments. Cleavage at methionyl bonds followed the general procedure of Gross (1967). The protein was dissolved in 72% formic acid (10 mg/mL) and treated with CNBr (20 mg/mL) at room temperature for 15–20 h. The reaction mixture was then diluted tenfold with water and lyophilized.

Arginyl bonds were selectively cleaved with trypsin after succinylation of lysyl residues. Succinylation was carried out at room temperature in a pH-stat (pH 7–8) as described by Yaoi et al. (1964), except that 6 M guanidine hydrochloride was used instead of 7 M urea.

Tryptophanyl bonds were cleaved by a modification of the method of Omenn et al. (1970). Protein (10 mg) was suspended in 0.8 mL of glacial acetic acid and 5–10 mg of BNPS-skatole was added. After the reagent had completely dissolved, 0.2 mL of water was added and the mixture was stirred at room temperature for 6 h in the dark. Following dilution with an equal volume of water, excess reagent was extracted three times with an equal volume of 1-chlorobutane. Clear separation of the two phases was effected by centrifugation and the aqueous layer was lyophilized.

Asp-Pro bonds were cleaved as described by Fraser et al. (1972). Protein (0.05–5 mg/mL) was dissolved in 10% acetic acid containing 7 M guanidine hydrochloride, adjusted to pH 2.5 with pyridine, and incubated at 37 °C for 4 days. The products were either desalted on Sephadex G-25 (in 9% formic acid) prior to separation or directly separated by gel filtration on larger pore columns.

Digestion with Tos-PheCH_2Cl-trypsin, thermolysin, or α-chymotrypsin (the latter in 1 mM *p*-aminobenzamidine) was carried out at pH 8.0 either in a pH-stat (until acid production ceased), or in 0.1 M NH_4HCO_3 for 15–20 h (37 °C). Digestion with pepsin took place at 37 °C in 9% formic acid overnight. Staphylococcal protease digests were incubated for 18 h at 37 °C either in 0.1 M NH_4HCO_3, pH 8.0, or in 0.1 M pyridine–acetate, pH 4.0 (Houmard & Drapeau, 1972; Drapeau et al., 1972), depending upon the solubility of the substrate. All digests contained 1% protein and a 2:100 molar ratio of enzyme to substrate.

Enzymatic digests were usually separated either on a column of Ag 50W-X8 using a double linear gradient of pyridine acetate buffers (Bradshaw et al., 1969) or on a column of Sephadex G-25 (or G-50) Superfine in 9% formic acid or 0.1 M NH_4HCO_3, pH 8.0. Further purification of peptides utilized columns of AG1-X2 (Wikler et al., 1970), AG 50W-X2 (Schroeder, 1967), or DEAE-Sephadex A-25 (Salnikow et al., 1973) in conjunction with volatile buffers. Paper electrophoresis at pH 3.6 or 6.5 (Ryle et al., 1955; Bennett, 1967) was also used for purification of small peptides.

Large fragments were separated by gel filtration on columns of Sephadex (of various grades) in 9% formic acid or 0.1 M NH_4HCO_3, pH 8.0. Insoluble fragments required the presence of either 7 M urea or 6 M guanidine hydrochloride. SP-Sephadex C-25 was also used in 7 M urea as follows. The matrix was allowed to swell in water and then successively washed with 5–10 volumes each of 1 N HCl, water, 1 N NaOH, and water on a glass filter. The air-dried product was then suspended in 7 M urea–0.1 M sodium formate, pH 2.9 (7 mol of urea was dissolved in 9% formic acid, the volume adjusted to 1 L, and 4 g of NaOH added), and packed in a column (2.5 × 50 cm). After equilibration at room temperature with 300 mL of the same buffer at 80 mL/h, the sample (100–500 mg) was applied in 5–10 mL of the same buffer. The column was eluted at 80 mL/h by application of a linear gradient established between 1 L of that buffer and 1 L of 7 M urea–0.75 M sodium formate, pH 3.9 (the second buffer is made by adding 26 g of NaOH to 1 L of the first buffer).

Preparation of Fragments of CB14 and CB17. One gram of a cyanogen bromide digest of *S*-carboxymethylphosphorylase was fractionated on a column (5 × 140 cm) of Sephadex G-50 Superfine using 9% formic acid. Fractions of 15 mL were collected at 60 mL/h. After monitoring the column at 280 nm, six fractions were pooled as described by Saari & Fischer (1973; pools A through F in Figure 1 of their paper). Fraction

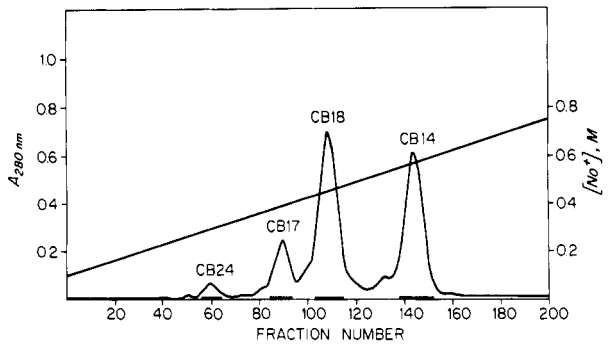

FIGURE 1: Fractionation of 350 mg of fraction C (from Figure 1 in Saari & Fischer, 1973) from CNBr-treated *S*-carboxymethylphosphorylase on a column (2.5 × 50 cm) of SP-Sephadex C-25. Fragments were eluted at room temperature at 80 mL/h by a linear 2-L gradient of 0.1 M sodium formate, 7 M urea, pH 2.9, to 0.75 M sodium formate, 7 M urea, pH 3.9. Fractions of 10 mL were collected and pooled as indicated by horizontal bars.

C (approximately 350 mg) was further separated on a column of SP-Sephadex C-25 into four fractions as shown in Figure 1. Each of these fractions was desalted on a column (2.5 × 40 cm) of Sephadex G-25 Fine in 9% formic acid and lyophilized. $NaDodSO_4$ gel electrophoresis and sequencer analysis indicated that each fraction was homogeneous, and corresponded to CB17 (40 mg), CB18 (65 mg), CB14 (65 mg), and to a new fragment, CB24 (30 mg). The sequences of CB18 and CB24 are reported separately by Hermann et al. (1978) and Titani et al. (1978).

Limited Proteolysis of Phosphorylase b. The procedure of Raibaud & Goldberg (1973) was followed, but with some modifications, i.e., soluble enzymes were used for greater convenience in large scale digestion, and proteolysis was terminated by the addition of trichloroacetic acid instead of phenylmethanesulfonyl fluoride.

(1) Small Scale Experiments. A suspension of crystalline phosphorylase *b* (approximately 10 mg/mL) was dialyzed at 5 °C for 15–20 h against 0.05 M Tris acetate, pH 8.5, containing 0.01 M β-mercaptoethanol. After dilution to 4.8 mg/mL with the same buffer, AMP (6.9 mg/mL) and an equal volume of a protease solution (32.2 μg/mL) in the same buffer were added. Following incubation at 5 °C for 16 h, the protein was precipitated by adding an equal volume of 5% trichloroacetic acid. The precipitate, collected by centrifugation, was reduced, carboxymethylated, and, after dialysis against water, lyophilized.

(2) Large-Scale Digestion with Subtilisin BPN′. Approximately 1 g of phosphorylase *b* in 100 mL was dialyzed at 5 °C for 16 h against 6 L of 0.05 M Tris–acetate, pH 8.5, containing 0.01 M β-mercaptoethanol. Following dilution to 210 mL with the same buffer and addition of 306 mg of AMP, 2 mg of subtilisin, dissolved in 210 mL of the same buffer without β-mercaptoethanol, was added with gentle stirring. After 12 h at 5 °C, the protein was precipitated by addition of 620 mL of 5% trichloroacetic acid; the product was reduced, carboxymethylated, dialyzed, and lyophilized. The yield of product was 90–95% by weight.

(3) Separation of the Products of Limited Proteolysis. Ten- to twenty-milligram samples were fractionated either on a column (1.5 × 85 cm) of Sephadex G-150 in 9% formic acid containing 7 M urea or on a column (1.5 × 50 cm) of Sephadex G-150 in 0.1% $NaDodSO_4$. Larger samples (e.g., 400 mg) were dissolved in 40 mL of 6 M guanidine hydrochloride and separated on a column (5.0 × 85 cm) of Sephadex G-150 equilibrated and developed with 9% formic acid containing 7 M urea. In each case the eluate was monitored by absorbance at 280 nm; the fractions were pooled, dialyzed, and lyophilized.

Results

General Strategy. Titani et al. (1975) have shown that fragment CB14 is derived from the amino terminus of phosphorylase. Raibaud & Goldberg (1973) have shown that subtilisin cleaves phosphorylase *b* within this fragment and within CB17 to generate a 30 000 dalton segment and a carboxyl-terminal 70 000 dalton segment. The proof of sequence of the 30 000 dalton segment (L_s) is given below, together with those of CB14 and CB17. First, the products of limited proteolysis are characterized and segment L_s is isolated. Then methionyl bonds in segment L_s are cleaved with CNBr to generate 8 primary fragments, and in whole phosphorylase to generate CB14 and CB17. Finally, detailed analysis of each fragment and certain overlapping structures (Figure 2) provide the sequence of the amino-terminal 349 residues.

Limited Proteolysis of Phosphorylase b with Various Proteases. Figure 3 shows the $NaDodSO_4$ gel electrophoresis patterns of the trichloroacetic acid insoluble products generated by mild proteolysis of phosphorylase *b* by subtilisin BPN′, *Streptomyces* alkaline protease, elastase, trypsin, Pronase, chymotrypsin, or thermolysin. No appreciable proteolysis was observed under the same conditions with plasmin, blood coagulation factor X_{1a}, or a protease from Russell's viper venom. From these results, the effectiveness of limited proteolysis of native phosphorylase *b* appears to fall into three general types.

Type I. Subtilisin, *Streptomyces* alkaline protease, and elastase each cleave the *b* form of phosphorylase to yield two primary fragments which will be designated as "light" and "heavy" segments.

Type II. Trypsin or Pronase yields one major fragment which is slightly smaller than the type I heavy segment. Although a few minor fragments are also observed, it appears that the rest of the molecule is cleaved into small peptides which are not precipitated by trichloroacetic acid.

Type III. Chymotrypsin or thermolysin slowly yields essentially one major fragment, slightly larger than the heavy fragment generated by type I proteases.

Separation of the Fragments Generated by Various Proteases. Trichloroacetic acid insoluble fragments of phosphorylase *b* were generated by limited proteolysis by various proteases and separated on analytical columns of Sephadex G-150 in 0.1% $NaDodSO_4$. As is to be expected from the corresponding $NaDodSO_4$ gel patterns (Figure 3), similar elution profiles were obtained with products of digestion by all type I proteases; these differed from the profiles of products of digestion by type II and III proteases. In large scale separations of segments generated by subtilisin (Figure 4), the samples were dissolved in 6 M guanidine hydrochloride (in 9% formic acid) and the columns developed with 7 M urea, 9% formic acid. Otherwise, most of the light segment aggregated and was eluted immediately before the heavy segment.

Characterization of Fragments Generated by Type I and II Proteases. The molecular weights of four large fragments isolated from various digests were estimated from duplicate $NaDodSO_4$ gel electrophoretic analyses. Table I summarizes both their molecular weights and their amino-terminal tripeptide sequences. The molecular weights of the three heavy segments, H_s, H_{st}, and H_e (Table I), are similar (68 000–70 000), while the heavy segment H_t-I (from the tryptic digest) is slightly smaller (65 000). The values for the three light segments, L_s, L_{st}, and L_e, are also similar (30 000–31 000).

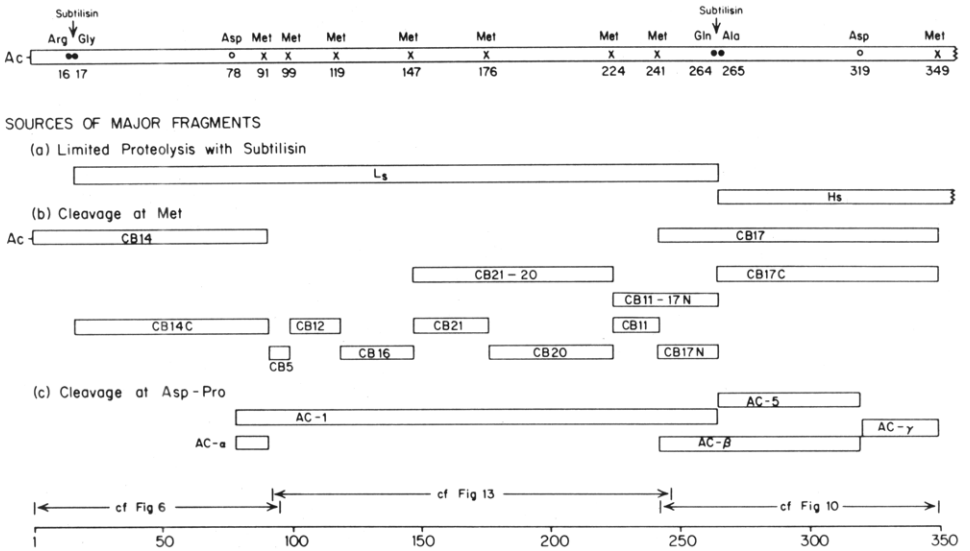

FIGURE 2: Diagrammatic summary of the major fragments generated for sequence analysis of the amino-terminal 349 residues of the phosphorylase subunit. The top bar represents this portion of the molecule and the residues (numbers below the bar) that are important for its fragmentation. L_s and H_s denote the light segment and a portion of the heavy segment generated by subtilisin cleavage. The prefix CB- indicates a fragment generated by cleavage with cyanogen bromide. The prefix AC- identifies a fragment generated by acid cleavage of L_s (AC-1 and AC-5), CB14 (AC-α), or CB17 (AC-β and AC-γ). Details of the proof of sequence are found in Figures 6, 10, and 13 for the regions indicated in the lower portion of this diagram.

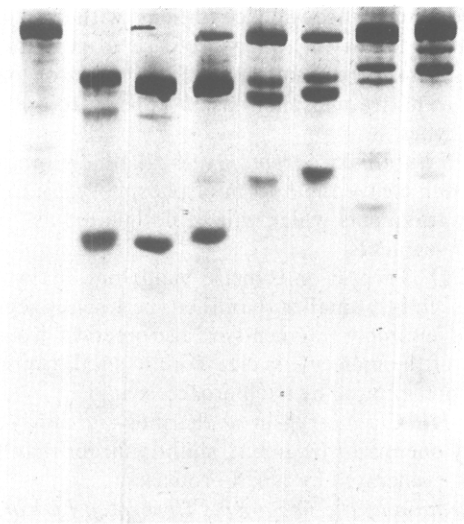

FIGURE 3: $NaDodSO_4$ disc gel electrophoresis of rabbit muscle phosphorylase *b* and the products generated by limited proteolysis. Left to right: intact protein; digestion products with subtilisin BPN′, *Streptomyces* alkaline protease, elastase, trypsin, Pronase, chymotrypsin, and thermolysin. Samples of 10–20 μg were applied to the gels. The concentration of cross-linking agent was 0.6%.

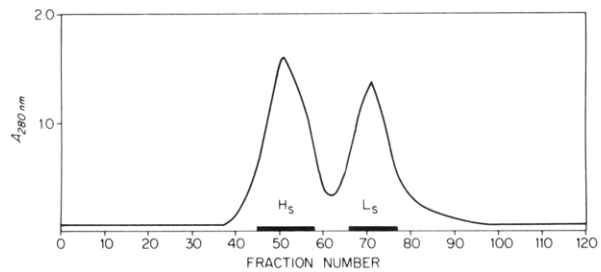

FIGURE 4: Separation of segments L_s and H_s from a subtilisin digest of native phosphorylase *b*. Four hundred milligrams was dissolved in 20 mL of 6 M guanidine hydrochloride and fractionated on a column (5 × 100 cm) of Sephadex G-150 equilibrated with 7 M urea in 9% formic acid. Fractions of 10 mL were collected at 13 mL/h.

These values are in good agreement with those for the corresponding fragments generated by Raibaud & Goldberg (1973) from phosphorylase *a* by subtilisin Carlsberg. In all cases except tryptic digestion, the sum of the molecular weights of the heavy and light segments approximates the molecular weight of the whole protein. In the case of tryptic digestion, the sum of the molecular weights of two large fragments, H_t-I and H_t-II (a minor fragment) is larger than the molecular weight of phosphorylase; thus, segment H_t-II is considered to be a degraded form of segment H_t-I.

The amino acid compositions of the two segments, H_s and L_s, isolated from the subtilisin digest, are listed in Table II. Similar results (not shown) were obtained with fragments from the two other type I proteases. The compositions are similar to those of fragments I and II of Raibaud & Goldberg (1973), although there are some differences in values for serine, proline, valine, methionine, and isoleucine.

Sequencer analysis indicated that each of the three type I proteases cleaved at different points. However, with the exception of H_t-II, short amino-terminal sequences could be determined in each case, from which ultimately the sites of cleavage in the phosphorylase molecule could be deduced. These results and those shown in Figure 3 indicate that native phosphorylase *b* was primarily cleaved in two restricted regions by subtilisin, *Streptomyces* alkaline protease or elastase, but the specific peptide bonds cleaved depended upon the substrate specificity of each enzyme. Since the α-amino group of phosphorylase is acetylated (Titani et al., 1975) and since neither the light nor the heavy segment is blocked, one of the regions in phosphorylase *b* which was cleaved must be very close to the amino terminus of the native enzyme.

Cleavage of Segment L_s with Cyanogen Bromide and Separation of Fragments. Approximately 7 μmol (205 mg) of the light segment L_s was cleaved with cyanogen bromide as described in Materials and Methods. Three equal aliquots of the products were fractionated in 9% formic acid on a column (2.5 × 110 cm) of Sephadex G-50 Superfine (Figure 5A). Of the seven fractions pooled, only fraction III approached purity, which was finally achieved by rechromatography on the same column (fragment CB20). Its unique amino-terminal sequence

TABLE I: Limited Proteolysis of Rabbit Muscle Phosphorylase *b*.

enzyme used for proteolysis	light segment			heavy segment		
		mol wt[a]	amino-terminal sequence[b]		mol wt[a]	amino-terminal sequence[b]
subtilisin BPN′	L_s	30 000	Gly_{17}-Leu-Ala - - -(80%) Ala_{19}-Gly-Val - - - (20%)	H_s	68 000	Ala_{265}-Val-Leu - - - (80%) Ile_{263}-Gln-Ala - - - (10%) Leu_{267}-Asp-Arg - - - (10%)
Streptomyces alkaline protease	L_{st}	31 000	Ala_{19}-Gly-Val - - - (80%) Gly_{17}-Leu-Ala - - - (20%)	H_{st}	68 000	Ala_{265}-Val-Leu - - - (80%)
elastase	L_e	30 000	Ser_{14}-Val-Arg - - - (50%) Gly_{20}-Val-Glu - - - (50%)	H_e	70 000	Gly_{261}-Tyr-Ile - - - (80%)
trypsin				H_t-I H_t-II	65 000 45 000	Val_{278}-Leu-Tyr - - - (60%) not determined

[a] By $NaDodSO_4$ gel electrophoresis. [b] By Sequencer analysis. Approximate ratios are given in cases where more than one sequence was observed. The subscript refers to the residue number in whole phosphorylase (Figure 14).

TABLE II: Amino Acid Compositions of Segments L_s and H_s.[a]

	L_s	H_s	$L_s + H_s$	rabbit muscle phosphorylase	
				by amino acid analysis[g]	from sequence[h]
Asx	29.4	63.1	92.5	95.2	96
Thr[b]	15.8	23.0	38.8	33.2	35
Ser[b]	5.5	21.7	27.2	28.7	29
Glx	28.9	65.1	94.0	99.0	95
Pro	10.2	27.5	37.7	41.2	36
Gly	20.8	28.3	49.1	48.1	48
Ala	19.0	44.0	63.0	63.2	63
cystine/2[c]	2.7	6.6	9.3	8.2	9
Val[d]	15.8	42.1	57.9	59.5	62
Met[e]	6.1	13.5	19.6	21.2	21
Ile[d]	7.6	41.9	49.5	47.1	49
Leu	23.9	56.7	80.6	78.9	79
Tyr	15.1	19.5	34.6	35.6	36
Phe	11.3	30.3	41.6	37.3	38
His	6.7	15.0	21.7	20.2	22
Lys	10.5	35.8	46.3	45.6	48
Arg	15.3	45.2	60.5	61.9	63
Trp[f]	6.5	5.2	11.7	12.5	12

[a] Residues/molecule on the basis of 19 and 44 residues of alanine in segments L_s and H_s. [b] Extrapolated to zero time. [c] Estimated as carboxymethylcysteine. [d] Values after 96-h hydrolysis. [e] Estimated as methionine sulfone after performic acid oxidation. [f] Determined by the method of Hugli & Moore (1972). [g] From Table II in Sevilla & Fischer (1969), recalculated for mol wt 97 000. [h] From Titani et al. (1977).

was shown to be Glu-Glu-Ala and its composition is given in Table III.

Fraction VII was fractionated on a column of AG 50W-X2 (Figure 5C), which yielded two major components, each with a composition identical with that of fragment CB5 of Saari & Fischer (1973). Apparently they differ only in that one has homoserine at the carboxyl terminus and the other has homoserine lactone.

Since fractions V and VI were difficult to redissolve in dilute acetic acid, each (14 mg of fraction V and 30 mg of fraction VI) was separately dissolved in 2.5 mL of 50% acetic acid (containing a few drops of 88% formic acid). Upon tenfold dilution with water, precipitation occurred. Each suspension was adjusted to pH 4.0 with pyridine and stored at 5 °C for 2 days. The precipitate was then collected by centrifugation, washed with water, and lyophilized. The supernatant from fraction V was refractionated on the same column of Sephadex G-50 to yield a major fraction (V-1) which had a composition (Table III) similar to that of fragment CB16 of Saari & Fischer and a single amino-terminal sequence, Glu-Glu-Leu. The supernatant from fraction VI was separated by ion-exchange chromatography into three fragments (Figure 5B). Two of these (VI-2 and VI-3) corresponded to CB11 and CB12 of Saari & Fischer and contained amino-terminal sequences of Pro-Tyr-Asp and Val-Asn-Leu, respectively. The third fraction, VI-1, yielded a single amino-terminal sequence, Arg-Leu-Trp, which was identical with that of CB17. However, the amino acid composition is quite different from that of CB17 and the molecular weight is smaller, suggesting that fraction VI-1 represents the amino-terminal portion of CB17 (CB17N), generated by subtilisin cleavage in its interior. Since the precipitates from fractions V and VI had identical amino acid compositions and amino-terminal sequences (Ala-Thr-Leu), they were combined and designated CB21.

Sequencer analysis of fraction I (after rechromatography on the original column) indicated that it represented the amino-terminal portion of segment L_s because it had the same ratio (Table I) of two amino-terminal sequences, Gly-Leu-Ala and Ala-Gly-Val. It will be shown later that these two sequences overlap within the structure Arg-Gly-Leu-Ala-Gly-Val-Glu_{22} in the interior of CB14. Hence, fraction I is designated CB14C.

Thus, five of the seven pooled fractions in Figure 5A yielded eight fragments, designated CB20, CB5, CB16, CB11, CB12,

TABLE III: Amino Acid Composition of Cyanogen Bromide Fragments.[a]

	CB14	CB14C	CB5	CB12	CB16	CB21	CB21-20	CB20	CB11	CB11-17N	CB17N	CB17	CB17C
Residues	1-91	17-91	92-99	100-119	120-147	148-176	148-224	177-224	225-241	225-264	242-264	242-349	265-349
Figure	1	5-A	5-C	5-B	5-A	5-A	5-A	5-A	5-B	5-A	5-B	1	c
Fraction		I	VII-1, VII-2	VI-3	V	V and VI[b]	II	III	VI-2	IV	VI-1		
Amino acids													
Asx	9.1 (9)	6.7 (8)	1.1 (1)	4.1 (5)	3.9 (3)	1.6 (1)	5.7 (5)	4.1 (4)	4.4 (4)	7.9 (9)	5.2 (5)	16.1 (18)	12.0 (13)
Thr	5.2 (5)	4.7 (5)	2.0 (2)	1.1 (0)	0.6 (0)	0.8 (1)	3.8 (4)	2.6 (3)	1.9 (2)	2.0 (2)	0.5 (0)	3.2 (3)	3.6 (3)
Ser	3.5 (4)	1.4 (1)		0.3 (0)	0.9 (1)	0.2 (0)	0.9 (1)	1.2 (1)	0.2 (0)	0.9 (1)	0.9 (1)	4.5 (5)	3.9 (4)
Hse	0.8 (1)	0.7 (1)	0.8 (1)	1.1 (1)	0.7 (1)	0.7 (1)	0.6 (2)	1.0 (1)	0.6 (1)	0.6 (1)	0.1 (0)	0.6 (1)	0.5 (1)
Glx	8.9 (9)	6.1 (6)	0.9 (1)	3.3 (3)	6.7 (6)	3.3 (3)	9.6 (10)	6.9 (7)	0.2 (0)	1.4 (1)	0.7 (1)	9.9 (9)	8.6 (8)
Pro	3.0 (3)	2.3 (2)			0.3 (0)	0.1 (0)	3.1 (3)	2.9 (3)	3.0 (3)	3.8 (4)	2.3 (1)	6.1 (6)	5.0 (5)
Gly	3.0 (3)	3.5 (3)	0.8 (1)	1.3 (1)	5.1 (5)	5.5 (6)	9.2 (9)	3.4 (3)	1.4 (1)	2.9 (3)	2.1 (2)	4.1 (4)	3.0 (2)
Ala	4.1 (4)	4.7 (4)		3.3 (3)	3.4 (3)	3.0 (3)	7.0 (7)	4.0 (4)	0.3 (0)	2.2 (2)	2.0 (2)	9.0 (9)	7.0 (7)
CmCys				0.9 (1)	1.0 (1)	0.7 (1)	0.5 (1)			0.2 (0)		(1)	0.4 (1)
Val	6.9 (7)	6.3 (6)		1.2 (1)		0.1 (0)	3.8 (5)	3.9 (5)	2.6 (3)	3.1 (4)	1.6 (1)	6.7 (7)	5.9 (6)
Ile	2.6 (3)	2.3 (2)		0.1	1.0 (1)	2.3 (3)	2.2 (3)	0.5 (0)	0.1 (0)	0.8 (1)	0.7 (1)	5.1 (5)	4.8 (4)
Leu	8.9 (9)	8.0 (8)	0.8 (1)	4.0 (4)	5.4 (5)	2.6 (2)	6.3 (5)	3.2 (3)	0.3 (0)	2.2 (2)	2.1 (2)	10.9 (11)	9.4 (9)
Tyr	5.5 (7)	5.9 (7)		1.0 (1)		2.2 (3)	3.9 (5)	1.9 (2)	1.4 (2)	1.9 (3)	1.0 (1)	2.4 (3)	1.9 (2)
Phe	3.9 (4)	4.3 (4)			1.0 (1)	1.7 (2)	4.1 (4)	2.0 (2)	0.2 (0)	1.9 (2)	2.0 (2)	8.4 (9)	6.6 (7)
His	5.2 (5)	4.6 (5)					1.2 (2)	1.7 (2)			0.1 (0)	1.1 (1)	1.3 (1)
Lys	6.7 (7)	4.8 (5)		0.4 (0)		0.8 (1)	3.0 (3)	1.9 (2)	0.4 (0)	2.0 (2)	2.1 (2)	6.7 (7)	5.1 (5)
Arg	10.0 (10)	7.3 (7)	1.0 (1)		1.2 (1)	0.9 (1)	3.6 (4)	3.0 (3)	1.0 (1)	2.0 (2)	1.3 (1)	7.8 (8)	6.9 (7)
Trp	(1)	(1)				(1)	(4)	(3)		(1)	(1)	(1)	
No. of Residues	91	75	8	20	28	29	77	48	17	40	23	108	85
% Yield	~40	59	29	37	18	41	5	60	40	5	25	~20	~20

[a] Residues/molecule by amino acid analysis or (in parentheses) from the sequence (Figure 14). [b] Precipitated (see text). [c] Isolated by Titani et al. (1978) from segment H_s as their fraction II-2 in Figure 2B.

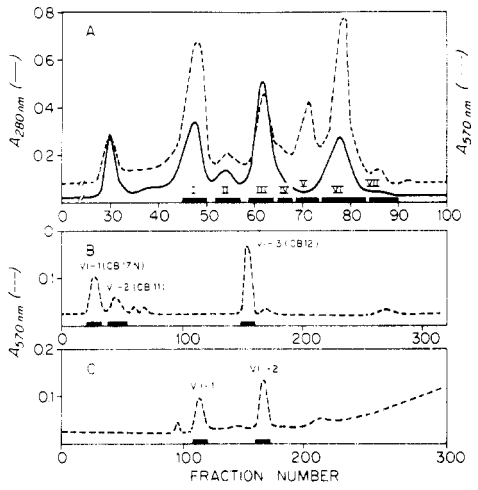

FIGURE 5: Separation of CNBr fragments of segment L_s (70 mg). Columns were monitored at 280 nm (solid line) or with ninhydrin after alkaline hydrolysis of 10% of the effluent (broken line) and fractions pooled as indicated by horizontal bars. (A) Initial separation on a Sephadex G-50 superfine column (2.5 × 115 cm) eluted with 9% formic acid at 20 mL/h. Aliquots of 5 mL were collected. (B) Further separation of the acid-soluble portion of fraction VI on a DEAE-Sephadex A-25 column (0.9 × 100 cm) at 39 °C. Peptides were eluted at 23 mL/h with a three-chambered linear gradient of (i) 0.2 M pyridine-acetate pH 6.05, (ii) 2.0 M, pH 6.1, and (iii) 2.0 M, pH 5.05 (150 mL each), followed by a two-chambered gradient of 2.0 M, pH 5.05, and 50% acetic acid (50 mL each). Aliquots of 2 mL were collected. (C) Purification of fraction VII on an AG 50W-X2 column (0.9 × 100 cm) at 50 °C. Peptides were eluted at 23 mL/h with a linear gradient of 0.2 M pyridine-acetate, pH 3.1, and 2.0 M, pH 5.0 (200 mL each), followed by a second linear gradient of 2.0 M, pH 5.0, and 8.0 M pyridine (100 mL each). Aliquots of 2 mL were collected.

CB17N, CB21, and CB14C. Their amino acid compositions and the extent of successful amino-terminal analysis of each are given in Tables III and IV. The sum of their amino acid compositions is in good agreement with that of the intact light segment L_s (Table II).

In addition to these eight major CB fragments, rechromatography of fractions II and IV (not shown) yielded two minor fragments, each homogeneous by Sequencer analysis. The fragment from fraction II had the same amino-terminal sequence as CB21, but its composition was essentially identical with the sum of those of CB21 and CB20 (Table III), indicating that this fragment (CB21-20) had resulted from incomplete cleavage between CB21 and CB20. Similarly, fraction IV appeared to overlap CB11 and CB17N (CB11-17N in Table III).

Amino Acid Sequence of CB14 (Residues 1-91). This cyanogen bromide fragment contains 91 amino acid residues and is located at the amino terminus of phosphorylase. Titani et al. (1975) established the sequence of the first 18 residues and showed that the amino terminus is blocked by an acetyl group. The remainder of the sequence is derived by sequence analysis of the amino terminus of segment L_s, by subfragmentation of CB14C and by enzymatic digestion of CB14 itself. The strategy is summarized in Figure 6.

Sequencer analysis of segment L_s demonstrated an 80/20 mixture of two related fragments (Figure 7). The difference in their concentrations permitted allocation of the products of stepwise degradation to a major and a minor fragment, the minor fragment simply lacking the amino-terminal dipeptide of the major fragment. Apparently subtilisin cleaved primarily

TABLE IV: Automated Edman Degradations of Fragments Derived from Residues 1 through 349.

Fragment analyzed (residue no.)	Amount applied (nmol)	Techniques used: Degradation[a] Identification[b]	Continuous sequence proven	Identification tentative	Identification poor or lacking
L_s (17-264)	70	B GL	17-31	23,25,30	22,26,28,29
"	510	A GL	17-45	36,38,42	
CB14C-BNPS-3 (68-91)	500	B GL	68-76		
"	500	B GL	68-90		70,84,86,88,89
L_s-AC-1 (79-264)	135	B GL	79-95	87,95	86,88,90,94
CB5 (92-99)	230	B HP	92-97		
L_s-T-1 (94-138)	575	B GL	94-124		121
CB12 (100-119)	200	B GL	100-106		
CB16 (120-147)	450	B GL	120-135	130,131,132,134,135	
"	470	B GL	120-144	142,143,144	140
L_s-T-2 (139-160)	200	B GL	139-154	146,149	
CB21 (148-176)	800	B GL	148-175	173,174,175	
"	400	B GL	148-161		149,160
L_s-T-3 (170-184)	300	B GL	170-183	180	181,182
CB20 (177-224)	1650	A GL	177-203		
"	1625	B GL	177-216	211,214,215	210
CB20-TS-α (206-224)	250	B GL	206-214	208	
"	100	C GL,HP	206-220		206,215,217,218
L_s-T-4 (215-234)	300	B GL	215-230		228
CB11 (225-241)	200	B GL	225-238		
CB11-17N (225-264)	360	B GL	225-246	240,245	228
L_s-T-5 (235-242)	130	C GL	235-241	240	235,236
CB17 (242-349)	800	A GL	242-272	255,258,269,270	
CB17N (242-264)	200	B GL	242-263	263	
H_s (265-841)	95	A GL	265-275	267,268	269,273
CB17C (265-349)	500	A GL	265-288	276	281
"	400	A GL,HP	265-292		283,286,289,291
CB17-T-6 (278-289)	150	C HP	278-288		278
CB17-T-7 (290-292)	200	B* GL,HP	290-292		
CB17-SP-3 (291-296)	70	B* HP	291-296		
CB17-T-9 (293-308)	125	B GL,HP	293-306	303,304	
CB17-P-1 (307-315)	50	B GL,HP	307-312		
CB17-TS-6 (310-322)	200	B GL,HP	310-321	314,319,320,321	312,313,317
CB17-T-12 (315-322)	110	B* HP	315-321		317,319
CB17-AC-γ (320-349)	250	B GL,HP	320-337	331	335
CB17-T-14 (323-349)	112	C HP	323-349		323,331,335,340,342
"	200	B* HP	323-329		
CB17-T-15 (332-349)	200	C GL,HP	332-349	342	332
"	200	B GL,HP	332-348		342

[a] Automated Edman degradation by (A) liquid phase (Sequencer) protein program, (B) liquid phase (Sequencer) peptide program where an asterisk indicates that polybrene was included as an inert carrier, or (C) solid phase (Sequemat). [b] Identification by gas–liquid chromatography (GL) and spot tests for histidine and arginine or by high performance liquid chromatography (HP).

on the amino side of glycine to expose the amino-terminal sequence Gly-Leu-Ala-Gly-Val- and to a lesser extent on the amino side of alanine to generate the minor sequence Ala-Gly-Val-. This analysis yielded positive identification for 26 of 29 degradation cycles of the major fragment (Figure 7 and Table IV), which overlaps the amino-terminal 18-residue sequence reported by Titani et al. (1975) as indicated in Figure 6. Thus, the amino-terminal sequence can be tentatively extended to residue 45.

The sequence of the carboxyl-terminal region of CB14 was partly elucidated by analysis of a fragment (BNPS-3), generated by cleavage of CB14C (2 μmol) at its single tryptophanyl residue. Gel filtration of the cleavage products yielded BNPS-3 in fraction III of Figure 8A. Duplicate Edman degradations established the sequence of most of its 24 residues (Figure 6). An Asp-Pro sequence was found in the middle of this peptide, providing a useful locus for mild acid cleavage. This was accomplished using segment L_s which was first succinylated to block the α-amino group. The products of acid cleavage were fractionated (Figure 8B) and a fraction was found (AC-1) which yielded a single sequence corresponding to the interior of BNPS-3 and continuing through the sequence Phe-X-Met-Gly-X-Leu-. Thus, the information from AC-1 and BNPS-3 provided not only a tentative sequence of the

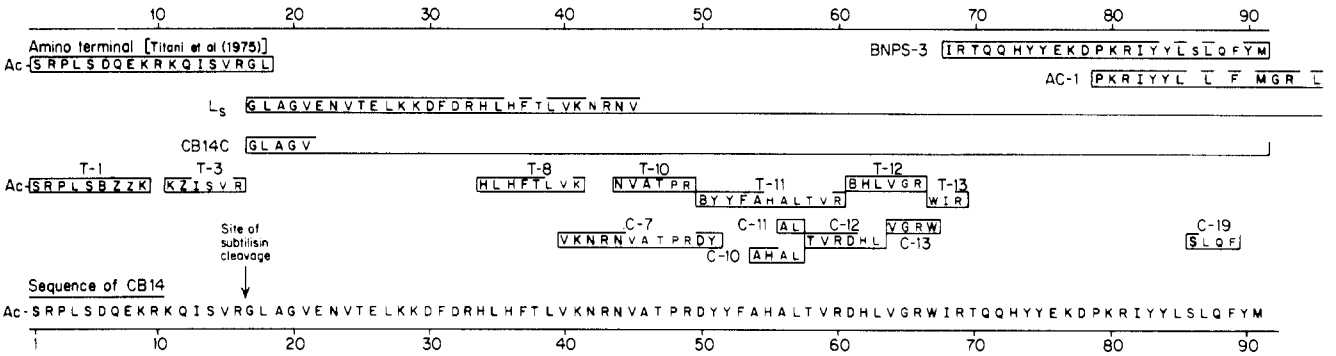

FIGURE 6: Summary proof of the sequence of fragment CB14. The one-letter code within the bars designates amino acid residues in that peptide. Large capital letters indicate amino acids identified after Edman degradation or by carboxypeptidase digestion. Smaller capital letters indicate placement by composition and cleavage specificity. The length of each bar indicates the length of the peptide (except segment L_s and fragment AC-1, which both continue to residue 264); enclosure of the top of the bar indicates that that portion of the sequence is proven; gaps in the upper enclosure indicate portions of sequence not identified. Methionine was identified as homoserine. The upper portion of the diagram summarizes the minimum proof of sequence. The lower portion summarizes the derived sequence. One-letter amino acid abbreviations are: A (alanine), B (aspartic acid or asparagine), 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), Y (tyrosine), and Z (glutamic acid or glutamine). Analysis of fragment AC-1 provides information extending beyond CB14 to a contiguous fragment (CB5 in Figure 13).

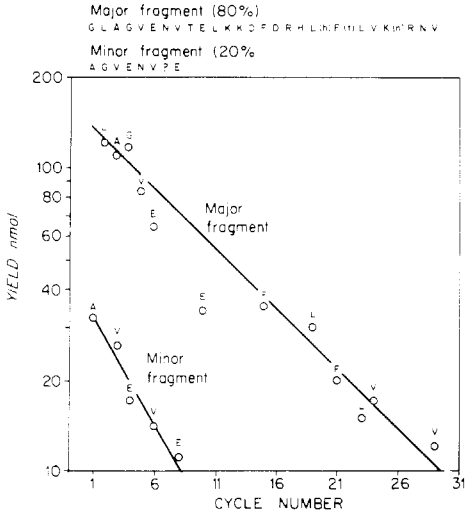

FIGURE 7: Yields of phenylthiohydantoins during Edman degradation of segment L_s in the sequencer. The data fall into two groups indicating that a major and a minor fragment were being degraded simultaneously. The minor sequence could not be recognized after eight cycles.

carboxyl-terminal residues 68–91 (Figure 6), but also an overlap into the contiguous cyanogen bromide fragment which later will be shown to be CB5.

The sequence of residues 34–69 was established by analyzing sets of overlapping peptides isolated from a tryptic and a chymotryptic digest of CB14 (Table V and Figure 9). As illustrated in Figure 6, this middle portion of CB14 included 5 tryptic peptides (T-8, T-10, T-11, T-12, and T-13) and 5 chymotryptic peptides (C-7, C-10, C-11, C-12, and C-13). Other peptides isolated from these digests corresponded in composition to those expected from established portions of the sequence (Figure 6), but are not necessary to prove the sequence. Two of these peptides are shown in Figure 6 to correspond to the amino-terminal portion of CB14, as reported by Titani et al. (1975). Another peptide (C-19), which was neutral in charge, completes the carboxyl-terminal 24-residue sequence by confirming serine-86 and glutamine-88.

The composition of the carboxyl-terminal portion was confirmed by amino acid analysis of a fragment (CB14-AC-α), cleaved from CB14 at Asp-Pro (78–79) by dilute acid and isolated by gel filtration (Figure 8C).

The sequence of CB14 thus derived is in good agreement with its composition (Table III), lending support to the alignments and overlaps illustrated in Figure 6.

Amino Acid Sequence of CB17 (Residues 242–349). This portion of the sequence of the enzyme was determined by analysis of CB17, isolated from the whole molecule, CB17N from segment L_s and CB17C from segment H_s (the isolation of CB17C is described by Titani et al., 1978). In addition subpeptides were generated by treatment with mild acid and by digestion with various proteolytic enzymes. The specific strategy of the sequence determination is illustrated in Figure 10. In the following presentation, the residue numbers refer to the 108 residues in fragment CB17.

Sequencer analysis of CB17 (10 mg; 0.8 μmol) yielded the amino-terminal sequence of 31 residues except for tentative identifications in cycles 14, 17, 28, and 29 (Table IV). Analysis of CB17N (200 nmol) yielded the entire sequence of this fragment except for the carboxyl terminus, glutamine-23, thus establishing the 14th and 17th residues (Figure 10) as lysine and asparagine, respectively (Table IV).

The combined results of duplicate analyses of CB17C (500 and 400 nmol each) on the Sequencer indicated that subtilisin cleaves at one major site and two minor sites. A similar ambiguity in cleavage site had been seen at the amino terminus of segment L_s. The major sequence observed in CB17C (80% of the degradation products) began with the sequence Ala-Val-Leu-Asp-Arg, which corresponds to residues 24–28 of CB17. Two minor sequences (5–15% level) were also distinguished by techniques analogous to those used above for segment L_s in Figure 7. One, beginning with the sequence Leu-Asp-Arg-, is simply 2 residues shorter than the major component. The other, beginning with Ile-Gln-Ala-Val-Leu-, is 2 residues longer than the major component. These analyses were facilitated by allowing the penultimate glutamine of the second minor component to cyclize to the pyroglutamyl derivative by interrupting the degradation cycle for 16 h *after* removal of the amino-terminal isoleucine. Under these conditions only the major component and the minor component Leu-Asp-Arg- continued to degrade, and the second minor component could be identified by comparison with the original mixture.

Since the amino-terminal sequence of CB17 overlaps the sequence of CB17N with that of the major component of CB17C, an uninterrupted sequence of the first 41 residues of

TABLE V: Amino Acid Compositions[a] of Peptides Obtained from CB14.

	Tryptic Digest (Fig. 9-A)													Chymotryptic Digest (Fig. 9-B)									
	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9	T-10	T-11	T-12	T-13[c]	T-14	T-15	C-7	C-10	C-11	C-12	C-13	C-19	CB14-AC-α[d]	BNPS-3[e]
Residue Number (Fig. 14)	1-9	1-10	11-16	12-16	17-28	17-29	30-33	34-41	42-43	44-49	50-60	61-66	67-69	70-78	78-80	40-51	54-57	56-57	58-63	64-67	86-89	79-91	68-91
Fig. No. Pool No.	Fig. 9-A IV	VIII	VIII	II	III	VI	XV	XVIII	XII	VII	XIII	XIII		XIII	VIII	Fig. 9-B XVI	X	VIII	XVI	XXII	VI	Fig. 8-C III	Fig. 8-A III
Purification Procedure[b]		AG 1, HVE 37	AG 1, HVE 37				HVE 3.7		HVE 3.7	AG 1	AG 1	AG 1			AG 1, HVE 3.7	AG 1, HVE 3.7	AG 1,	AG 1	AG 1, HVE 3.7		AG 1		
Amino Acids																							
Asx	1.13 (1)	1.14 (1)			1.27 (1)	1.16 (1)	2.09 (2)		1.07 (1)	1.12 (1)	0.98 (1)	1.07 (1)		1.00 (1)	1.00 (1)	3.27 (3)			1.16 (1)			0.29 (0)	1.18 (1)
Thr					0.93 (1)	0.99 (1)		1.04 (1)		0.98 (1)	0.90 (1)			0.95 (1)		1.23 (1)			0.97 (1)				0.92 (1)
Ser	1.76 (2)	2.12 (2)	0.84 (1)	0.90 (1)	0.24 (0)	0.10 (0)															0.99 (1)	1.10 (1)	0.74 (1)
Glx	1.99 (2)	2.28 (2)	1.15 (1)	1.21 (1)	2.15 (2)	1.86 (2)					0.21 (0)			2.58 (3)							1.05 (1)	0.98 (1)	4.40 (4)
Pro	1.14 (1)	0.88 (1)		0.12 (0)						1.12 (1)					1.02 (1)	1.10 (1)						1.02 (1)	1.25 (1)
Gly				0.24 (0)	1.87 (2)	1.86 (2)						1.03 (1)								0.90 (1)			0.14 (0)
Ala					1.00 (1)	1.00 (1)				1.02 (1)	1.97 (2)					1.00 (1)	1.85 (2)	0.97 (1)					0.25 (0)
Val			1.00 (1)	1.00 (1)	1.87 (2)	1.86 (2)		1.03 (1)		1.00 (1)	0.86 (1)	0.93 (1)				2.22 (2)			1.00 (1)	1.00 (1)		0.34 (0)	0.18 (0)
Hse																						0.74 (1)	0.63 (1)
Ile			1.08 (1)	0.92 (1)	0.13 (0)	0.13 (0)							1.00 (1)									1.03 (1)	1.85 (2)
Leu	1.00 (1)	1.00 (1)			1.95 (2)	1.76 (2)		2.00 (2)			1.00 (1)	1.00 (1)				0.36 (0)	1.00 (1)	1.00 (1)	0.91 (1)		1.00 (1)	2.00 (2)	2.00 (2)
Tyr											2.06 (2)			1.99 (2)		1.04 (1)						2.68 (3)	2.30 (5)
Phe							1.00 (1)	0.87 (1)			1.09 (1)										0.96 (1)	0.95 (1)	1.15 (1)
His								1.88 (2)			1.06 (1)	0.94 (1)		1.06 (1)		0.25 (0)	0.94 (1)		1.00 (1)				0.75 (1)
Lys	0.99 (1)	1.24 (1)	1.02 (1)	0.24 (0)	0.97 (1)	1.76 (2)		0.90 (1)						0.90 (1)	0.94 (1)	0.90 (1)						1.04 (1)	2.05 (2)
Arg	0.97 (1)	2.00 (2)	1.08 (1)	1.01 (1)			1.01 (1)		1.00 (1)	1.05 (1)	1.05 (1)	1.07 (1)	1.00 (1)			2.25 (2)			0.84 (1)	1.24 (1)		0.72 (1)	1.72 (2)
Trp													N.D. (1)							0.75[f] (1)			
Yield %	68	3	4	38	30	53	14	40	23	48	23	23	66	12	2	8	8	18	6	19	15	13	15

[a] Residues/molecule by amino acid analysis or (in parentheses) from the sequence (Figure 14). [b] AG 1 and HVE 3.7 denote chromatography on an AG 1X2 column and high-voltage paper electrophoresis at pH 3.7, respectively. [c] T-13 was missed from the first digest but isolated from a replicate digest by gel filtration on Sephadex G-25 superfine (2.5 × 200 cm) in 9% formic acid. [d] Derived by treatment of CB14 with mild acid (pH 2.5) (Figure 8C). [e] Derived by treatment of CB14C with BNPS-skatole (Figure 8A). [f] Residues/molecule released by carboxypeptidase A.

TABLE VI: Amino Acid Compositions[a] of Peptides Obtained from CB17.

	Tryptic Digest (Fig. 11-A)															Tryptic Digest of Succinylated CB17 (Fig. 11-B)				Other Digests				
	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9	T-10	T-11	T-12	T-13	T-14	T-15	TS-4	TS-5	TS-6	TS-7	AC-β[e]	AC-γ[e]	SP-2[f]	SP-3[f]	P-1[g]
Residue Number (Fig. 14)	242 & 309	243-247	248-255	256-269	270-277	278-289	290-292	293-294	293-308	310-311	312-314	315-322	323-331	323-349	332-349	278-292	293-308	310-322	323-349	242-319	320-349	288-290	291-296	307-315
Fig. No. Pool No.	Fig. 11-A VIII	VIII	V	V	V	V	VII	VI	III	VII	VI	V	V	I	II	Fig. 11-B III	III	III	I	Fig. 11-D II	Fig. 11-D III	Fig. 11-C VI	Fig. 11-C IV	
Purification Procedure[b]	HVE 3.7	HVE 3.7	AG 50, HVE 3.7	[c]	AG 50, HVE 3.7	AG 50, HVE 3.7	HVE 3.7	HVE 3.7		HVE 3.7	HVE 3.7	AG 50, HVE 3.7	AG 50, HVE 3.7			AG 50	AG 50	AG 50				HVE 3.7	HVE 3.7	HVE 6.5
Amino Acids																								
Asx			3.47 (3)	3.32 (3)	2.00 (2)	3.06 (3)			1.11 (1)			1.04 (1)	2.81 (3)	4.46 (5)	2.10 (2)	2.82 (3)	1.08 (1)	1.06 (1)	4.91 (5)	12.39 (13)	4.91 (5)			
Thr				0.10 (0)		0.12 (0)			1.01 (1)				0.61 (1)	2.00 (2)	1.22 (1)		0.81 (1)		1.82 (2)	1.31 (1)	1.82 (2)			
Ser		0.95 (1)		0.26 (0)	0.98 (1)	0.16 (0)			0.31 (0)	0.13 (0)	1.60 (2)		0.13 (0)	1.18 (1)	1.30 (1)	0.54 (0)	0.27 (0)	1.74 (2)	1.36 (1)	3.60 (4)	1.14 (1)	0.16 (0)	0.26 (0)	2.04 (2)
Glx				1.21 (1)	1.21 (1)	1.07 (1)	0.89 (1)		2.81 (3)				0.35 (0)	2.19 (2)	2.15 (2)	2.20 (2)	2.91 (3)	0.47 (0)	2.24 (2)	7.23 (7)	2.08 (2)	1.05 (1)	2.13 (2)	0.18 (0)
Pro			1.02 (1)	0.14 (0)		1.10 (1)			0.10 (0)			1.13 (1)	1.00 (1)	2.86 (3)	2.03 (2)	1.04 (1)	0.38 (0)	1.10 (1)	2.78 (3)	2.18 (2)	3.14 (4)			
Gly				1.96 (2)		0.99 (1)			0.32 (0)			1.00 (1)	0.32 (0)	0.37 (0)	0.12 (0)	0.87 (1)	0.56 (0)	1.33 (1)	0.50 (0)	4.00 (4)	0.35 (0)	1.00 (1)	0.19 (0)	0.48 (0)
Ala		1.04 (1)	0.99 (1)	1.08 (1)	1.00 (1)	0.22 (0)			1.93 (2)				1.00 (1)	2.87 (3)	2.16 (2)	0.55 (0)	2.00 (2)	0.23 (0)	3.00 (3)	6.00 (6)	3.00 (3)			0.32 (0)
Cys[d]												0.60 (1)						0.68 (1)		0.35 (1)				
Val				1.85 (2)		0.90 (1)			1.38 (2)			1.13 (1)	0.30 (0)	1.12 (1)	1.34 (1)	0.94 (1)	1.65 (2)	1.00 (1)	0.85 (1)	4.72 (5)	1.98 (2)			
Hse														0.60 (1)	0.54 (1)				0.53 (1)		0.59 (1)			
Ile				0.96 (1)	1.09 (1)				0.95 (1)				0.13 (0)	2.01 (2)	2.13 (2)	0.51 (0)	1.11 (1)	0.15 (0)	1.90 (2)	3.24 (3)	2.00 (2)			1.21 (1)
Leu		1.00 (1)	1.00 (1)	1.16 (1)	1.09 (1)	1.00 (1)	1.00 (1)	1.00 (1)	2.00 (2)	0.23 (0)			0.53 (0)	3.00 (3)	3.00 (3)	2.00 (2)	2.18 (2)	0.68 (0)	2.97 (3)	7.95 (8)	3.18 (3)		1.96 (2)	0.16 (0)
Tyr				0.77 (1)		0.86 (1)			0.68 (1)				0.17 (0)			0.67 (1)	0.79 (1)			2.77 (3)				
Phe			0.96 (1)	1.20 (1)		2.00 (2)			0.89 (1)	1.00 (1)		0.90 (1)	1.83 (2)	1.88 (2)	0.19 (0)	1.69 (2)	1.28 (1)	1.97 (2)	1.68 (2)	6.92 (7)	1.96 (2)			1.91 (2)
His									0.11 (0)					0.97 (1)	1.00 (1)				0.99 (1)		1.00 (1)			
Lys		1.06 (1)	0.95 (1)	0.23 (0)		1.04 (1)	0.11 (0)	1.14 (1)	1.04 (1)	1.11 (1)	1.00 (1)		1.00 (1)	1.03 (1)	0.30 (0)	0.82 (1)	1.03 (1)	1.88 (2)	0.92 (1)	5.76 (6)	1.26 (1)	0.96 (1)	1.03 (1)	2.00 (2)
Arg	1.00 (1)			1.00 (1)	1.07 (1)		1.03 (1)		0.94 (1)	0.25 (0)		2.00 (2)	0.19 (0)			1.47 (1)	1.13 (1)	2.37 (2)		6.90 (7)	1.14 (1)		1.00 (1)	2.06 (2)
Trp		N.D. (1)																		N.D. (1)				
Yield %	30	18	10	29	10	5	30	17	4	17	30	15	10	42	27	30	13	20	60	47	63	10	8	30

[a] Residues/molecule by amino acid analysis or (in parentheses) from the sequence (Figure 14). [b] AG 50, HVE 3.7 and 6.5 denote chromatography on an AG-50WX2 column and high voltage paper electrophoresis at pH 3.7 and pH 6.5. [c] Pool V contained one peptide (T-4) insoluble at pH 3.1. Others were separated from the soluble fraction as indicated. [d] Estimated as *S*-carboxymethylcysteine. [e] Derived by treatment of CB17 with mild acid (pH 2.5) (Figure 11D). [f] Derived from a digest of CB17C with staphylococcal protease. [g] Derived from a peptic digest of CB17C. The most basic peptide was isolated by HVE 6.5.

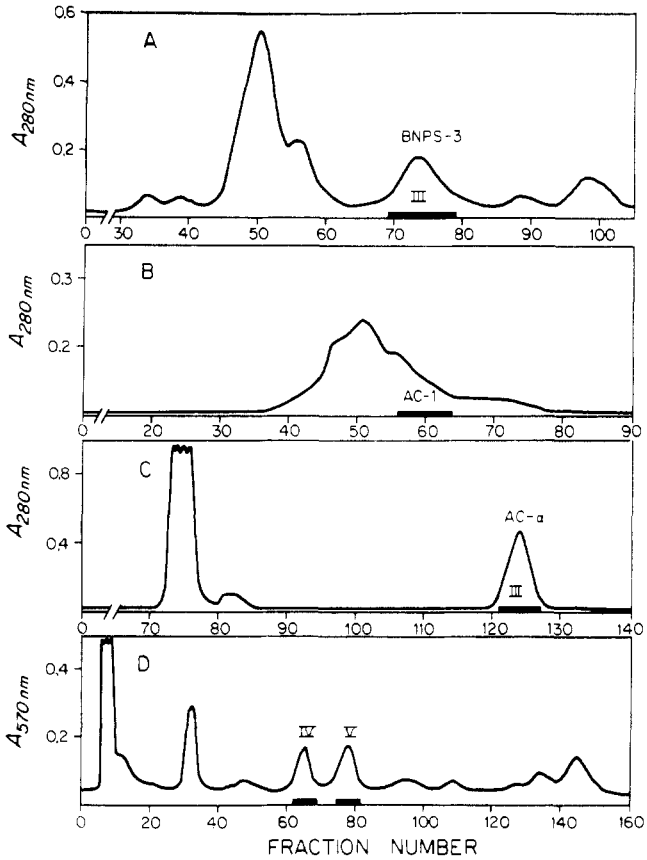

FIGURE 8: (A) Separation of a fragment (BNPS-3) generated from CB14C (17 mg) by cleavage at a tryptophanyl bond. A column (1.5 × 90 cm) of Sephadex G-50 superfine was eluted with 9% formic acid at 11 mL/h and monitored at 280 nm. Fractions of 2 mL were collected and pooled as indicated by horizontal bars. (B) Separation of fragments generated from succinylated segment L_5 (30 mg) by cleavage at aspartylproline bonds. The reaction mixture was applied to a column (2.5 × 110 cm) of Sephadex G-75 fine and eluted with 9% formic acid containing 7 M urea at 22 mL/h. Fractions of 5 mL were collected. (C) Fragments generated from CB14 (50 mg), by cleavage of an aspartylproline bond, were separated on a column (2.5 × 200 cm) of Sephadex G-25 Superfine in 9% formic acid at 20 mL/h. Aliquots of 5 mL were collected. (D) Separation of the products of digestion of succinylated CB20 (1 μmol) on a column (0.9 × 50 cm) of DEAE-Sephadex A-25 at 39 °C. Peptides were eluted at 23 mL/h by a three-chambered gradient of (i) 0.2 M pyridine-acetate, pH 6.05, (ii) 2.0 M, pH 6.1, and (iii) 2.0 M, pH 5.05 (150 mL each), and the column was stripped with 50% acetic acid (100 mL). Column effluent (13%) was monitored by the ninhydrin reaction after alkaline hydrolysis. Fractions of 2 mL were collected.

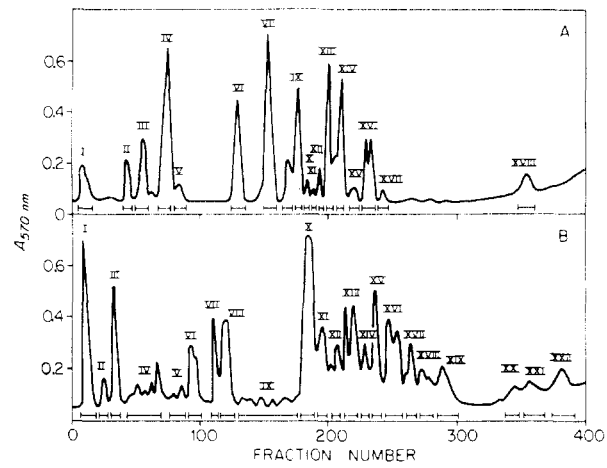

FIGURE 9: Separation of tryptic (A) and chymotryptic (B) peptides from 5-μmol digests of CB14 on a column (2.0 × 20 cm) of AG 50W-X8 at 50 °C. Peptides were eluted at 80 mL/h by three successive linear gradients: (i) 0.05 M pyridine-acetate, pH 2.4, to 0.5 M, pH 3.75 (each 500 mL), (ii) 0.5 M, pH 3.75, to 2.0 M, pH 5.0 (each 500 mL), and (iii) 2.0 M, pH 5.0, to 8.0 M pyridine (each 200 mL). The column was monitored by ninhydrine after alkaline hydrolysis. Fractions of 6 mL were collected and pooled as indicated by the horizontal bars.

CB17 was obtained (Figure 10), and six of the next 10 residues are tentatively identified.

The remainder of the sequence of CB17 was determined by analyzing sets of overlapping peptides generated by peptic digestion of CB17, by tryptic digestion (before and after succinylation) of CB17, and by digestion of CB17C with staphylococcal protease. Twenty-seven peptides were purified, of which 22 are indicated in Table VI. Thirteen of these provided data to complete the sequence of CB17 (Figure 10); the remainder yielded data in complete accord with that sequence.

The carboxyl-terminal sequence was determined by two analyses of T-15, one on the Sequencer (200 nmol), and the other on the Sequemat (200 nmol) after attachment of the terminal homoserine to TETA resin. The combined data yielded the complete sequence from residues 91 to 108. This sequence was extended back to threonine-82 by analysis of T-14 and by the observation that its composition is equal to the sum of those of peptides T-13 and T-15 (Table VI). Thus, peptide T-13 is contiguous with the carboxyl-terminal peptide T-15.

This sequence was partly confirmed and extended 3 residues by analysis of a fragment (CB17-AC-γ) generated from CB17 (0.8 μmol) by mild acid cleavage at pH 2.5. Separation of the products (Figure 11D) yielded a fraction which contained 30 residues (Table VI) and the amino-terminal sequence Pro-Val-Arg-Thr-Asn-Phe, which overlaps peptide T-14 and places aspartic acid at position 89. Thus it appears that an Asp-Pro dipeptide at positions 78–79 is susceptible to acid cleavage.

To complete the sequence of residues 42–78, it is helpful to consider the location of the 5 isoleucine residues in CB17. Four of them have been placed at residues 22, 34, 93, and 104. The fifth served as a unique marker of an unresolved portion of the fragment. It is found in peptides T-9, TS-5, and P-1. The two tryptic peptides are identical, ending in the sequence Gln-Asp-Ile-Arg. Peptide P-1 begins with Ile-Arg-Arg-Phe-Lys-Ser and overlaps the carboxyl terminus of T-9. The carboxyl-terminus of P-1 was identified by digestion with carboxypeptidase Y (2 h digest, 37 °C, 0.1 M pyridine-acetate, pH 6; yield: 1.0 Phe, 0.55 Lys, 1.25 Ser). Peptide TS-6 corresponds to the carboxyl terminus of P-1. Although the complete sequence of TS-6 was not determined (Figure 10), the partial sequence is sufficient to establish the order of peptide T-10, T-11, and T-12 in the sequence Phe-Lys-Ser-Ser-Lys-Phe-Gly-X-Arg. Analysis of T-12 provided the partial sequence Phe-Gly-X-Arg-X-Pro-Val-Arg, where the two unidentified residues, (X), are cysteine and either aspartic acid or asparagine. Since an Asp-Pro bond was already indicated in a sequence Asp-Pro-Val-Arg by the generation of CB17-AC-γ (*vide supra*), and since the Arg-Asp bond would resist tryptic attack of peptide T-12, it is concluded that T-12 and T-14 are overlapped by CB17-AC-γ. This leaves cysteine at position 76, in accord with the Gly-Cys-Arg-Asp peptide reported by Zarkadas et al. (1968). Thus, a continuous sequence is established from leucine-52 to methionine-108 and it only remains to complete the sequence from residues 42 to 51 in CB17.

Peptide T-6 (150 nmol) was analyzed on the Sequemat after attachment of the carboxyl-terminal lysine to DITC-3-aminopropyl glass (Table IV). Although the amino- and carboxyl-terminal amino acids are not identified by this procedure, the remainder of the sequence was determined and lysine

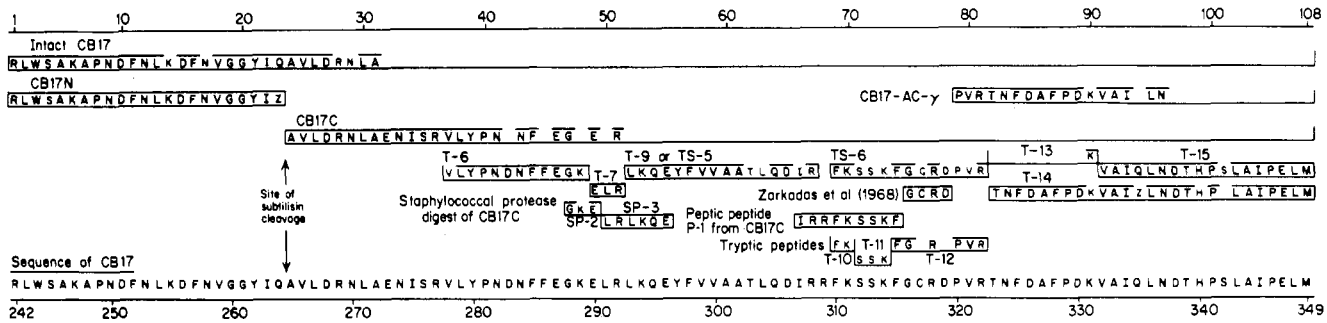

FIGURE 10: Summary proof of the sequence of fragment CB17 (108 residues), using the same annotations as in Figure 6. For simplicity, only those peptides crucial to the proof are illustrated. The upper portion of the diagram summarizes the simplest proof of sequence. The derived sequence is shown at the bottom with the residue numbers of the whole molecule. Subtilisin cleavage separates segment L_s (residues 17–264) from segment H_s (residues 265–841).

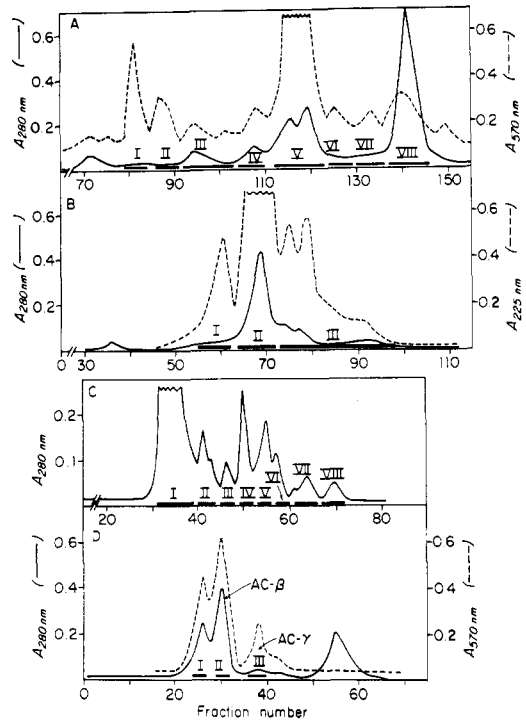

FIGURE 11: (A) Separation of tryptic peptides from CB17 (3.2 μmol) on a column (2.5 × 200 cm) of Sephadex G-25 superfine with 9% formic acid at 24 mL/h. The column was monitored at 280 nm (solid line) and by ninhydrin after alkaline hydrolysis (broken line). Fractions of 6 mL were collected and pooled as indicated by horizontal bars. (B) Separation of tryptic peptides from succinylated CB17 (2 μmol) on a column (2.5 × 110 cm) of Sephadex G-50 superfine with 0.1 M NH_4HCO_3, pH 8.0, at 20 mL/h. The column was monitored at 280 nm (solid line) and 225 nm (broken line). Fractions of 5 mL were collected. (C) Separation of a digest of CB17C (1 μmol) by staphylococcal protease on a column (1.5 × 85 cm) of Sephadex G-25 Superfine with 9% formic acid at 8.6 mL/h. The digest was first lyophilized and resuspended in 3 mL of 9% formic acid; insoluble peptides were discarded after centrifugation, and the supernatant was applied to the column. Fractions of 2 mL were collected and pooled as indicated. (D) Fragment generated by cleavage of CB17 (160 nmol) at an aspartylproline bond was separated on a column (1.5 × 85 cm) of Sephadex G-50 Superfine with 9% formic acid at 10 mL/h. The column was monitored at 280 nm (solid line) and by ninhydrin after alkaline hydrolysis (broken line). Fractions of 2 mL were collected and pooled as indicated.

placed on the basis of tryptic specificity. The resulting sequence coincides with the partial sequence of residues 37–48 in CB17C (Figure 10) and extends that structure to residue 48, ending with Phe-Glu-Gly-Lys. All 4 glycyl residues are now placed (19, 20, 47, and 75) and peptide SP-2 can be aligned only with glycine-47 to yield Gly-Lys-Glu-49.

The critical overlap from glutamic acid-49 to leucine-52 is provided by two peptides, T-7 and SP-3, both of which contain arginine. Of the 8 arginyl residues in CB17, 7 have been placed in sequence (1, 28, 36, 67, 68, 76, 81) and one tentatively assigned to residue 51. Neither T-7 nor SP-3 can be derived from any of the established arginyl sequences and both are consistent with an arginine at residue 51. Assuming that trypsin cleaves at lysine-48 and arginine-51, T-7 extends the amino-terminal sequence to Gly-Lys-Glu-Leu-Arg-51 (Figure 10). Peptide SP-3 overlaps this sequence (Leu-Arg) and the amino terminus of T-9 (Leu-Lys-Gln-Glu), completing the sequence of fragment CB17.

Structural Analysis of the Interior of Segment L_s (Residues 92–241). Sequence analysis of CB14 and CB17 (vide supra) has established the amino-terminal 91 residues of phosphorylase (which overlaps segment L_s by 75 residues; cf. Figure 2), a short overlap (Gly-Arg-X-Leu) into residues 92–95, and the carboxyl-terminal sequence of segment L_s (residues 242–264 in CB17N). The remainder of segment L_s consists of the sequences represented by fragments CB5, CB11, CB12, CB16, CB20, and CB21, all of which were isolated from that segment. In this section is presented structural information derived from these six CB-fragments and from five methionine-containing peptides isolated from a tryptic digest of segment L_s. These two subsets of peptides are shown to form overlapping sequences accounting for residues 92–241 and completing the proof of sequence of segment L_s.

To generate the tryptic peptides, segment L_s (4 μmol) was digested with 0.075 μmol of trypsin for 3 h at pH 8.0 in a pH-stat and the digest separated into four fractions by gel filtration on Sephadex G-50 (Figure 12A). Amino acid analysis indicated the presence of methionine only in fractions II and III, which were separately refractionated on the same column (Figures 12B and 12C). Fraction II-1 yielded 0.575 μmol of a single component (L_s-T-1) containing 45 residues including 2 methionines (Table VII) and an amino-terminal sequence beginning with Thr-Leu-Gln-Asn. Sequenator analysis of this fragment placed 30 residues in sequence (Figure 13 and Table IV), ending with Met-Glu-X-Leu-Glu-Glu. Fraction III was more complex and subfractionation yielded four fractions (Figure 12C) of which only the first three contained methionine. Fraction III-1 was essentially identical with fraction II-1 (Figure 12B). Fractions III-2 and III-3 were further purified on DEAE-Sephadex (Figures 12D and 12E) and Ag1-X2 (Figure 12F). Four methionine-containing peptides were thus isolated and analyzed both for their amino acid compositions (Table VII) and for their amino-thermal sequences (Table IV). Each of these peptides contained a single methionine which was placed in sequence, providing information which is used below to overlap the cyanogen bromide fragments.

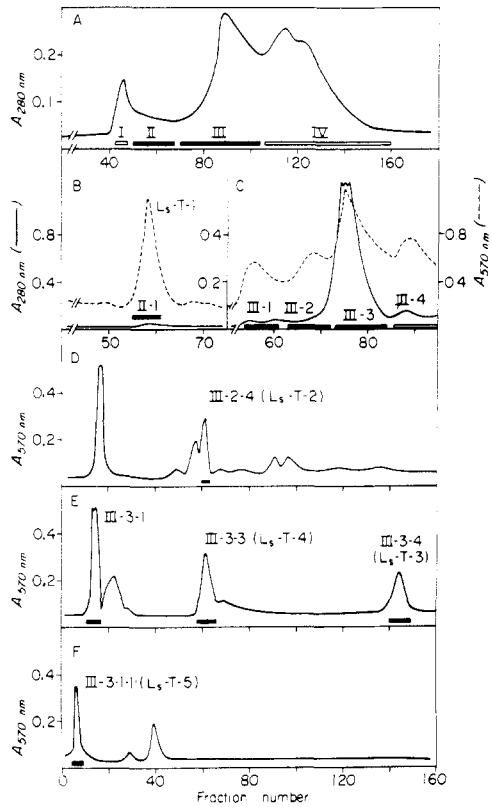

FIGURE 12: (A) Separation of tryptic peptides from segment L_s (120 mg) on a column (2.5 × 110 cm) of Sephadex G-50 Superfine in 0.1 M NH_4HCO_3 (pH 8.0) at 27 mL/h. Fractions of 4.5 mL were collected and pooled as indicated by horizontal bars. (B and C) Rechromatography of fractions II (20 mg) and III (51 mg) on the same column as that used in A. The column was monitored at 280 nm and by ninhydrin after alkaline hydrolysis. Fractions of 5 mL were collected and pooled as indicated. (D and E) Further separation of fractions III-2 and III-3 from C on a column (0.9 × 100 cm) of DEAE-Sephadex A-25 at 39 °C. Peptides were eluted at 23 mL/h by a three-chambered gradient of (i) 0.2 M pyridine-acetate, pH 6.05, (ii) 2.0 M, pH 6.10, and (iii) 2.0 M, pH 5.05 (150 mL each). Thirteen percent of each column effluent was monitored with ninhydrin after alkaline hydrolysis. Fractions of 2 mL were collected and pooled as indicated. (F) Separation of fraction III-3-1 from E on a column (0.9 × 50 cm) of AG 1-X2 at 35 °C. Peptides were eluted at 23 mL/h by a linear gradient of 3% pyridine and 0.5 M pyridine-acetate, pH 6.0 (200 mL each). The column was monitored as in E. Fractions of 2 mL were collected.

TABLE VII: Amino Acid Compositions of Tryptic Peptides, Containing Methionine, Obtained from Segment L_s and from CB20.

Peptide	L_s-T-1	L_s-T-2	L_s-T-3	L_s-T-4	L_s-T-5	CB20-TS-α
Residues	94–138	139–160	170–184	215–234	235–242	206–224
Figure	12-B	12-D	12-E	12-E	12-F	8-D
Fraction	II-1	III-2-4	III-3-4	III-3-3	III-3-1-1	IV and V
Amino acids						
Asx	7.3 (8)	1.9 (1)	2.0 (2)	2.2 (2)	1.9 (3)	1.2 (1)
Thr	2.8 (2)	1.6 (1)		2.3 (2)	1.0 (1)	1.7 (2)
Ser		1.2 (1)				0.9 (1)
Glx	10.1 (10)	0.8 (0)	3.1 (3)	1.2 (1)		2.9 (3)
Pro				2.7 (3)		
Gly	6.1 (6)	3.4 (3)	2.0 (2)	1.3 (1)		1.2 (1)
Ala	4.0 (4)	5.0 (5)	1.1 (1)	2.4 (1)		2.0 (2)
CmCys	1.4 (1)	1.0 (1)	0.6 (1)			
Val	0.6 (1)			3.1 (4)	1.8 (2)	3.5[a] (4)
Met	2.0 (2)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	0.7[b] (1)
Ile	1.1 (1)	1.3 (1)	1.0 (1)			
Leu	8.2 (8)	4.2 (4)	0.9 (1)	1.4 (1)		1.1 (1)
Tyr	1.0 (1)	2.0 (2)		2.0 (2)		
Phe		1.0 (1)				
His						0.9 (1)
Lys						1.0 (1)
Arg	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	
Trp[c]			(2)	(1)		(1)
No. of Residues	45	22	15	20	8	19
Yield (%)	41	47	28	38	12	35

[a] 96-h hydrolysis. [b] Determined as homoserine. [c] From sequence.

Of the six cyanogen bromide fragments (CB5, CB11, CB12, CB16, CB20, and CB21), only CB5 possessed an amino-terminal sequence which corresponded to the Gly-Arg-X-Leu overlap provided by analysis of fragment AC-1. Thus CB5 follows CB14 (Figure 2). The sequence of CB5 was reported by Cohen et al. (1973) to be Gly-Arg-Thr-Leu-(Gln, Asn, Thr)-Met, which was further resolved in this investigation to Gly-Arg-Thr-Leu-Gln-Asn-(Thr, Met), the last six residues overlapping the amino-terminal sequence of L_s-T-1 (Figure 13). Sequencer analysis of L_s-T-1 not only provided the remaining sequence of CB5 through methionine 99 in a sequence Thr-Met-Val-Asn-Leu-Ala, but also extended the sequence through a second methionine in the sequence Met-Glu-X-Leu-Glu-Glu. The sequence between these two methionyl residues corresponds in composition to CB12 which has an amino-terminal Val-Asn-Leu-Ala. The sequence after the second methionine corresponds to the amino terminus of CB16 (Glu-Glu-Leu-Glu-Glu). Thus, four cyanogen bromide fragments were placed in the order CB14-CB5-CB12-CB16 and the sequence was extended to glutamic acid-124.

Continued Sequencer analysis of CB16 extended the sequence from glutamic acid-124 through glycine-137 and continued with the partial sequence Arg-Leu-X-Ala-(Cys-Phe-Leu_{144}), the identification of the last 3 residues being tentative. This partial sequence from leucine-139 to leucine-144 corresponds to the amino-terminal sequence of L_s-T-2. Analysis of L_s-T-2 for 16 residues placed methionine as the ninth residue followed by Ala-Thr-Leu-Gly-Leu-Ala-Ala, which in turn overlaps the amino terminus of CB21. With analysis of 25 residues of CB21, the sequence was extended to glycine-172. Peptide L_s-T-3 overlapped this sequence by an Ile-Cys-Gly sequence (residues 170–172) and provided definitive data through methionine-176 and the sequence Glu-Glu-Ala-179. The latter three residues correspond to the amino-terminal sequence of a 48-residue fragment, CB20.

Extended duplicate analyses of fragment CB20 established the identity of the first 32 residues, and continued with the partial sequence -Thr-X-(Gln)-Gly-Ala-(Lys-Trp)-Val-216, where the three residues in parentheses were only tentatively identified. Since all three arginyl residues in CB20 had been unequivocally placed (residues 184, 193, and 205), fragment CB20 (1.0 μmol) was succinylated and cleaved by trypsin to generate a peptide representing residues 206–224. This peptide (TS-α) was isolated in both fractions IV and V of Figure 8D. Its composition (Table VII) agrees with the difference between the composition of CB20 and the sequence of residues 177–205. The combined results of Sequencer and Sequemat degradation (Table IV) yielded the sequence from valine-206 to lysine-214 and placed Gln-Val at residues 219 and 220. Thus, the sequence of CB20 was established from its amino terminus (glutamic acid-177) to lysine-214; tryptophan is tentatively identified at residue 215 and followed by the sequence Val-X-X-Gln-Val-220. This partial sequence was recognized at the amino terminus of the tryptic peptide L_s-T-4, i.e., Trp-Val-Asp-Thr-Gln-Val-Val-Leu-Ala-Met-Pro-Tyr-Asp. The first 10 residues of the sequence represent the last 10 residues of CB20 as indicated both by the partial sequence overlap and the agreement between the observed composition of CB20 and

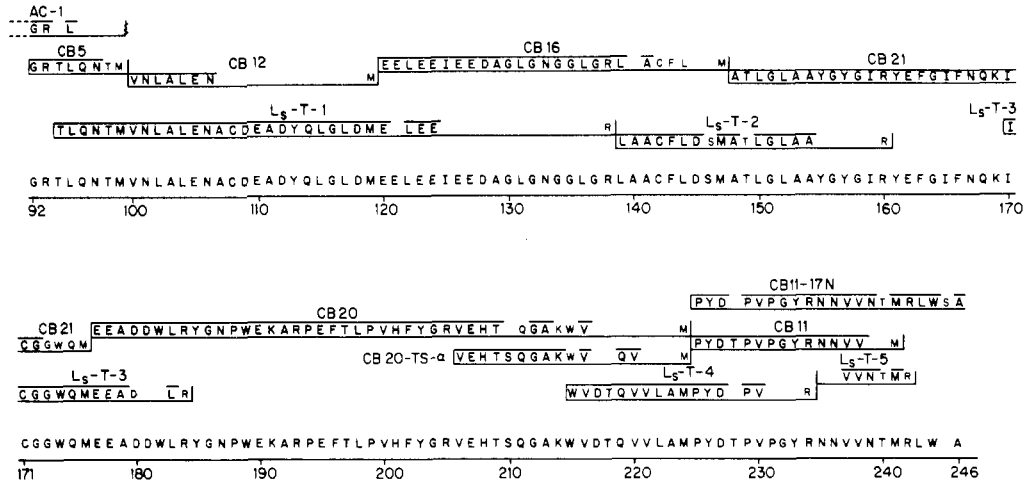

FIGURE 13: Summary proof of the sequence of residues 92 to 246, using the same annotations as in Figure 6. Part of the analysis of fragment AC-1 (residues 92–95) is included to demonstrate the contiguity of CB14 (Figure 6) and CB5. Residues 242–246, as placed in fragment CB11-17N, overlap the amino terminus of CB17 (Figure 10).

```
 1                  10          *  ↓    20
Ac-S-R-P-L-S-D-Q-E-K-R-K-Q-I-S-V-R-G-L-A-G-V-E-N-V-T-
      30                  40                  50
-E-L-K-K-D-F-D-R-H-L-H-F-T-L-V-K-N-R-N-V-A-T-P-R-D-
                  60                  70
-Y-Y-F-A-H-A-L-T-V-R-D-H-L-V-G-R-W-I-R-T-Q-Q-H-Y-Y-
      80                  90                  100
-E-K-D-P-K-R-I-Y-Y-L-S-L-Q-F-Y-M-G-R-T-L-Q-N-T-M-V-
            *  110                 120
-N-L-A-L-E-N-A-C-D-E-A-D-Y-Q-L-G-L-D-M-E-E-L-E-E-I-
      130                 140  *              150
-E-E-D-A-G-L-G-N-G-G-L-G-R-L-A-A-C-F-L-D-S-M-A-T-L-
         *        160                 170
-G-L-A-A-Y-G-Y-G-I-R-Y-E-F-G-I-F-N-Q-K-I-C-G-G-W-Q-
      180                 190                 200
-M-E-E-A-D-D-W-L-R-Y-G-N-P-W-E-K-A-R-P-E-F-T-L-P-V-
                  210                 220
-H-F-Y-G-R-V-E-H-T-S-Q-G-A-K-W-V-D-T-Q-V-V-L-A-M-P-
      230                 240                 250
-Y-D-T-P-V-P-G-Y-R-N-N-V-V-N-T-M-R-L-W-S-A-K-A-P-N-
                  260          ↓      270
-D-F-N-L-K-D-F-N-V-G-G-Y-I-Q-A-V-L-D-R-N-L-A-E-N-I-
      280                 290                 300
-S-R-V-L-Y-P-N-D-N-F-F-E-G-K-E-L-R-L-K-Q-E-Y-F-V-V-
                  310                 320
-A-A-T-L-Q-D-I-R-R-F-K-S-S-K-F-G-C-R-D-P-V-R-T-N-F-
      330                 340
-D-A-F-P-D-K-V-A-I-Q-L-N-D-T-H-P-S-L-A-I-P-E-L-M-
```

FIGURE 14: The amino acid sequence in one-letter code (see Figure 6) of residues 1 through 349 of rabbit muscle phosphorylase. Asterisks locate serine-14 (S) which is phosphorylated by ATP in the presence of phosphorylase kinase, tyrosine-155 (Y) which is labeled by an analogue of AMP, and cysteine residues 108 and 142 (C) which are involved in the association of subunits. Vertical arrows indicate the major sites of limited proteolysis by subtilisin (residues 16–17 and 264–265). The amino terminus is acetylated (Ac).

that derived by sequence analysis (Table III). Thus, the sequence of phosphorylase is extended to aspartic acid-227 (Figure 13).

The tripeptide sequence Pro-Tyr-Asp, following methionine in L_s-T-4, provides an overlap with the amino terminus of CB11. Of the 17 residues in fragment CB11, 14 were placed by Sequencer analysis, including the single arginine in the structure -Arg-Asn-Asn-Val-Val-238. This placement leads to the alignment of L_s-T-5, whose sequence X-X-Val-Val-Asn-(Thr)-Met-Arg is consistent with that position. The composition of CB11 supports the placement of threonine at residue-240. More convincing data were obtained in this region by analysis of a minor fragment obtained from the CNBr digest by rechromatography of fraction IV (Figure 5A) on the same column. Amino-terminal analysis yielded a sequence identical with that of the first 9 residues of CB11, followed by Arg-Asn-Asn-Val-Val-Asn-(Thr)-Met-Arg-Leu-Trp-(Ser)-Ala- (Figure 13). This fragment is designated CB11-17N in Tables III and IV, indicating its overlap with the amino terminus of CB17 (Arg-Leu-Trp-Ser-Ala in Figure 10). Thus, a continuous 349-residue sequence is provided from CB14 through CB5, CB12, CB16, CB21, CB20, and CB11 to CB17 (Figure 14).

The alignment of CB21 and CB20 was confirmed by analysis of a minor fragment from the CNBr digest. As mentioned earlier, fraction II (Figure 5A) revealed the same amino terminus as CB21 but a composition equal to the sum of those of CB21 and CB20.

Site of AMP Binding. It was reported by Anderson & Graves (1973) and by Anderson et al. (1973) that in the AMP-binding site of phosphorylase *b* a tyrosine residue was modified by an affinity label analogue of AMP. Phosphorylase *b* (22.9 mg) labeled with [^{14}C]-8-[*m*-(*m*-fluorosulfonylbenzamido)benzylthio]adenine (68 500 cpm) was generously donated by Dr. Donald Graves (Iowa State University). The product was treated with subtilisin BPN′ and then with cyanogen bromide as already described. (Control experiments indicated that the label was stable in 72% formic acid.) Approximately 85% of the radioactivity was recovered in segment L_s (5.82×10^3 cpm), and about 10% (0.69×10^3 cpm) in the void volume of the column (which contained undigested protein, and/or aggregated light segment). In the primary separation of CB fragments (Figure 5A), the radioactivity was recovered in fraction II (0.87×10^3 cpm; 15% of that applied), which had been already identified as a minor fragment overlapping CB21 and CB20, and in an area encompassing fractions V and VI (3.90×10^3 cpm; 67% of that applied) corresponding to fragments CB11, CB12, CB16, CB17N, or CB21. Most of the radioactivity (80%) in fractions V and VI was recovered in CB21, isolated as before by precipitation with dilute acid. By amino acid analysis recovery of CB21 was 47% and the recovery of radioactivity was 45.6%, confirming that CB21 contained the primary site of modification.

Sequencer analysis of a mixture of labeled CB21 (11 nmol; 312 cpm) and cold carrier CB21 (400 nmol) yielded the expected amino-terminal sequence (for 13 residues), but no radioactivity was recovered with the phenylthiohydantoins in chlorobutane extracts. After these 13 cycles the radioactivity could be extracted from the cup with ethyl acetate. The Sequencer program was then modified by substituting ethyl ac-

etate for chlorobutane and the analysis repeated using 22 nmol of labeled CB21 (624 cpm). Radioactivity was recovered in the ethyl acetate extracts of cycle 8 (62%), cycle 9 (13%), and cycle 10 (3%), indicating that tyrosine-155 (the eighth residue in CB21) had been labeled.

Discussion

Each subunit of rabbit muscle glycogen phosphorylase contains 841 amino acid residues, representing one of the longest polypeptide chains yet subjected to sequence analysis. Preliminary attempts to apply conventional analytical procedures were largely unsuccessful because of the complexity of the mixtures of peptides generated by chemical or enzymatic fragmentation. The strategy finally adopted is based on the generation of a small number of very large fragments by specific cleavage techniques. Fowler & Zabin (1977), facing similar difficulties in their analysis of β-galactosidase (1021 residues), applied immunological techniques to identify key fragments. Here, we first cleaved the protein by limited proteolysis into two segments (L_s and H_s), and then applied specific fragmentation techniques to segment L_s (this communication) and segment H_s (Titani et al., 1978).

The relatively high resistance of phosphorylase to proteolytic attack formed the basis of the primary fragmentation process applied in this investigation. Of the 840 peptide bonds in phosphorylase, only 9 are hydrolyzed by the four different proteases listed in Table I. These bonds are clustered in two regions (residues 14–20 and 261–278) encompassing less than 3% of the mass of the molecule. Within these protease-sensitive regions the specific cleavage site is apparently dictated by the primary specificity of the protease.

Cleavage by subtilisin generated two major segments (L_s and H_s) and thus divided the analytical problem into two major projects, each comparable in complexity to the sequence analysis of a conventional large protein. Analysis of segment L_s relied principally on Edman degradation of 9 primary fragments, derived by cleavage by cyanogen bromide of methionyl bonds or by acid of the single Asp-Pro bond. In this manner ca. 70% of the 247 residues in segment L_s were identified. Specific cleavage by trypsin of arginyl bonds generated fragments which provided most of the additional sequence data including sequences overlapping the primary fragments. The remaining amino acid residues were identified by the more conventional (and tedious) approach of generating small peptides from enzymatic digests. The sequence of segment L_s is aligned with the amino terminus of the whole molecule by the overlapping sequence data of Titani et al. (1975) and with the amino terminus of segment H_s by the overlap provided by fragment CB17 (Figure 2).

Two other general aspects of the methodology merit comments, i.e., the separation procedures and the techniques of Edman degradation. The techniques which are conventionally applied to the purification of native proteins or to the separation of small peptides were generally not directly applicable in the present work since mixtures of large denatured polypeptide fragments tended to aggregate and precipitate. However, limited cleavage produced relatively simple mixtures of large fragments which obviated the use of high resolution techniques. It was usually sufficient to dissolve the mixture in a denaturing solvent and to fractionate it by gel filtration. Many fragments were soluble in 9% formic acid but acidic fragments required more alkaline solvents (e.g., 0.1 M NH_4HCO_3, pH 8). In certain other cases, the addition of urea (in acid) was necessary. Ion-exchange chromatography resolved complex mixtures of small peptides as well as the two large segments generated by subtilisin cleavage. In the latter instance it was necessary to include 7 M urea in the solvent and to employ a hydrophilic matrix (SP-Sephadex) for fractionation. Two peptides (CB21 and T-4 from CB17) were conveniently separated from mixtures by selective precipitation.

It is now generally accepted that it is more efficient to apply extended Edman degradation to a single large fragment than to several shorter ones. The data presented in this communication were derived over a period of several years in the course of which new degradation procedures and new instrumentation were developed. A few early sequence analyses involved manual degradation, whereas several, more recent analyses took advantage of automated degradation on a solid phase support (Table IV). The vast majority of degradations were carried out in the automated liquid phase instrument (Sequencer). With large fragments, stepwise yields of 93–96% were routinely observed. During the course of this work the capability of retaining small peptides in the spinning cup was improved by refinement of programs of extraction and drying (cf., Walsh et al., 1978; Hermodson et al., 1977). Most recently polybrene was included as an inert carrier (Tarr et al., 1978).

The extent of successful degradation of each fragment is influenced by any of three other limiting factors: (1) the susceptibility of the peptide to gradual nonspecific acid hydrolysis during degradation; (2) the decreased yields at residues of proline, glutamine, or of asparagine in the sequence Asn-Gly (Walsh et al., 1978); and (3) the reliability of the method for detecting degradation products. In the early phase of this work phenylthiohydantoins were identified by gas/liquid chromatography as their silylated derivatives in conjunction with spot tests for arginine and histidine. Later this procedure was supplemented or replaced by more quantitative identifications by high performance liquid chromatography. In each case (Table IV) identifications were judged to be proven if a single phenylthiohydantoin appearing in a given cycle clearly rose above the background of the previous cycle and dropped in the next cycle. We regard these observations as semiquantitative in nature and hesitate to document apparent yields, especially from gas/liquid chromatography. Even in the case of high performance liquid chromatography data, appropriate subtraction of background phenylthiohydantoins and correction for incomplete reaction in each cycle introduce rather arbitrary adjustments of yield which are more realistically expressed by the qualitative judgments in Table IV than by the seemingly more precise quantitations shown in Figure 7.

Any proof of primary structure should at best be regarded as a firm working hypothesis of the actual structure (Hirs et al., 1960). Most of the sequence deduced for residues 1–349 is proven in an unambiguous fashion by replicate or overlapping analyses, but certain data are derived only from single analyses (e.g., residues 77–78, 107–119, 162–169, 297–306, and 312–314). In each case the sequence data are consistent with the composition of a small peptide derived from the corresponding region. In addition, identifications of lysine and placements of amides were more reliable by high performance liquid chromatography than by gas–liquid chromatography (Table IV), whereas identifications of serine and threonine (which tend to be destroyed during analysis) are less certain by either method. However, supportive arguments are based on the composition of small peptides, electrophoretic mobilities, replicate analyses, or overlapping sequences. For example, identifications of serine residues at positions 86, 146, and 342, threonine at 240 and 303, and glutamine at 88 and 264 were all based on a combination of tentative identifications by gas–liquid chromatography and compositional arguments.

The three weakest points in the present proof of sequence

appear to be cysteine-317, aspartic acid-319, and asparagine-235, 236. The cysteine and aspartic acid residues at positions 317 and 319 were not observed directly as their phenylthiohydantoins. Their presence in this region of fragment CB17 was indicated by the compositions of two overlapping subpeptides, T-12 and TS-6 (Table VI), but Edman degradation did not yield identifiable products corresponding to residues 317 or 319 (Figure 10). Our placement is based on the sequence of a cysteinyl tetrapeptide, isolated by Zarkadas et al. (1968) from phosphorylase, which does not correspond in composition to any of the other eight cysteinyl residues of phosphorylase (Titani et al., 1978). The placement of aspartic acid at 319 is also supported by the observations that fragment AC-γ was generated from CB17 (Figure 10) and AC-4 from segment H_s (Titani et al., 1978) by treatment known to exclusively cleave Asp-Pro bonds. The location of the two contiguous asparagine residues at positions 235 and 236 (Figure 13) is not entirely satisfactory. These placements were made on the basis of Sequencer analyses of fragments CB11 and CB11-17N (Table IV). Although the amino acid composition of peptide CB11 (residues 225-251; Table III) is consistent with this placement, that of L_s-T-5 (residues 235-252, Table VII) indicates two asparagine residues rather than three in this region and Sequemat analyses failed to identify residues 235 and 236. Thus, a measure of uncertainty remains in this region of the sequence.

Several residues of interest in the enzymatic function of this molecule have been placed in sequence by this analysis. The location of the site of phosphorylation at serine-14 has already been discussed by Titani et al. (1975). In addition, the tyrosine residue which was affinity labeled by the AMP analogue of Anderson & Graves (1973) is now located at residue 155. Battell et al. (1968) and Avramovic-Zikic et al. (1970) described two cysteine residues which react rapidly with iodoacetamide without changing either the catalytic function or the aggregation of the oligomeric form of the enzyme. These can now be placed at positions 171 and 317. The slower alkylation of two other cysteine residues, now identified as residues 108 and 142, correlates with both dissociation of the oligomeric enzyme and abolition of enzymatic activity (Battell et al., 1968). The present identification of cysteine-142 in the sequence Ala-Ala-Cys-Phe is at variance with the earlier identification by Battell et al. (1968) in the sequence Ala-Cys-Ala-Phe, but Titani et al. (1978) report that this is the only cysteine residue of nine in phosphorylase which can account for the composition of the peptide in the earlier study.

Acknowledgments

The authors wish to thank Dr. Edmond H. Fischer for the donation of rabbit muscle phosphorylase *b*, for his encouragement and valuable discussion throughout the present work, and Dr. Donald J. Graves for donating the affinity-labeled phosphorylase *b*. Thanks are also due to Dr. John C. Saari, Dr. Mark A. Hermodson, and Mrs. Sheila Balian for contributions to the initial phase of this work, to Richard Olsgaard and Richard Granberg for excellent technical assistance, to Roger D. Wade for his assistance in the solid phase Edman degradations, and to Rosa Mae Macdonald for help with the high performance liquid chromatography.

References

Ambler, R. T. (1967) *Methods Enzymol. 11,* 436.
Anderson, R. A., & Graves, D. J. (1973) *Biochemistry 12,* 1895.
Anderson, R. A., Parrish, R. F., & Graves, D. J. (1973) *Biochemistry 12,* 1901.
Avramovic-Zikic, O., Smillie, L. B., & Madsen, N. B. (1970) *J. Biol. Chem. 245,* 1558.
Battell, M. L., Zarkadas, C. G., Smillie, L. B., & Madsen, N. B. (1968) *J. Biol. Chem. 243,* 6202.
Bennett, J. C. (1967) *Methods Enzymol. 11,* 330.
Bradshaw, R. A., Garner, W. H., & Gurd, F. R. N. (1969) *J. Biol. Chem. 244,* 2149.
Bridgen, P. J., Cross, G. A. M., & Bridgen, J. (1976) *Nature (London) 263,* 613.
Cavins, J. F., & Friedman, M. (1970) *Anal. Biochem. 35,* 489.
Cohen, P., Saari, J. C., & Fischer, E. H. (1973) *Biochemistry 12,* 5233.
Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238,* 622.
Crewther, W. G., & Inglis, A. S. (1975) *Anal. Biochem. 68,* 572.
Drapeau, G. R., Boily, Y., & Houmard, J. (1972) *J. Biol. Chem. 247,* 6720.
Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1,* 80.
Fischer, E. H., Pocker, A., & Saari, J. C. (1970) *Essays Biochem. 6,* 23.
Fischer, E. H., Heilmeyer, L. M. G., Jr., & Haschke, R. H. (1971) *Curr. Top. Cell Regul. 4,* 211.
Fowler, A. V., & Zabin, I. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74,* 1507.
Fraser, K. J., Poulsen, K., & Haber, E. (1972) *Biochemistry 11,* 4974.
Graves, D. J., & Wang, J. H. (1972) *Enzymes, 3rd Ed. 7,* 435.
Green, A. A., & Cori, G. T. (1943) *J. Biol. Chem. 143,* 21.
Gross, E. (1967) *Methods Enzymol. 11,* 238.
Hayashi, R. (1977) *Methods Enzymol. 47,* 84.
Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H., & Walsh, K. A. (1978) *Biochemistry 17* (second of three papers in a series in this issue).
Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry 11,* 4493.
Hermodson, M. A., Schmer, G., & Kurachi, K. (1977) *J. Biol. Chem. 252,* 6276.
Hirs, C. H. W. (1967) *Methods Enzymol. 11,* 197.
Hirs, C. H. W., Moore, S., & Stein, W. H. (1960) *J. Biol. Chem. 235,* 633.
Houmard, J., & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69,* 3506.
Hugli, T. E., & Moore, S. (1972) *J. Biol. Chem. 247,* 2828.
Konigsberg, W., & Hill, R. J. (1962) *J. Biol. Chem. 237,* 2547.
Lewis, U. J., Williams, D. E., & Brink, N. G. (1956) *J. Biol. Chem. 222,* 705.
Omenn, G. S., Fontana, A., & Anfinsen, C. B. (1970) *J. Biol. Chem. 245,* 1895.
Pisano, J. J., & Bronzert, T. J. (1972) *Anal. Biochem. 45,* 43.
Raibaud, O., & Goldberg, M. E. (1973) *Biochemistry 12,* 5154.
Ryle, A. P., Sanger, F., Smith, L. F., & Kitai, R. (1955) *Biochem. J. 60,* 541.
Saari, J. C., & Fischer, E. H. (1973) *Biochemistry 12,* 5225.
Salnikow, J., Liao, T., Moore, S., & Stein, W. H. (1973) *J. Biol. Chem. 248,* 1480.
Schroeder, W. A. (1967) *Methods Enzymol. 11,* 351.
Sevilla, C. L., & Fischer, E. H. (1969) *Biochemistry 8,* 2161.

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