

Complete Amino-Acid Sequence of Actin of Rabbit Skeletal Muscle

(cyanogen bromide cleavage/maleation/myofibrillar protein/*N*^γ-methyl histidine)

MARSHALL ELZINGA*, JOHN H. COLLINS*, W. MICHAEL KUEHL†, AND ROBERT S. ADELSTEIN†

* Department of Muscle Research, Boston, Biomedical Research Institute, Boston, Massachusetts 02114 and Harvard Medical School, Boston, Massachusetts; and † Laboratory of Biochemistry, National Heart and Lung Institute, Bethesda, Maryland 20014

Communicated by C. B. Anfinsen, June 4, 1973

ABSTRACT The complete amino-acid sequence of actin of rabbit skeletal muscle was determined. The actin polypeptide chain is composed of 374 residues, including one residue of the unusual amino acid *N*^γ-methyl histidine, and has a calculated molecular weight of 41,785. The sequence of actin was determined by isolating the peptides produced by cleavage of the protein with cyanogen bromide, determining the sequence of these peptides, and establishing their order within the molecule. This study represents the first complete determination of the amino-acid sequence of a myofibrillar protein. Comparison of this sequence with peptides from actins isolated from different sources indicates that the sequence of actin is highly conserved.

Actin is a protein that plays a major role in the force-generating systems of muscle and other cells. It is a globular protein that can exist as monomeric G-actin in the absence of salt, or as a double-stranded polymer, F-actin, in the presence of physiological concentrations of salt. Actin is made up of a single polypeptide chain and one molecule each of nucleotide (ATP or ADP) and divalent cation (Mg^{2+} or Ca^{2+}).

The location and function of actin is most clearly understood in vertebrate skeletal muscle; the actin units are assembled into double-stranded helices which, together with tropomyosin and the troponin complex, constitute the thin filaments. Extensions of myosin molecules in the thick filaments can interact with actin, forming crossbridges between the filaments. During the lifetimes of attachment of these crossbridges, force is generated, which results in sliding of the filaments past each other. Thus, actin serves as the site of crossbridge attachment and also as a support for the regulatory proteins, tropomyosin and troponin.

Besides being found in skeletal, smooth, and cardiac muscle, actin-like proteins with properties that are very similar to those of muscle actin have been isolated from several other sources, including amoeba (1), blood platelets (2), fibroblasts (3), brain (4), and cytoplasm of numerous embryonic cells (5). The exact role of actin in each of these cell types is under investigation, but the ability of cytoplasmic actin to interact with myosin from muscle (1, 2, 4) implies that in nonmuscle cells it is also involved in force generation.

Actin clearly occupies a central role in biological movement, and, as a necessary step toward understanding its function on a molecular level, we have determined the amino-acid sequence of actin isolated from rabbit skeletal muscle.

Present address: W. Michael Kuehl, Albert Einstein College of Medicine, Bronx, N.Y. 10461.

Present address: Robert S. Adelstein, Cardiology Branch, National Heart and Lung Institute, Bethesda, Md. 20014.

MATERIALS AND METHODS

Actin Preparation. Rabbit back and leg muscle was used as a source of actin (6, 7). Purified actin gave a single band on Na dodecyl sulfate-gel electrophoresis. Reduction and alkylation or aminoethylation of sulfhydryl groups have been described (6, 7).

Cyanogen Bromide Cleavage. Protein or peptides were dissolved in 70% formic acid and treated with a 50- to 125-fold molar excess of cyanogen bromide (CNBr) over methionine residues for 16-24 hr at 23°. Under these conditions, the methionine was usually quantitatively converted to homoserine (6). In some cases complete conversion was insured by treatment of the protein with 2-mercaptoethanol, in order to convert residual methionine sulfoxide to methionine, and again treating the protein with cyanogen bromide (7).

Separation of CNBr Peptides. Peptides CB-1 through CB-13 (see Table 1) were first fractionated on Sephadex G-50 equilibrated with 25% acetic acid (Fig. 1a) or 20% formic acid and were then purified by ion-exchange chromatography (6-8). The sequences of the purified cyanogen bromide fragments were determined by studying the products of enzymic digestion of the peptides with trypsin, chymotrypsin, and/or thermolysin.

Peptides CB-15, CB-16, and CB-17 were separated from the other peptides as a group on Sephadex G-10 (Fig. 1b). The separation was based upon insolubility of these three peptides at pH 3-10 (see *Results*). The sequences of these peptides were determined by digesting the entire mixture with trypsin or chymotrypsin and then by digesting the partially resolved mixtures with pepsin. The products of the various enzymic digestions were purified by gel filtration and ion-exchange chromatography.

Sequence Determination. The amino-acid sequences of the peptides were determined by quantitative subtractive Edman degradations done as described (6). Assignment of asparagine and glutamine residues was made by amino-acid analysis after digestion with aminopeptidase M (6), or, in some cases, by determination of the charge on a peptide by high-voltage paper electrophoresis. The sequences of residues 238-248 and 256-268 were determined with the use of an automatic solid phase sequencer (9, 10). Identification of these residues, including amides, was based upon chromatographic mobility of the PTH derivatives (10).

Isolation of Overlap Peptides. Native G-actin was digested in the presence of 0.5 mM ATP, 0.1 mM $CaCl_2$, and 0.75 mM

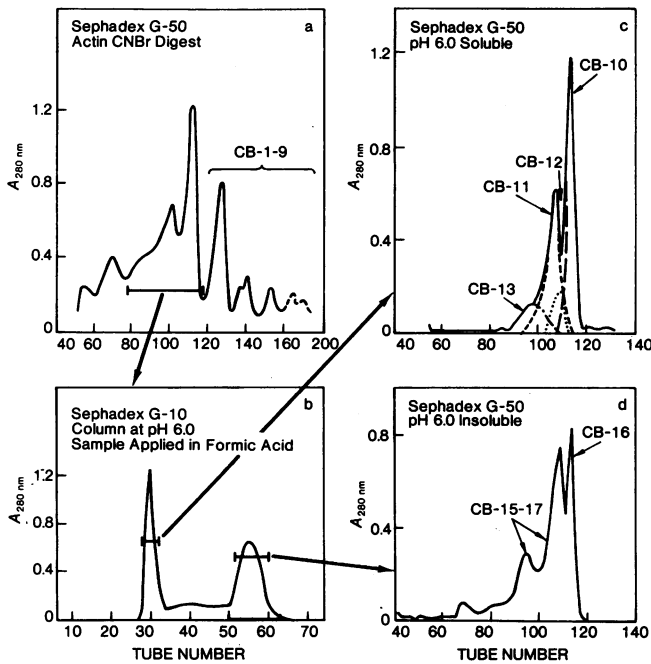


FIG. 1. Gel filtration of the cyanogen bromide peptides of actin. (a) Entire digest on Sephadex G-50 (25% acetic acid, 1.9 \times 400 cm). CB-1-CB-9 were purified by ion-exchange chromatography (not shown). The region enclosed by the bar was pooled, dried, dissolved in 5 ml of 70% formic acid, and applied to 1.9 \times 200-cm column of Sephadex G-10 (b). Rechromatography of the G-10 peaks under conditions described for a is shown in c and d (22).

2-mercaptoethanol for 19 hr with 1% trypsin at 23°. The digest was then lyophilized, dissolved in 10 M urea, and digested for an additional 24 hr in 2 M urea. After the second digestion the material was reduced with 2-mercaptoethanol and alkylated in 10 M urea with iodoacetamide (7). Two groups of peptides resulted, those soluble in 10% formic acid and those soluble only in 50–100% formic acid.

Methionine-containing peptides were isolated from the soluble fraction by gel filtration, ion-exchange chromatography, paper chromatography, and paper electrophoresis. The peptides that contained methionine were identified with a platonic iodide spray on paper (11). The fraction that was soluble only in 50–100% formic acid yielded two methionine peptides; these were purified by repeated gel filtration on Sephadex G-50 (fine) in 50% formic acid.

Tryptic Digestion of Maleated Actin. Carboxymethylated actin was treated with powdered maleic anhydride in 70 mM NaBO₃ (pH 8–8.6) (12). The extent of maleation was quantitative, as monitored by dansylation of the product. Since maleated amino groups do not react with dansyl chloride, recovery of lysine after acid hydrolysis and amino-acid analysis afforded a quantitative measure of the extent of maleation of the actin. The maleated actin was then digested with trypsin (digestion occurring only at the arginine residues), and the tryptic peptides were separated by (i) DEAE-cellulose chromatography with a linear pyridine-acetic acid gradient of decreasing pH and increasing ionic strength, (ii) paper electrophoresis (pH 8.6), and (iii) descending paper chromatography.

Chymotryptic Digestion of Actin. Direct proof for the peptide sequence CB-11-CB-14 was obtained by isolation of a chromatryptic peptide from actin that contained residues 106–127. Performic acid-oxidized actin (250 mg) was digested with 5 mg of α -chymotrypsin for 4 hr at 23°. The digest was chromatographed on Sephadex G-50, SP-Sephadex G-25, Bio-Gel P-6, and again on SP-Sephadex.

RESULTS

Characterization of the Cyanogen Bromide Peptides. The initial step in determining the amino-acid sequence of actin was cleavage of the polypeptide chain with cyanogen bromide. Since actin contains 16 methionine residues, 17 CNBr peptides were expected. Thirteen of the CNBr peptides, CB-1 through CB-13 (see Table 1), were readily soluble at all pH values and were purified in desired quantities by gel filtration and ion-exchange chromatography. Their sequences were determined by appropriate enzymic digestions with trypsin, chymotrypsin, and thermolysin, followed by subtractive Edman degradation of the resulting peptides.

Peptides CB-15, CB-16, and CB-17 presented special difficulties. Although they had been isolated and characterized (7), the low yields and uncertain purity made sequence studies on these preparations difficult. These peptides were insoluble in the pH range 3–10, and aggregated during gel filtration in 25% acetic acid. These properties essentially precluded purification by ion-exchange chromatography under the conditions previously used to purify other peptides of similar size, i.e., CB-10 through CB-13.

The method finally used was to separate the "insoluble" peptides from the others, using a Sephadex G-10 column equilibrated and run with a pH 6.0, 0.01 M pyridine-acetic acid buffer. The fraction that contained CB-10 through CB-17 was dissolved in 70% formic acid, and the insoluble peptides were "washed off" the column by the band of formic acid. Rechromatography of the mixture on Sephadex G-50 yielded pure CB-16 (less than 10% contamination as determined by amino-acid analysis), and a mixture of CB-15 and CB-17 (see Fig. 1). Yields of these peptides were greater than 70%.

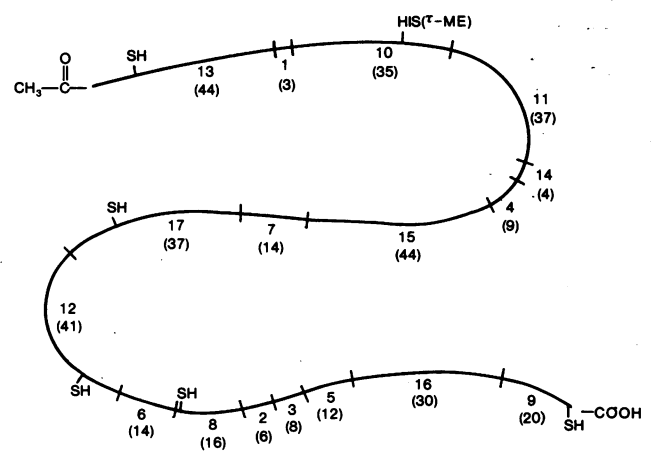


FIG. 2. Ordering of the cyanogen bromide peptides. *Upper numerals* designate the cyanogen bromide peptide number (see Table 1) and the numerals in *parentheses* indicate the number of residues in each peptide. *Vertical lines* drawn through the polypeptide chain indicate the location of the methionine residues. The appropriate positions of the 5-SH groups and *N*⁷-methyl histidine are shown.

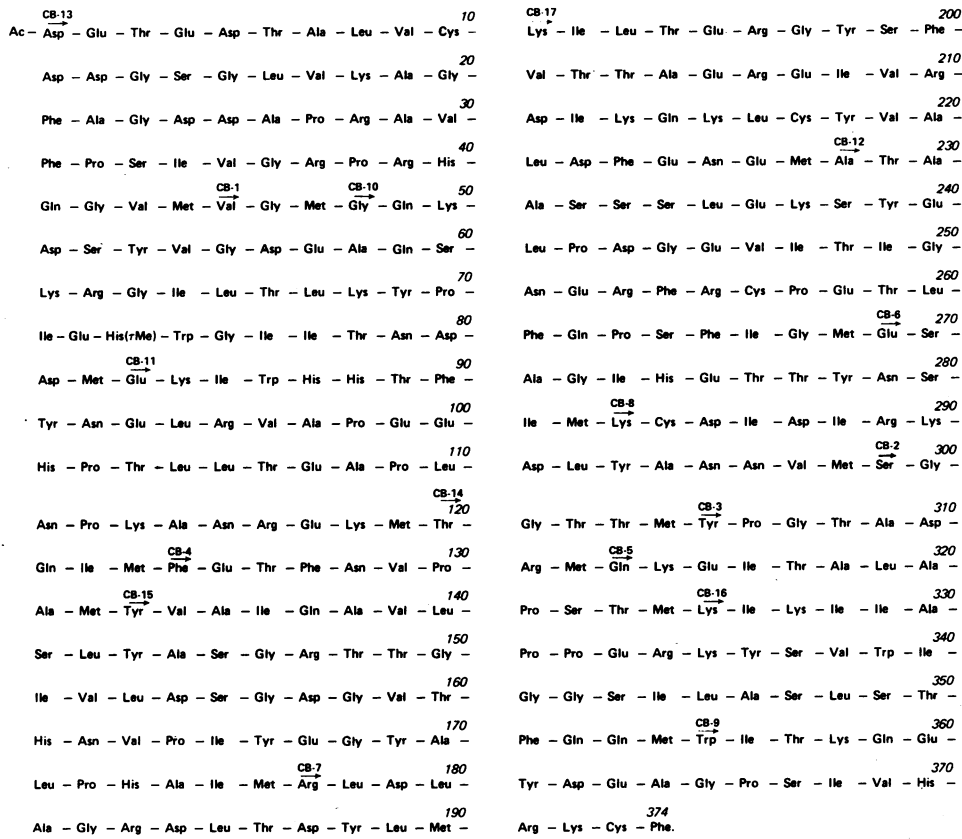


Fig. 3. Amino-acid sequence of actin of rabbit skeletal muscle.

Tryptic, chymotryptic, and peptic peptides of CB-15, CB-16, and CB-17 were generated by digesting mixtures of all three peptides, CB-15 plus CB-17, or CB-16 with the appropriate enzymes. A total of 68 proteolytic peptides were purified from

these digests. The sequences or partial sequences of many of these were determined, and by inspection of the sequences or compositions of overlapping peptides, unequivocal sequences were deduced for CB-15, CB-16, and CB-17.

TABLE 1. Amino-acid composition of the cyanogen bromide peptides of actin

	CB-1	CB-2	CB-3	CB-4	CB-5	CB-6	CB-7	CB-8	CB-9	CB-10	CB-11	CB-12	CB-13	CB-14	CB-15	CB-16	CB-17	Total Residues
Lys					1			2	2	3	3	1	1			3	3	19
His						1			1		3		1		2			8
His-(rMe)										1								1
Arg			1				2	1	1	1	2	2	3		1	1	3	18
Cys								1	1			1	1				1	5
Asp*			1	1		1	3	5	1	5	3	2	6		3		3	34
Thr		2		1	2	2	1		1	2	3	3	2	1	3	1	3	27
Ser		1			1	2			1	2		5	2		3	4	1	22
Glu†				1	2	2			3	4	6	6	3	1	2	3	6	39
Pro			1	1	1				1	1	4	3	3		2	2		19
Gly	1	2	1			1	1		1	4		3	6		5	2	1	28
Ala			1	1	2	1	1	1	1	1	3	3	5		5	2	2	29
Val	1			1				1	1	1	1	1	5		5	1	3	21
Ile			1		1	2		2	2	4	1	3	1	1	4	5	3	30
Leu					1		4	1		2	4	3	2		4	2	3	26
Tyr			1			1	1	1	1	2	1	1			4	1	2	16
Phe				2					1		1	3	2			1	2	12
Trp									1	1	1							4
Hse	1	1	1	1	1	1	1	1		1	1	1	1	1	1	1	1	16
No. of Residues	3	6	8	9	12	14	14	16	20	35	37	41	44	4	44	30	37	374
N-Terminal Residue	Val	Ser	Tyr	Phe	Gln	Glu	Arg	Lys	Trp	Gly	Glu	Ala	CH ₃ C(=O)Asp	Thr	Tyr	Lys	Lys	

* Includes 11 residues of Asn. † Includes 11 residues of Gln.

The remaining CNBr peptide, CB-14, had an N-terminal threonine and was, therefore, not quantitatively released after initial cyanogen bromide treatment. It apparently remained attached to the C-terminus of CB-11 during the initial G-50 fractionation. It was obtained pure, in about 35% yield, by lyophilizing the region that contained CB-11, suspending this fraction in dilute acetic acid, briefly exposing it to pH 12.1 by titrating the solution with KOH, and finally lowering the pH to 4.4 by addition of glacial acetic acid. The insoluble peptides (CB-15–CB-17) were removed by centrifugation, and the supernatant was passed over a 400-cm Sephadex G-50 column. The tetrapeptide, CB-14, was found in the anticipated position.

Ordering of the Cyanogen Bromide Peptides (Fig. 2). In order to establish the arrangement of the CNBr peptides in the intact polypeptide chain, the "overlap" or methionine-containing peptides were isolated from a tryptic digest of actin and characterized. In general, cyanogen bromide cleavage of the tryptic peptides yielded fragments that could be identified as arising from one of the CNBr peptides. In some cases it was necessary to isolate either the amino-terminal or carboxyl-terminal tryptic peptide of the original CNBr peptide in order to identify the overlap. Because lysine was the amino terminus of three of the seventeen CNBr peptides, not all of the overlaps could be resolved by tryptic digestion of the unmodified protein. This problem was resolved by treating actin with maleic anhydride before tryptic digestion in order to prevent hydrolysis at the lysine residues. Characterization of the resulting overlap peptides, which were completely soluble at neutral pH, permitted resolution of the ambiguous overlaps. The sequence CB-11–CB-14 (overlap of CB-11 and CB-14) was established by isolation from performic acid-oxidized actin of a chymotryptic peptide that, by composition and tryptic subpeptides, was identified as residues 106 through 127.

DISCUSSION

The complete amino-acid sequence of actin of rabbit skeletal muscle is presented in Fig. 3. Proof that the sequence is complete rests primarily upon the fact that analysis of the methionine-containing peptides permits an unambiguous alignment of the 17 CNBr peptides into a polypeptide chain that has a blocked (acetylated) N-terminus and a C-terminal phenylalanine (14). The chain contains 374 residues, and has a calculated molecular weight of 41,785.

The molecular weight of 41,785 (or 42,300 including Ca^{2+} and ATP) is slightly lower than has been estimated from its thermodynamic properties (15). Although the presence of an additional prosthetic group besides ATP cannot be entirely ruled out, we think that the 5–10% higher molecular weight found by ultracentrifugation reflects errors inherent in the technique.

The sequence contains several interesting features, although in the absence of information about the tertiary structure of actin any assignment of functional roles to individual residues is premature. There are some unusually highly charged regions, notably the acetylated N-terminus where four of the first five residues contain sidechain carboxyl groups. The first 25 residues carry a net negative charge of seven (at pH 7.0).

There are five free sulfhydryl residues, at positions 10, 217, 256, 284, and 373. The results of Bridgen (16) on trout actin indicate that residues 10, 284, and 373 react rapidly with

iodoacetamide but that only residue 10 reacts rapidly with iodoacetate. Residue 373 reacts first with *N*-ethylmaleimide. Residue 217 is apparently not available for reaction in the native molecule. The four chymotryptic peptides isolated by Lusty and Fasold from actin of rabbit skeletal muscle (17) are easily identified in the sequence; of particular interest is residue 373, a cysteine whose reactivity is reduced when actin interacts with myosin (17).

*N*⁷-Methyl histidine is an unusual amino acid, found only in actin and myosin (18); its biological function is obscure. It is probably synthesized by enzymic methylation of the histidine at position 73 (13, 19). The possibility that it is accessible to an enzyme after synthesis of the protein suggests that this residue is in contact with solvent. Tyrosine 69, which is in the same region, is the first tyrosine to be nitrated when actin is treated with tetranitromethane (20).

Since actin, in a form that interacts with rabbit myosin, has been identified in a wide variety of cell types and a range of organisms from ameba to humans, it is of interest to examine the relationships among actins from different sources. Bridgen (21) has characterized four –SH peptides from trout actin, and they are very similar in sequence to those from rabbits. For instance, the carboxyl terminal 13 residues are identical in the two proteins, with the exception of one Leu → Ile replacement. The amino-terminal 18-residue peptide from trout is identical in composition to the rabbit peptide, and there are two other regions, one of ten residues and the other of 14, in which there are no more than one or two substitutions. Thus, in peptides representing 55 residues, there appear to be only 3–5 substitutions. Weihing and Korn (1) have isolated and analyzed three cyanogen bromide peptides from ameba actin; although the ameba peptides are not completely pure, the compositions are strikingly similar to the rabbit peptides. These structural comparisons corroborate what was suspected from consideration of the biochemical properties of actin—that the protein is highly conserved throughout nature.

The study described here represents the first determination of the complete amino-acid sequence of a myofibrillar protein. In its functional role in muscle (and presumably in other cells), actin monomers interact with other proteins in a highly specific fashion, and apparently strong (noncovalent) bonds are formed. These strong specific interactions include actin–actin interaction to form actin helical polymers, which are the backbone of the thin filaments, interaction with myosin during the force-generating step, and interaction with the regulatory proteins tropomyosin and troponin. It may be expected that knowledge of the complete amino-acid sequence of actin will facilitate an understanding of these interactions and their role in muscle contraction.

We thank Mary Anne Conti, Nora Jackman, and Gedelita Wilshire for excellent technical assistance. R.S.A. and W.M.K. thank Dr. W. W. Kielley for helpful discussions. M.E. is an Established Investigator of the American Heart Association. This work was supported in part by grants from the National Institutes of Health (AM-14728) and the American Heart Association.

1. Weihing, R. R. & Korn, E. D. (1972) *Biochemistry* **11**, 1538–1543.
2. Adelstein, R. S. & Conti, M. A. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 599–605.
3. Yang, Y.-Z. & Perdue, J. F. (1972) *J. Biol. Chem.* **247**, 4503–4509.

4. Berl, S., Puszkin, S. & Nicklas, W. J. (1973) *Science* **179**, 441-446.
5. Bray, D. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 567-571.
6. Elzinga, M. (1970) *Biochemistry* **9**, 1365-1374.
7. Adelstein, R. S. & Kuehl, W. M. (1970) *Biochemistry* **9**, 1355-1364.
8. Elzinga, M. (1971) *Biochemistry* **10**, 224-229.
9. Laursen, R. A. (1971) *Eur. J. Biochem.* **20**, 89-102.
10. Laursen, R. A., Horn, M. J. & Bonner, A. G. (1972) *FEBS Lett.* **21**, 67-70.
11. Bennett, J. C. (1967) in *Methods in Enzymology*, ed. Hirs, C.H.W. (Academic Press, New York), pp. 330-339.
12. Butler, P. J. G., Harris, J. I., Hartley, B. S. & Leberman, R. (1969) *Biochem. J.* **112**, 679-689.
13. Asatoor, A. M. & Armstrong M. D. (1967) *Biochem. Biophys. Res. Commun.* **26**, 168-174.
14. Johnson, P. & Perry, S. V. (1968) *Biochem. J.* **110**, 207-216.
15. Rees, M. K. & Young, M. (1967) *J. Biol. Chem.* **242**, 4440-4458.
16. Bridgen, J. (1972) *Biochem. J.* **126**, 21-25.
17. Lusty, C. J. & Fasold, H. (1969) *Biochemistry* **8**, 2933-2940.
18. Vickery, H. B. (1972) *Advan. Protein Chem.* **26**, 82-172.
19. Hardy, M. F. & Perry, S. V. (1969) *Nature* **223**, 300-302.
20. Lehrer, S. S. & Elzinga, M. (1972) *Fed. Proc.* **31**, 1627.
21. Bridgen, J. (1971) *Biochem. J.* **123**, 591-600.
22. Elzinga, M. & Collins, J. H. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 1-7.