References

Bartlett, G. R. (1959), *J. Biol. Chem. 234*, 466.
Burton, A. J., and Carter, H. E. (1964), *Biochemistry 3*, 411.
Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956), *Anal. Chem. 28*, 350.
Fiske, C. H., and Subbarow, Y. (1925), *J. Biol. Chem. 66*, 375.
Gahan, L. C., Sandford, P. A., and Conrad, H. E. (1967), *Biochemistry 6*, 2755.
Gmeiner, J., Lüderitz, O., and Westphal, O. (1969), *Eur. J. Biochem. 7*, 370.
Heath, E. C., Mayer, R. M., Edstrom, R. D., and Beaudreau, C. A. (1966), *Ann. N. Y. Acad. Sci. 133*, 315.
Hersberger, C., David, M., and Binkley, S. B. (1968), *J. Biol. Chem. 243*, 1585.
Huber, C. N., Scobell, H. D., Tai, H., and Fisher, E. E. (1968), *Anal. Chem. 40*, 207.
Koeltzow, D. E., Epley, J. D., and Conrad, H. E. (1968), *Biochemistry 7*, 2920.
Leloir, L. F., and Cardini, C. E. (1957), *Methods Enzymol. 3*, 840.
Lüderitz, O., Jann, K., and Wheat, R. (1968), *Comp. Biochem. 26A*, 105.
Lüderitz, O., Staub, A. M., and Westphal, O. (1966), *Bacteriol. Rev. 30*, 192.
Nikaido, H. (1969), *Advan. Enzymol. 31*, 77.
Osborn, M. J. (1963), *Proc. Nat. Acad. Sci. U. S. 50*, 499.
Osborn, M. J. (1969), *Annu. Rev. Biochem. 38*, 501.
Ryhage, R., and Stenphages, E. (1960), *Ark. Kemi 15*, 551.
Scher, M., Lennarz, W. J., and Sweeley, C. C. (1968), *Proc. Nat. Acad. Sci. U. S. 59*, 1313.
Shin, J. M. (1962), *Anal. Chem. 34*, 1164.
Shively, J. E., and Conrad, H. E. (1970), *Biochemistry 9*, 33.
Wright, A., Dankert, M., Fennessey, P., and Robbins, P. W. (1967), *Proc. Nat. Acad. Sci. U. S. 57*, 1798.

Amino Acid Sequence around 3-Methylhistidine in Rabbit Skeletal Muscle Actin*

Marshall Elzinga

ABSTRACT: Actin contains 1 mole of the unusual amino acid 3-methylhistidine [2-amino-3-(1-methyl-4-imidazolyl)propanoic acid] per mole of protein. All of the 3-methylhistidine is found in only 1 of the 17 cyanogen bromide peptides of actin; this peptide contains a total of 35 amino acid residues, and the determination of its primary structure is reported in this paper. The sequence is: Gly-Gln-Lys-Asp-Ser-Tyr-Val-Gly-Asp-Glu-Ala-Gln-Ser-Lys-Arg-Gly-Ile-Leu-Thr-Leu-Lys-Tyr-Pro-Ile-Glu-3-methylhistidine-Trp-Gly-Ile-Ile-Thr-Asn-Asp-Asp-Hse.

The unusual amino acid, 2-amino-3-(1-methyl-4-imidazolyl)propanoic acid (3-methylhistidine), has been identified as a natural constituent of the myofibrillar proteins actin and myosin (Asatoor and Armstrong, 1967; Johnson *et al.*, 1967). There is one residue of 3-methylhistidine in the single polypeptide chain of rabbit muscle actin while in myosin, which is composed of two heavy chains and two (or three) light chains, there is an average of one residue of 3-methylhistidine in each of the two heavy subunits in white skeletal muscle myosin.

Studies on peptides from actin have suggested that the 3-methylhistidine is present as a single fully methylated residue; all the 3-methylhistidine in actin is found in only 1 of the 17 peptides that are released when actin is cleaved with cyanogen bromide (Elzinga, 1970; Adelstein and Kuehl, 1970). Previously Johnson *et al.* (1967) found that when the soluble portion of a tryptic digest of *S*-β-carboxymethylactin was chromatographed on a column of Dowex 1, 3-methylhistidine was localized in one of the effluent peaks.

The presence of derivatives of both lysine and histidine has been shown in a variety of proteins besides actin. Among these are histones, which contain ε-*N*-acetyllysine, ε-*N*-monomethyllysine, and ε-*N*-dimethyllysine (Ogawa *et al.*, 1969; DeLange *et al.*, 1969), and possibly 3-methylhistidine (Gershey *et al.*, 1969); myosin, with 3-methylhistidine, ε-*N*-monomethyllysine, and ε-*N*-trimethyllysine (Trayer *et al.*, 1968; Kuehl and Adelstein, 1969; Hardy and Perry, 1969; Huszar and Elzinga, 1969); and wheat germ cytochrome *c*, with ε-*N*-trimethyllysine (DeLange *et al.*, 1970).

In the case of actin and myosin there is evidence that the methylated derivatives of both histidine and lysine are not directly incorporated into the polypeptide chains, but rather the methyl groups are added enzymatically at specific positions, with *S*-adenosylmethionine serving as the methyl donor (Asatoor and Armstrong, 1967; Trayer *et al.*, 1968).

As part of a study of the complete amino acid sequence of rabbit skeletal muscle actin, the sequence of the cyanogen bromide peptide that contains 3-methylhistidine has been determined. It is anticipated that this basic structural informa-

* From the Department of Muscle Research, Boston Biomedical Research Institute and the Department of Neurology, Harvard Medical School, Boston, Massachusetts 02114. *Received August 24, 1970.* This work was carried out while the Department of Muscle Research was a part of the Retina Foundation. The author is an Established Investigator of the American Heart Association, Inc. The work was supported by research grants from the Medical Foundation of Boston, the National Institutes of Health (H-5949 and 1-S01-FR-05527), and the National Science Foundation. A preliminary report of this work was presented at the 3rd International Biophysics Congress, Cambridge, Mass., August 1969.

tion will facilitate investigations designed to elucidate the function of this unusual amino acid, including chemical modification experiments directed toward either 3-methylhistidine or amino acids that are near it in the protein, and comparative amino acid sequence studies. In the present report the sequence around the 3-methylhistidine residue in actin is described, and in the following paper (Huszar and Elzinga, 1971) the sequence of this peptide is compared with the amino acid sequence around 3-methylhistidine in myosin.

Methods and Materials

Preparation of CB-10. Peptide CB-10 was prepared from rabbit skeletal muscle actin as described previously (Elzinga, 1970) with two exceptions. First, the actin was isolated from acetone powder by the method of Rees and Young (1967); this procedure incorporates a G-200 gel filtration step for the final purification of the actin. Secondly, CB-10 was purified using a column of SE-Sephadex rather than DEAE-cellulose. SE-Sephadex (G-25) was washed with 1 N HCl and 1 N NaOH and equilibrated with 25% acetic acid. A 0.9 × 15 cm column was used at 37°. The sample containing 2–6 μmoles of peptides CB-10 and CB-12 was applied under gravity flow in 2–4 ml of 25% acetic acid. A linear gradient in pyridine concentration was applied over a total volume of 300 ml by using 150 ml of 25% acetic acid and 150 ml of pH 4 pyridine–acetic acid buffer. The pH 4 buffer was prepared by mixing appropriate volumes of pH 3.1 and pH 5.0 pyridine–acetic acid buffers (Schroeder, 1967). A free-flow rate of 20 ml/hr was attained by adjusting the height of the buffer reservoir. Peptides were detected by monitoring the column at 280 nm and at 570 nm following alkaline hydrolysis and reaction with ninhydrin (Hirs, 1967).

Enzyme Digestion. CB-10 and some of the smaller peptides were digested with trypsin, chymotrypsin, thermolysin, aminopeptidase M, and carboxypeptidases A and B using the general procedures described previously (Elzinga, 1970). In addition, T-2 was digested with subtilopeptidase A (Protease, subtilopeptidase A, lot 28B-2340, Sigma Chemical Co.) under the following conditions. The peptide was dissolved to a concentration of 1 μmole/ml in 1 mM Tris, pH 7.5. Enzyme (50 μg, 50 μl of a solution containing 1 mg/ml in water) was added and the mixture was incubated at 25° for 4 hr. The digestion was terminated by lowering the pH to 3 by addition of glacial acetic acid. The solution was dried by evaporation under reduced pressure, and the peptides were separated by paper chromatography using Whatman No. 3MM paper. The chromatograms were developed using the organic phase of a 1-butanol–acetic acid–water (4:1:5) mixture as the solvent.

Edman Degradation. This procedure was carried out as described previously (Elzinga, 1970).

Amino Acid Analyses. These were carried out as described previously (Elzinga, 1970) with one change. In the absence of histidine, 3-methylhistidine was detected between lysine and ammonia on the standard basic column (0.6 × 10 cm) developed at 60° with the use of a pH 5.35 0.35 M sodium citrate buffer.

Results

Amino Acid Composition. The 3-methylhistidine-containing peptide, designated CB-10, was isolated from rabbit skeletal muscle actin. The amino acid composition of the peptide, as reported earlier (Elzinga, 1970), is given in Table I. The

TABLE I: Amino Acid Compositions of Peptides.

Amino Acid	CB-10	T-1	T-2	T-3	T-4	T-5	T-1-C-1	T-1-C-2	T-1-C-3	TL-1[d]	C-1	C-2	C-3	C-4
Lysine	3.08 (3)	0.15	1.00 (1)	0.93 (1)	0.95 (1)						0.91 (1)	1.28 (1)		1.20 (1)
3-Methylhistidine	0.84 (1)	0.98 (1)							0.95 (1)	0.77 (1)				0.95 (1)
Arginine	1.04 (1)					1.00 (1)						1.05 (1)		
Aspartic acid	4.78 (5)	2.80 (3)	2.07 (2)				2.00 (2)	1.04 (1)			1.05 (1)	1.10 (1)		0.20
Threonine	1.87 (2)	0.85 (1)		0.88 (1)				1.02 (1)					0.95 (1)	0.15
Serine	1.98 (2)	0.19	2.14 (2)								0.84 (1)	0.98 (1)		0.20
Glutamic acid	3.91 (4)	1.21 (1)	2.07 (2)		0.99 (1)				1.05 (1)	1.06 (1)	1.28 (1)	2.18 (2)		1.05 (1)
Proline	0.90 (1)	0.88 (1)							1.06 (1)					0.97 (1)
Glycine	3.77 (4)	1.17 (1)	1.09 (1)	1.07 (1)	1.07 (1)			1.06 (1)		0.88 (1)	1.09 (1)	2.13 (2)		0.32
Alanine	0.89 (1)	0.20	1.00 (1)									1.03 (1)		
Valine	1.22 (1)	0.11	0.89 (1)									0.84 (1)		
Isoleucine	3.84 (4)[a]	2.94 (3)[a]		1.12 (1)				1.91 (2)[a]	1.02 (1)	1.06 (1)		0.90 (1)		0.92 (1)
Leucine	1.98 (2)	0.17		2.26 (2)								0.87 (1)	1.04 (1)[c]	
Tyrosine	1.98 (2)	1.07 (1)	0.83 (1)						0.86 (1)		0.81 (1)			0.88 (1)
Homoserine[b]	0.80 (1)	0.86 (1)					0.84 (1)							
Tryptophan	0.90 (1)[c]	(1)[c]							0.11 (1)	0.65 (1)				(1)[c]
Yield (%)		85	60	65	90	90	65	38	25	30	43	37	40	33

[a] Obtained from 96-hr hydrolysates. [b] Homoserine plus homoserine lactone. [c] Estimated spectrophotometrically. [d] Aminopeptidase M digest.

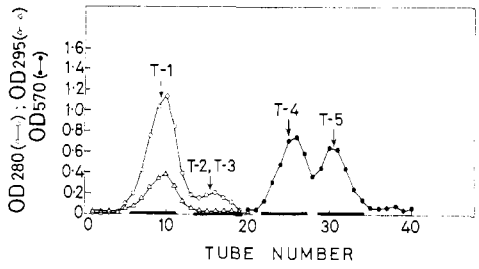

FIGURE 1: Sephadex G-50 chromatography of a tryptic digest of CB-10. The column size was 1.9 × 400 cm, the solvent was 25% acetic acid, the column was run at 23°, the flow rate was 12 ml/hr, and the fraction size was 5 ml. A volume of 550 ml was passed through the column before collection was begun.

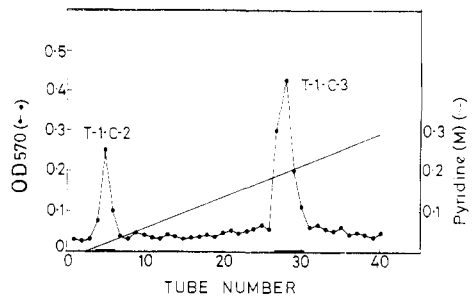

FIGURE 3: Phosphocellulose chromatography of peak T-1-C-2, T-1-C-3 from the column shown in Figure 2. The column was 0.9 × 15 cm and was run at 40°. The column was developed by applying a linear gradient from 0.0 to 0.5 M pyridine in 25% acetic acid over a total volume of 200 ml; 3-ml fractions were collected.

molar extinction coefficient of CB-10 at 280 nm in 10^{-4} M Tris, pH 8.0, was about 6600 and the amino acid analysis showed that there were two residues of tyrosine in the peptide.

If one assumes a molar extinction coefficient of 1100 for tyrosine (Fruton and Simmonds, 1961) the two tyrosine residues would account for 2200 absorption units. This leaves about 3800, which suggests that the peptide contains one tryptophan. Subsequent degradation experiments confirmed that the peptide indeed contains a single tryptophan residue.

Sequence Determination. CB-10 was first digested with trypsin (enzyme–substrate, 1:20 by weight), and the mixture of peptides was chromatographed on Sephadex G-50 (Elzinga, 1970), as shown in Figure 1. The peptides, designated T-1, T-4, and T-5, were obtained essentially pure after a single gel filtration. The peptides in the peak designated T-2 and T-3 were resolved on a Dowex 50-X2 column. The compositions of the tryptic peptides are given in Table I; together they account for all of the amino acid residues in CB-10.

T-1. T-1 was the only peptide that contained homoserine and its absorption properties indicated that it also contained the single tryptophan residue. This conclusion was based upon the fact that of the five peptides, T-1 had most of the 280-nm absorption and was the only peak to absorb at 295 nm, a wavelength at which tryptophan but not tyrosine absorbs (Fruton and Simmonds, 1961). Two steps of Edman degradation on T-1 gave the following results. (In presenting the results of subtractive Edman degradations the amino acid assumed to be lost at a given degradation step is indicated by boldface type; homoserine was assumed to be the carboxyl-terminal amino acid in each peptide in which it was found. The value given for homoserine is actually homoserine plus homoserine lactone. A dash (—) indicates that the amino acid was not determined quantitatively. The sequence data are summarized in Figure 5.) Step 1:[1] **Tyr, 0.17;** Pro, 0.86; Ile, 3.06; Glx, 1.16; 3-methylhistidine, 0.98; Gly, 1.09; Thr, 0.85; Asx, 2.71; Hse, 0.74. Step 2: Tyr, 0.18; **Pro, 0.06;** Ile, 2.93; Glu, 1.36; 3-methylhistidine, 0.90; Gly, 1.30; Thr, 1.07; Asx, 3.00, Hse, —.

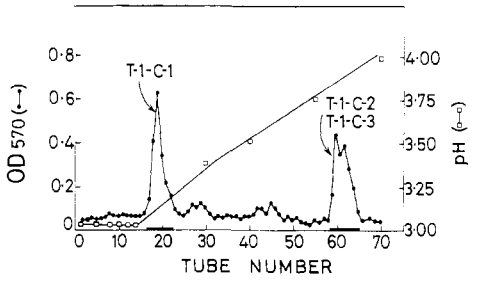

FIGURE 2: Dowex 50-X2 chromatography of a chymotryptic digest of T-1. The column was 0.9 × 60 cm and was run at 40°. Elution was accomplished using a linear gradient of 250 ml each of pH 3.1, 0.2 M and pH 5.0, 2.0 M pyridine–acetic acid buffers; 4-ml fractions were collected.

These results established that the amino terminus of T-1 was Tyr-Pro: T-1 was then digested with chymotrypsin (enzyme–substrate, 1:50, 4 hr). The peptide mixture obtained was chromatographed as shown in Figures 2 and 3, and three pure peptides were obtained. Their amino acid compositions are given in Table I and together they account for all of the amino acids in T-1.

T-1-C-1. This peptide contained homoserine, which indicated that it arose from the carboxyl terminus of T-1, and two residues of either aspartic acid or asparagine. T-1-C-1 was digested with aminopeptidase M, an enzyme which can digest peptides without removing amides, and an amino acid analysis of the aminopeptidase M digest showed that the peptide contained two residues of aspartic acid, no asparagine, and one residue of homoserine. Placing homoserine at the C terminus, the sequence of T-1-C-1 is Asp-Asp-Hse.

T-1-C-2. This pentapeptide was shown by aminopeptidase M digestion to have one residue of asparagine, and its sequence was established by doing four steps of Edman degradation: Step 1: **Gly, 0.12;** Ile, 2.00; Thr, 1.04; Asn, 0.99. Step 2: Gly, 0.39; **Ile, 1.08;** Thr, 1.09; Asn, 1.08. Step 3: Gly, 0.15; **Ile, 0.11;** Thr, 1.01; Asn, 0.98. Step 4: Gly, —; Ile, —; **Thr, 0.15;** Asn, 1.00. Thus the sequence of T-1-C-2 is Gly-Ile-Ile-Thr-Asn.

T-1-C-3. This peptide contains six residues, including the amino acids (tyrosine and proline) that had been found at the amino terminus of T-1; therefore, T-1-C-3 was placed at the amino end of T-1. Since the first two residues of this peptide had already been established, the first two steps of Edman degradation were omitted, and the results of steps 3 and 4 were as follows: Step 3: Tyr, —; Pro, 0.10; **Ile, 0.40;** Glx, 0.97; 3-methylhistidine, 1.06; Trp, —. Step 4: Tyr, —; Pro, —; Ile, 0.25; **Glx, 0.42;** 3-methylhistidine, 1.00; Trp, —. Thus the first four residues were established to be Tyr-Pro-Ile-Glx. The two residues remaining were 3-methylhistidine and tryptophan, and from the specificity of chymotrypsin it appeared that tryptophan must be at the carboxyl terminus. For

[1] All peptides from T-1 that contained isoleucine were hydrolyzed for 96 hr before amino acid analysis; this extended hydrolysis time was necessary to completely release isoleucine.

reasons which we cannot explain at this time T-1-C-3 was refractive to digestion by carboxypeptidase A and it was not possible to prove the carboxyl-terminal sequence using this enzyme. The sequence information that was already available on T-1-C-3 and T-1-C-2 suggested that thermolysin would release from CB-10 a pentapeptide containing both 3-methylhistidine and tryptophan by hydrolysis of two peptide bonds involving isoleucine. Thus CB-10 was treated with thermolysin (enzyme–substrate, 1:20, 4 hr, 37°) and the digest was chromatographed on Sephadex G-50 and Dowex 50-X2. In each case the peak that contained both tryptophan, as determined by the Ehrlich stain on paper, and 3-methylhistidine, measured by amino acid analysis, was isolated. The pure peptide that was isolated was designated TL-1, and its amino acid composition after aminopeptidase M digestion is shown in Table I. This is a pentapeptide that contains one residue of glutamic acid, and both 3-methylhistidine and tryptophan. By inspection of known sequences from T-1-C-3 and T-1-C-2, the partial sequence of TL-1 could be written Ile-Glu-(3-methylhistidine, Trp)-Gly. Since tryptophan was destroyed during the acid hydrolysis step of the subtractive Edman degradation, it was necessary to place the tryptophan residue by difference; the first three steps of Edman degradation were performed with the following result: Step 1: **Ile, 0.00;** Glu, 1.00; 3-methylhistidine, 1.02; Trp, —; Gly, 0.99. Step 2: Ile, 0.00; **Glu, 0.29;** 3-methylhistidine; 1.05; Trp, —; Gly, 1.00. Step 3: Ile, 0.00; Glu, 0.23; **3-methylhistidine, 0.50;** Trp, —; Gly, 1.00.

Finally, four steps of Edman degradation were performed, and the residue was analyzed without acid hydrolysis; under these conditions the carboxyl-terminal amino acid from a pentapeptide would be released. We recovered 0.25 residues of free glycine, proving that TL-1 was indeed a pentapeptide with glycine at the carboxyl end. Tryptophan was placed between 3-methylhistidine and glycine by difference, and the sequence of TL-1 is Ile-Gly-3-methylhistidine-Trp-Gly.

By combining the sequence data on T-1, T-1-C-1, T-1-C-2, T-1-C-3, and TL-1, the sequence of T-1 may be written Tyr-Pro-Ile-Glu-3-methylhistidine-Trp-Gly-Ile-Ile-Thr-Asn-Asp-Asp-Hse.

T-2. The amino acid composition of T-2 is shown in Table I. In order to determine if any of the four dicarboxylic acids were present as amides, T-2 was digested with aminopeptidase M. Amino acid analysis of this digest indicated that the peptide contained two residues of aspartic acid, one residue of glutamic acid, and one residue of glutamine. T-2 was subjected to five steps of Edman degradation with the following results. Step 1: **Asx, 1.07;** Ser, 1.76; Glu, 2.25; Gly, 1.10; Ala, 0.81; Val, 1.12; Tyr, 0.90; Lys, —. Step 2: Asx, 1.12; **Ser, 1.14;** Glx, 2.31; Gly, 1.15; Ala, 0.83; Val, 0.86; Tyr, 0.87, Lys, —. Step 3: Asx, 1.05; Ser, 1.06; Glx, 2.13; Gly, 1.11; Ala, 0.79; Val, 0.99; **Tyr, 0.22;** Lys, —. Step 4: Asx, 1.02; Ser, 1.03; Glx, 2.12; Gly, 1.04; Ala, 0.92; **Val, 0.41;** Tyr, 0.28, Lys, —. Step 5: Asx, 1.15; Ser, 0.92; Glx, 2.11; **Gly, 0.68;** Ala, 0.93; Val, 0.34; Tyr, 0.20; Lys, —.

These results indicated that the partial sequence of T-2 was Asp-Ser-Tyr-Val-Gly-(Asp,Ser,Glu,Gln,Ala)Lys.

T-2 was then digested with carboxypeptidases A and B. The conditions of digestion and results are given in Table II. These results indicate that the sequence of the carboxyl-terminal segment of T-2 is Glu-Ala-(Gln,Ser)-Lys.

In order to confirm this sequence, T-2 was digested with subtilopeptidase A. The digest was partially resolved by paper chromatography, and one peptide (R_F 0.40 in the solvent 1-butanol–glacial acetic acid–water, 4:1:5) was isolated. It was designated T-2-S-1 and it had the composition Glx, 1.95; Ala, 1.06. Two steps of Edman degradation yielded the following results. Step 1: **Glx, 1.42;** Ala, 1.00. Step 2: Glx, 1.31; **Ala, 0.33.** This, together with the results of carboxypeptidase treatment of T-2, established the sequence of T-2-S-1 to be: Glu-Ala-Gln. Edman degradation, carboxypeptidase treatment, and subtilopeptidase digestion allow alignment of 10 of the 11 residues in this peptide; the remaining aspartic acid residue is placed in the center, between glycine and glutamic acid, by difference, and the sequence of T-2 is thus: Asp-Ser-Tyr-Val-Gly-Asp-Glu-Ala-Gln-Ser-Lys.

TABLE II: Digestion of T-2 with Carboxypeptidases.

	Carboxypeptidase B, Then Carboxypeptidase A for		
	0 min	30 min	16 hr
Lysine	1.00	1.00	1.00
Serine		0.25	0.82
Glutamine		0.15	0.62
Alanine			0.25
Glutamic acid			0.16

T-3. The sequence of T-3 was determined by four steps of Edman degradation as follows. Step 1: **Gly, 0.22;** Ile, 1.01; Thr, 0.80; Leu, 1.79; Lys, —. Step 2: Gly, 0.24; **Ile, 0.12;** Thr, 1.26; Leu, 1.97; Lys, —. Step 3: Gly, 0.39; Ile, 0.15; Thr, 0.82; **Leu, 0.85;** Lys, —. Step 4: Gly, 0.27; Ile, 0.18; **Thr, 0.22;** Leu, 0.80; Lys, —. Lysine was placed at the carboxyl end because of the specificity of trypsin and the sequence of this peptide is thus: Gly-Ile-Leu-Thr-Leu-Lys.

T-4. T-4 was shown by aminopeptidase digestion to contain glutamine; two steps of Edman degradation were carried out as follows. Step 1: **Gly, 0.16;** Glx, 1.08; Lys, 1.00. Step 2: Gly, 0.08; **Glx, 0.26;** Lys, 1.00. Thus the sequence of this peptide is Gly-Gln-Lys.

T-5. T-5 was free arginine. The proof for this was that equal amounts of arginine were observed both before and after acid hydrolysis.

Alignment of the Tryptic Peptides. The five tryptic peptides from CB-10 were aligned by comparison of amino acid compositions of chymotryptic peptides from CB-10 with the amino acid sequence information that had been obtained from the tryptic peptides. CB-10 was digested with chymotrypsin (enzyme–substrate, 1:50, 16 hr) and the peptides were resolved on phosphocellulose (see Figure 4). The compositions of the chymotryptic peptides are given in Table I. The alignments may be followed by reference to Figure 3.

First, T-1 was placed at the carboxyl terminus of CB-10 because it contained homoserine. CB-10 contains one proline and one tryptophan, and both of these are in C-4; the composition of C-4 corresponds to that of T-1-C-3 plus one residue of lysine. C-3 has the composition Thr, Leu and inspection of the sequences of all of the tryptic peptides from CB-10 indicates that this peptide could only arise from T-3 by hydrolysis of the two peptide bonds that involve leucine. Cleavage at the Leu–Lys bond would leave lysine as part of C-4.

The composition of C-2 corresponds to the three residues at the amino terminus of T-3 plus eight residues from valine through lysine from T-2, and also the single residue of argi-

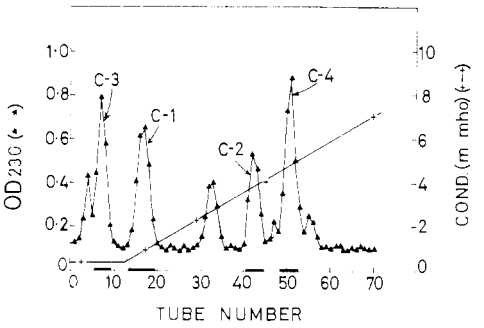

FIGURE 4: Phosphocellulose chromatography of a chymotryptic digest of CB-10. The column was run using a 1 mM, pH 3.8 sodium acetate buffer and a linear gradient of sodium chloride as described previously (Elzinga, 1970); 3-ml fractions were collected.

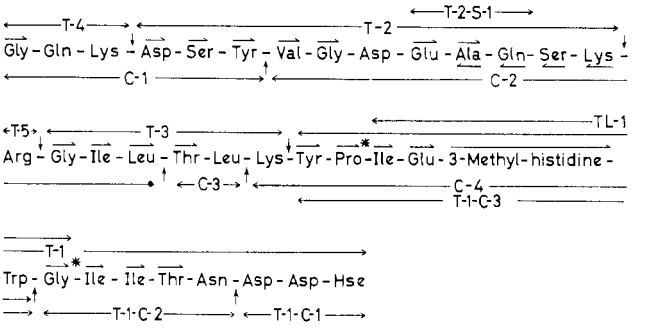

FIGURE 5: Summary of the sequence data on CB-10. A half-arrow pointing to the right indicates the residue was placed by Edman degradation; a half-arrow pointing to the left indicates that the residue was placed by carboxypeptidase digestion. An arrow pointing downward indicates a point of tryptic hydrolysis; an arrow pointing upward indicates a point of chymotryptic hydrolysis; asterisks indicate the points of hydrolysis by thermolysin that give rise to peptide TL-1.

nine. Clearly C-2 overlaps T-2 and T-3 and arginine can only be located between these two peptides. C-1 comprises T-4 and the three residues at the amino end of T-2; clearly it arises by chymotryptic hydrolysis of the Tyr–Val bond in T-2. Since all of the other peptides are aligned contiguously to make up the remainder of the peptide, C-1 is placed at the amino terminus of CB-10. Peptides representing the carboxyl-terminal region of CB-10 from glycine through homoserine were identified but, since no information regarding overlaps could be obtained from them, they were not extensively purified.

Discussion

It was the intention of this study to determine the primary structure around the single 3-methylhistidine in rabbit skeletal muscle actin, and a sequence of 35 amino acid residues, which represents one of the 17 cyanogen bromide fragments of actin, is reported. The distribution of amino acids in this peptide offers no obvious clues regarding the function of 3-methylhistidine. If the side chain of the 3-methylhistidine is involved in one of the characteristic functional properties of actin such as polymerization, nucleotide, or metal binding, actin–myosin interaction, or interaction with one or more of the other myofibrillar proteins, one would expect it to be located on the outside surface of the molecule. A segment of a polypeptide chain that lies at the surface of a protein is often characterized by the presence of charged amino acid side chains; however, the 17-residue segment from position 16 through position 32 (3-methylhistidine is at position 26) in this peptide is rather hydrophobic. Eight of the seventeen residues are isoleucine, leucine, tryptophan, or tyrosine, while the lysine at position 21 and glutamic acid at position 25 are the only residues that would carry a charge at neutral pH. This would suggest that a large part of the region is not in contact with solvent. The 3-methylhistidine residue is flanked by a glutamic acid and a tryptophan residue; one might expect the glutamic acid to be at the surface, while tryptophans are often buried in the interior of proteins. A definitive localization of this region may have to await X-ray diffraction studies although it should be possible to study the accessibility of the residues that have reactive side-chains by chemical means. Thus the presence of a tryptophan residue adjacent in the sequence to 3-methylhistidine suggests that the accessibility of this region to solvent may be studied by measuring the rate of oxidation of this tryptophan in the presence of mild oxidizing agents. A second potential approach to this region of actin may be to study the 3-methylhistidine directly by photooxidation which, under favorable conditions, results in selective oxidation of histidine residues. It is not clear whether 3-methylhistidine could be studied in this way.

In addition to providing basic information necessary to undertake studies directed toward an understanding of the function of 3-methylhistidine in actin, determination of the sequence of this peptide in actin has permitted a comparison of this peptide to the primary structure around 3-methylhistidine in myosin. As shown in the following paper (Huszar and Elzinga, 1971) the amino acid sequences around the 3-methylhistidine residues in actin and myosin are very different.

Acknowledgment

The author is grateful to Miss Marie-Anne Mattelaer for excellent technical assistance. Mr. Stephen Carabell contributed to the early phases of this project in partial fulfillment of the requirements for the degree of Bachelor of Arts with Honors in Biochemical Sciences, Harvard University, Cambridge, Mass., 1969.

Added in Proof

Johnson and Perry (1979) have recently published results of photooxidation studies on actin which suggest that 3-methylhistidine is not directly involved in the biological activities of actin.

References

Adelstein, R. S., and Kuehl, W. M. (1970), *Biochemistry 9*, 1355.

Asatoor, A. M., and Armstrong, M. D. (1967), *Biochem. Biophys. Res. Commun. 26*, 168.

DeLange, R. J., Fambrough, D. M., Smith, E. L., and Bonner, J. (1969), *J. Biol. Chem. 244*, 319.

DeLange, R. J., Glazer, A. N., and Smith, E. L. (1970), *J. Biol. Chem. 245*, 3325.

Elzinga, M. (1970), *Biochemistry 9*, 1365.

Fruton, J. S., and Simmonds, S. (1961), *in* General Biochemistry, New York, N. Y., Wiley.

Gershey, E. L., Haslett, G. W., Vidali, G., and Allfrey, V. G.

(1969), *J. Biol. Chem. 244*, 4871.
Hardy, M. F., and Perry, S. V. (1969), *Nature (London) 223*, 300.
Hirs, C. H. W. (1967), *Methods Enzymol. 11*, 325.
Huszar, G., and Elzinga, M. (1969), *Nature (London) 223*, 834.
Huszar, G., and Elzinga, M. (1971), *Biochemistry 10*, 229.
Johnson, P., Harris, C. I., and Perry, S. V. (1967), *Biochem. J. 105*, 361.
Johnson, P., and Perry, S. V. (1970), *Biochem. J. 119*, 293.
Kuehl, W. M., and Adelstein, R. S. (1969), *Biochem. Biophys. Res. Commun. 37*, 59.
Ogawa, Y., Quagliarotti, G., Jordan, J., Taylor, C. W., Starbuck, W., and Busch, H. (1969), *J. Biol. Chem. 244*, 4387.
Rees, M. K., and Young, M. (1967), *J. Biol. Chem. 242*, 4449.
Schroeder, W. A. (1967), *Methods Enzymol. 11*, 351.
Trayer, I. P., Harris, C. I., and Perry, S. V. (1968), *Nature (London) 217*, 452.

Amino Acid Sequence around the Single 3-Methylhistidine Residue in Rabbit Skeletal Muscle Myosin*

Gabor Huszar and Marshall Elzinga

ABSTRACT: As part of a study of the functional importance of methylated amino acids in myosin, a peptide that contains one residue of 3-methylhistidine has been isolated from rabbit skeletal muscle myosin. Enzymatically inactive subfragment-1 was prepared from white skeletal muscle myosin of adult rabbits by a modified procedure which utilizes ethanol precipitation. The subfragment-1 was reduced, S-alkylated using iodoacetamide, cleaved with cyanogen bromide, and finally digested with trypsin. After the chemical and enzymatic degradation the peptide mixture contained about 160 peptides; from this mixture the 3-methylhistidine peptide was isolated by a combination of gel filtration on Sephadex G-25 and Bio-Gel P-6, and ion-exchange chromatography using Dowex-50 and phosphocellulose. The amino acid sequence was determined to be: Leu-Leu-Gly-Ser-Ile-Asp-Val-Asp-3-methylhistidine-Gln-Thr-Tyr-Lys. The overall uncorrected recovery of the 3-methylhistidine peptide was about 40%. There was no indication of other peptides that contained 3-methylhistidine, or heterogeneity in the 3-methylhistidine peptide. The sequences around the single 3-methylhistidine residues in actin and myosin are different.

At least three unusual amino acids, methylated derivatives of histidine and lysine, are present in acid hydrolysates of myosin. Trayer *et al.* (1968) found about 2 moles of 3-methylhistidine [2-amino-3-(1-methyl-4-imidazolyl)propanoic acid] per 500,000 g of adult rabbit "white" skeletal muscle myosin. In addition to 3-methylhistidine, myosin contains methylated lysines (Hardy and Perry, 1969; Huszar and Elzinga, 1969a); adult rabbit white skeletal myosin contains both ϵ-*N*-monomethyllysine and ϵ-*N*-trimethyllysine in a ratio of about 1:2 (Kuehl and Adelstein, 1969).

The myosin molecule is generally believed to be composed of two heavy chains each having a molecular weight about 200,000 and two or three light chains with molecular weights of about 20,000. Limited proteolytic digestion of myosin liberates two large fragments, light meromyosin and heavy meromyosin. Further enzymatic cleavage of heavy meromyosin results in the formation of two molecules of heavy meromyosin subfragment-1 and one heavy meromyosin subfragment-2. Each subfragment-1 contains a portion of the heavy chain of myosin and one or two small subunits (Lowey *et al.*, 1969; see also recent reviews, Gergely, 1966, and Young, 1969). Both 3-methylhistidine and the methylated lysines have been localized in subfragment-1, which represents the globular part of the myosin molecule and contains both the ATPase and actin combining sites of myosin (Johnson *et al.*, 1967; Huszar and Elzinga, 1969a; Kuehl and Adelstein, 1969). The light chains of subfragment-1 were shown to contain no methylated amino acids (Huszar and Elzinga, 1969b; Kuehl and Adelstein, 1970).

Actin also contains 3-methylhistidine (Asatoor and Armstrong, 1967; Johnson *et al.*, 1967); there is one residue of the amino acid in the polypeptide chain and recent sequence studies have shown that it represents a single, fully methylated histidine (Elzinga, 1970, 1971; Adelstein and Kuehl, 1970). Thus actin and subfragment-1 of myosin, which are believed to interact during cross-bridge formation, each contain 3-methylhistidine, although there is no evidence that 3-methylhistidine is directly involved in this interaction.

The content of 3-methylhistidine has been found to vary depending upon the source of the myosin (Johnson *et al.*, 1969; Kuehl and Adelstein, 1970); "red" skeletal muscle myosin and cardiac muscle myosin contain essentially no methylated histidine. It has also been shown that myosin prepared from skeletal muscle of newborn rabbits has no 3-methylhistidine and has lower ATPase activity than adult myosin (Trayer *et al.*, 1968; Trayer and Perry, 1966). The appearance of 3-

* From the Department of Muscle Research, Boston Biomedical Research Institute and Department of Neurology, Harvard Medical School, Boston, Massachusetts 02114. *Received August 24, 1970.* This work was carried out while the Department of Muscle Research was a part of the Retina Foundation. This work was supported by grants from the National Institutes of Health (H-5949, 1-S01-FR-05527, and H-05811), the National Science Foundation, and the Medical Foundation of Boston; and was carried out during the tenure of an Established Investigatorship of the American Heart Association, Inc. (M. E.).