Proc. Natl. Acad. Sci. USA
Vol. 74, No. 10, pp. 4281–4284, October 1977
Biochemistry

Amino acid sequence of a myosin fragment that contains SH-1, SH-2, and N^{τ}-methylhistidine

(myofibrillar protein/sulfhydryl groups)

MARSHALL ELZINGA*†‡§ AND JOHN H. COLLINS†‡¶

* Biology Department, Brookhaven National Laboratory, Upton, Long Island, New York 11973; † Department of Muscle Research, Boston Biomedical Research Institute, and ‡ Department of Neurology, Harvard Medical School, Boston, Massachusetts 02114

Communicated by R. B. Setlow, July 25, 1977

ABSTRACT **A peptide having 92 amino acid residues and a calculated molecular weight of 10,478 was isolated from a cyanogen bromide digest of rabbit skeletal muscle myosin. It contained both proline and N^{τ}-methylhistidine, indicating that it arose from the portion of the heavy chain that folds to form most of the globular head of the myosin molecule. The amino acid sequence of the peptide included the two sulfhydryl groups whose alkylation modifies myosin's catalytic properties: SH-2 at position 11 in the peptide, and SH-1 at position 21. This proximity in the sequence means that SH-1 and SH-2 must be relatively close together in myosin, and several lines of evidence suggest that this region is near the catalytic or actin binding site(s) of myosin.**

Myosin is a large multisubunit protein, each molecule of which contains two similar heavy chains [molecular weight (M_r) about 200,000] and two each of two types of light chains (M_r about 20,000). The NH_2-terminal region of each heavy chain, along with the light chains, forms a globular "head," called subfragment 1 or S-1 (1) which is an enzyme with ATPase activity. The remainder of the molecule is almost entirely α-helical, and the two chains form a double-stranded coiled coil (the "tail") about 140 nm long. This tail can be divided into two segments. The COOH-terminal part (called light meromyosin or LMM) (2) is insoluble at physiological ionic strength and aggregates with other myosin tails to form the main shaft of the thick filament of muscle. The region that connects subfragment 1 and light meromyosin is called subfragment 2 or S-2 (3), and it seems to tether the head to the main shaft of the thick filament. The various fragments can be prepared by judicious use of proteolytic enzymes.

In carrying out its role as a force-generating transducer in muscle, the head of myosin can form a crossbridge between the thick and thin filaments by attaching to an actin molecule; it then probably rotates at the actin–myosin interface in such a way that the force thus generated propels the thick and thin filaments in opposite directions. The reaction between actin and myosin is cyclical, and each event is accompanied by the hydrolysis of one molecule of ATP. It now seems likely that variations of this system are also responsible for force generation in nonmyofibrillar systems in a wide variety of cells (4).

In order to understand the details of the enzymic properties and protein–protein interactions that are the bases for the functional properties of myosin, it is essential to describe this molecule in detail. The amino acid sequences of the light chains of myosin from rabbit skeletal muscle have been determined (5, 6) but equivalent information about the heavy chain is fragmentary. As part of a study in which we hope eventually to elucidate the complete sequence of the approximately 1700 amino acid residues in the myosin heavy chain, we report here the sequence of a 92-residue CNBr peptide that arises from the head of the myosin molecule. This peptide is of special interest because it contains the unusual amino acid N^{τ}-methylhistidine; skeletal muscle myosin is known to contain a single residue of this amino acid (7, 8), and the sequence of a 13-residue tryptic peptide containing N^{τ}-methylhistidine has previously been reported (9). The 92-residue peptide, which shall be referred to as p10, also contains two cysteinyl residues, and comparison of the sequences around them with shorter cysteinyl-containing peptides studied by others indicates that they represent the sulfhydryl groups SH-1 and SH-2 whose selective alkylation under defined conditions leads to well-characterized changes in the enzymic properties of myosin. SH-1 and SH-2 are only 10 residues apart in the sequence, confirming that they must be close together in the myosin head, and chemical modification work by others suggests that the region that contains SH-1 and SH-2 is at or near an ATP binding site and possibly the actin binding site as well in myosin (10–12).

METHODS

To prepare the peptide, 10 g of myosin was isolated from the back and leg muscles of New Zealand white rabbits as described (9). The myosin was reduced and alkylated with iodoacetate or iodoacetamide at reagent-to-protein ratios similar to those used for actin (9, 13). Reagents were removed by dialysis, and the protein was lyophilized, suspended in 70% (vol/vol) formic acid at a concentration of about 50 mg/ml, and treated with CNBr (1 g of CNBr per g of myosin) for 16 hr at 23–25 °. The viscous, cloudy suspension cleared after 1–2 hr of CNBr treatment. The solution was then diluted by addition of an equal volume of water, and the CNBr and much of the formic acid were removed by evaporation under reduced pressure to a volume of about 20 ml. The solution was diluted to 100 ml by addition of water, evaporated to 20 ml, diluted again to 100 ml with water, and lyophilized. The powder was then extracted twice with 100 ml of 25% (vol/vol) acetic acid, with centrifugation after each extraction. The supernatants were combined

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: M_r, molecular weight; p10, 92-residue CNBr peptide from head of myosin molecule.

§ To whom reprint requests should be addressed at Brookhaven National Laboratory.

¶ Present address: Department of Pharmacology and Cell Biophysics, University of Cincinnati, College of Medicine, 231 Bethesda Avenue, Cincinnati, OH 45267.

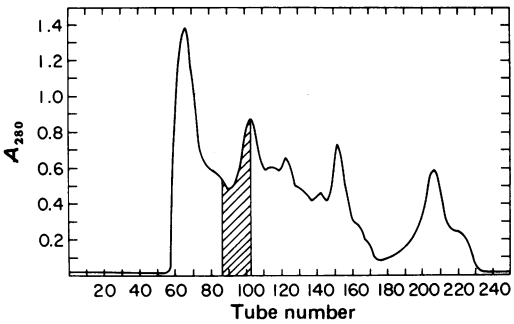

FIG. 1. Elution profile of the CNBr digest of rabbit skeletal muscle myosin gel filtered on a 5 × 150 cm column of Sephadex G-50 (fine). Flow rate, 75 ml/hr; temperature, 23°; fractions, 8 ml.

and divided into two equal fractions that were gel filtered sequentially on a 5 × 150 cm column of Sephadex G-50 (fine) in 25% acetic acid. A typical elution pattern is shown in Fig. 1.

The fraction identified by crosshatching was pooled, dried by rotary evaporation, and dissolved in 20 ml of water. Because of residual acetic acid in the sample, the pH of this solution was between 3 and 4. The solution was then titrated to pH 7.0 by dropwise addition of 1.0 M NaOH, which produced a heavy precipitate. The solution was centrifuged, and the precipitate, which contained p10, was dissolved in a minimum volume (usually 2–3 ml) of 25% acetic acid. This solution was then gel filtered on a 1.9 × 200 cm column of Sephadex G-50 (fine); the effluent was monitored by (*i*) measuring absorbance at 280 nm and (*ii*) subjecting a small aliquot (5–7 μl) from each tube to sodium dodecyl sulfate/polyacrylamide gel electrophoresis by the method of Swank and Munkres (14). The tubes that, by visual examination of the stained gels, appeared to be at least 90% pure p10 were pooled (Fig. 2) and used for subsequent studies.

The amino acid sequence was determined by automated Edman degradation of the intact peptide in a Beckman 890C sequencer (Beckman Instruments, Palo Alto, CA), combined with studies on tryptic and chymotryptic peptides with both the Beckman sequencer and a solid-phase (15, 16) sequencer (Sequemat, Inc., Watertown, MA). The derivatized amino acids obtained from the sequencers were identified by thin-layer chromatography (17) and amino acid analysis after regeneration of the amino acid by HI hydrolysis (18). Amides were assigned based upon thin layer chromatography of the derivatized amino acids and/or amino acid analysis after enzymic (aminopeptidase M) digestion. Phenylthiohydantoin-carboxymethylcysteines were identified by a combination of thin layer chromatography, in which they migrate with the same R_F as phenylthiohydantoin-aspartic acid, and HI hydrolysis which yields alanine. (A detailed description of the determination of the amino acid sequence of p10 will be published elsewhere.)

RESULTS AND DISCUSSION

The amino acid sequence of p10 is shown in Fig. 3. Numbering of the residues is consecutive from the NH_2 terminus of the peptide and does not reflect the positions of these residues in the intact myosin heavy chain. Sixty-three of the 92 residues are uncharged, and 32 of these have hydrophobic side chains. The peptide is basic, with a net charge of +7 at neutral pH. The N^{τ}-methylhistidine is at position 69 and the two cysteinyl residues are at positions 11 and 21.

It has long been recognized (19–21) that alkylation of myosin sulfhydryl groups modifies the catalytic properties of the enzyme, and studies by many investigators (22–28) have indicated that the observed effects accompany rather specific alkylation of two sulfhydryl groups. Sulfhydryl group SH-1 can be alkylated in the absence of bound nucleotides, with the result that the Ca^{2+}-ATPase activity is stimulated, and the EDTA(K^+)-ATPase activity is lost. In the presence of ADP, sulfhydryl group SH-2 also reacts, and myosin alkylated at both SH groups is devoid of ATPase activity. The sulfhydryl groups have been labeled with radioactive alkylating agents and, after enzymic digestion of the myosin, the labeled peptides have been isolated and sequenced. Yamashita *et al.* (28) have summarized the studies on the peptides that contain SH-1 and SH-2 and have concluded that the sequence around SH-2 is Arg-Cys-Asn-Gly-Val-Leu, and the sequence around SH-1 is Glu-Gly-Ile-Arg-Ile-Cys-Arg. The sequences correspond exactly to the sequence of residues 10–15 and 16–22, respectively, of p10. These facts are summarized in Fig. 4, and it seems clear that residue 11 represents SH-2 and residue 21 is SH-1.

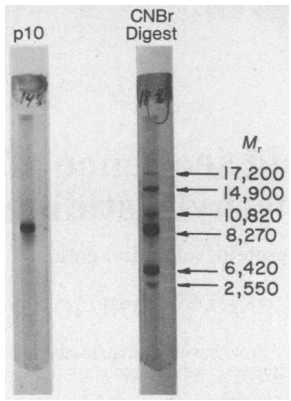

FIG. 2. Polyacrylamide gel electrophoresis of p10 (*Left*) and the peptide mixture that is generated by CNBr digestion of sperm whale apomyoglobin (*Right*). Gels contained 12.5% acrylamide and were prepared at a bis-acrylamide/acrylamide ratio of 1:10 (14). The calculated M_rs of the myoglobin fragments are listed at the right.

Although there is no direct evidence that either SH-1 or SH-2 is directly involved in actin or nucleotide binding (29), the evidence mentioned above suggests that they are at or near the catalytic site for myosin ATPase. It seems likely that a nucleotide binding site would contain both a hydrophobic pocket for the adenine moiety and a basic residue(s), probably arginine (30), with which a phosphate group(s) could interact electrostatically. The arginine residues at positions 10, 19, and 22 all are close to either SH-1 or SH-2, providing possible sites for binding of phosphates. The segment extending from residues 4 through 31 has a high concentration of hydrophobic sidechains (12 of 28) which could form the appropriate hydrophobic pocket. One scheme that seems consistent with the observed patterns of labeling is that SH-1 exists in a microenvironment that causes it to have a low pK_a (and thus high reactivity to alkylating agents); indeed, Takamori *et al.* (31) have presented evidence that the pK_a of SH-1 is about 6.3, and Schliselfeld (32) has observed that an ionizable group having a pK_a of about 6.5 is involved in the binding of ATP to myosin. When nucleotide is bound, the adenine moiety could bind to a site that includes some of the hydrophobic sidechains in the vicinity of SH-1, and a phosphate group(s) could interact with arginine 10. The proximity of the negatively charged phosphate group(s) could then shift the pK of SH-2 down, increasing its reactivity to alkylating agents.

This scheme is not inconsistent with the proposals of Harrington and his colleagues that SH-1 and SH-2 are 12–14 Å apart in native myosin (33) and that both are involved in

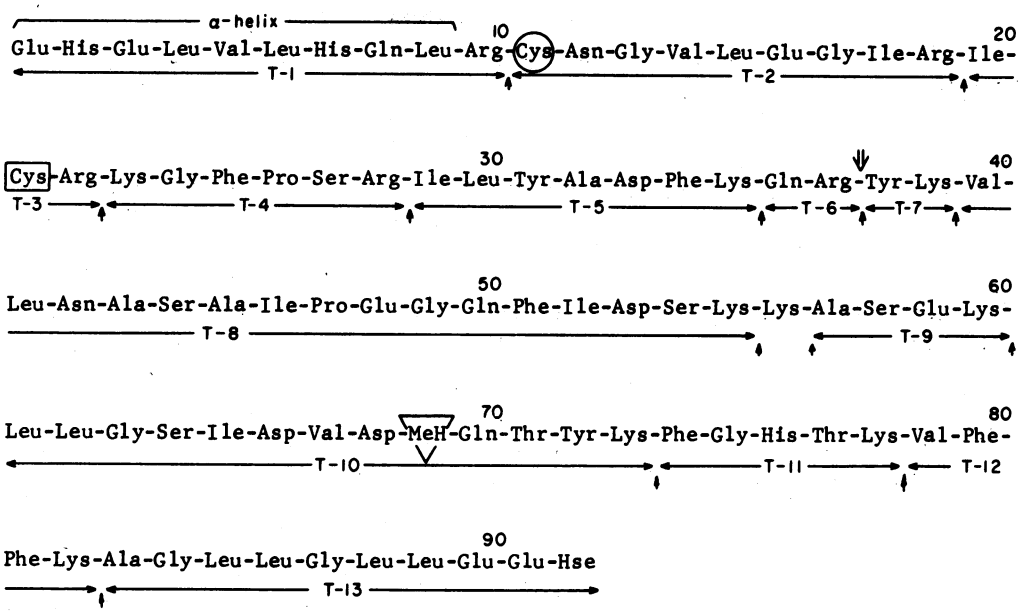

FIG. 3. The amino acid sequence of p10. SH-2 is enclosed by a circle, SH-1 by a rectangle, and N^{τ}-methylhistidine by a triangle. Upward pointing arrows designate points of tryptic digestion of unmodified p10; the arrow between Arg^{37} and Tyr^{38} designates the point of tryptic hydrolysis of citraconylated p10 that gives rise to a fragment (containing residues 38–92) that was useful for the determination of the sequence of the central portion of the peptide. Residues 1–9 are predicted to exist as an α-helix by application of structure prediction rules (33). The amino acid composition of p10 is listed in Table 1.

binding of Mg^{2+}-ATP at the catalytic site (34). Application of the rules of Chou and Fasman (35) leads to the prediction that the region between SH-1 and SH-2 is not an α-helix, is not part of a β-pleated sheet, and does not form a β-bend. Thus, the region could exist as a flexible loop with the two SH groups 12–14 Å apart. Studies on the conformation and nucleotide-binding capacity of p10 may help to define the role of this region in myosin function.

Our p10 is the longest continuous sequence from the myosin heavy chain that has been reported to date. The second longest is that of an 18-residue cysteine-containing peptide reported by Weeds and Hartley (36): Val-Arg-Cys-Ile-Ile-Pro-Asn-Glu-Thr-Lys-Thr-Pro-Gly-Ala-Met-Glu-His-Glu. There is a four-residue overlap (Met-Glu-His-Glu) between the COOH terminus of Weeds and Hartley's peptide and the NH_2 terminus of our p10, and these peptides very likely constitute a 107-residue segment of the heavy chain.

Although microheterogeneity of the heavy chains has been documented (37, 38), the apparent absence of microheterogeneity in this 107-residue segment suggests that microheterogeneity is not extensive and that the two (or more) heavy chains in a given preparation of rabbit skeletal muscle myosin are identical in amino acid sequence over large parts of the molecule.

Rabbit skeletal muscle myosin contains one residue of N^{τ}-methylhistidine, and it appears as residue 69 in this peptide. The tryptic peptide that represents residues 61 through 73 was isolated and sequenced previously (9), and the present results confirm the findings of that study. Because the methylhistidine is the product of post-translational enzymic methylation of histidine (39, 40), it seems likely that the region surrounding it is at or near the surface of myosin. The presence of this residue in the same peptide as SH-1 and SH-2 does not require that they be in the same region of the myosin molecule, because the 48 residues between them could span the entire myosin head.

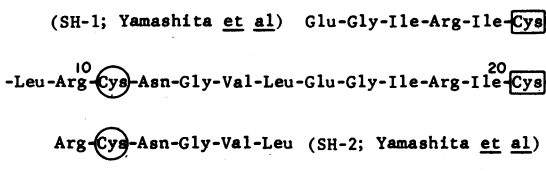

FIG. 4. Comparison of the sequences of proteolytic fragments that contain SH-1 and SH-2, described by Yamashita *et al.* (25), with the sequence of residues 9–23 of p10 (middle line). SH-2 is circled and SH-1 is enclosed by a rectangle.

Two other studies that involved the characterization of myosin peptides should be mentioned. Kubo *et al.* (41) isolated trinitrophenyl (TNP) peptides from subtilisin and Pronase digests of myosin that had been treated with trinitrobenzenesulfonate and found Ser-TNPLys-(Gly, Glu, Ser)-(Gly, Ala) and Asn-Pro-Pro-TNPLys. These peptides are not present in p10.

Table 1. The amino acid composition of p10

Amino acid	Residues
Cysteine	2
Aspartic acid	4
Asparagine	2
Threonine	2
Serine	5
Glutamic acid	7
Glutamine	4
Proline	2
Glycine	8
Alanine	5
Valine	5
Isoleucine	6
Leucine	12
Tyrosine	3
Phenylalanine	6
Lysine	9
Histidine	3
N^{τ}-Methylhistidine	1
Arginine	5
Homoserine	1
Total	92

Barany and Merrifield (42) synthesized an octapeptide that they suggested may be related in sequence to a region of the catalytic site of frog myosin. The sequence of this peptide was Thr-Ala-Cys-Gly-Gln-Lys-Ser-Pro, and it does not resemble any part of the sequence of p10. The peptides studied by both groups of investigators could of course be near the region occupied in p10 in the tertiary structure of myosin and yet not be contiguous in the primary structure.

Although an understanding of the relationship between the segment described in this paper and the intact head of myosin must await other analyses, the availability of this fragment should make possible several types of study. First, this peptide provides a handle for an interesting part of the myosin molecule, and comparative studies on equivalent segments of other myosins may be instructive in identifying, for example, residues that are essential for nucleotide binding. Second, chemical modification experiments can be directed at potentially interesting residues, including the sulfhydryl groups, and by isolating and characterizing this fragment after modification of residues that it contains it will be possible to evaluate with precision the sites or extents of reaction. Finally, this peptide may be useful in evaluating the existence of altered forms of myosin, particularly in the heart. For example, if in cardiac hypertrophy an "abnormal" myosin is synthesized as a result of activation of a normally inactive gene, the new myosin would probably differ in sequence from the normal myosin at several positions; the peptide described here is relatively easy to purify, seems to represent an important part of the molecule, and is large enough to permit the identification of even relatively infrequent differences in sequence between the normal and abnormal myosins.

The authors acknowledge the excellent contributions of Debbie Duffy Hight, Erlinda Capuno, John J. Collins, and Virginia Hayssen to various aspects of this work. The project was supported by Grants HL-17464 and HL-21471 from the National Institutes of Health, and PCM75-15313 from the National Science Foundation. The research carried out at Brookhaven National Laboratory was under contract with the U.S. Energy Research and Development Administration.

1. Mueller, H. & Perry, S. V. (1962) *Biochem. J.* **85**, 431–439.
2. Szent-Györgyi, A. G., Cohen, C. & Philpott, D. E. (1960) *J. Mol. Biol.* **2**, 133–142.
3. Lowey, S., Goldstein, L., Cohen, C. & Luck, S. M. (1967) *J. Mol. Biol.* **23**, 287–304.
4. Pollard, T. D. & Weihing, R. R. (1974) *Crit. Rev. Biochem.* **2**, 1–65.
5. Frank, G. & Weeds, A. G. (1974) *Eur. J. Biochem.* **44**, 317–334.
6. Collins, J. H. (1976) *Nature* **259**, 699–700.
7. Trayer, I. P., Harris, C. I. & Perry, S. V. (1968) *Nature* **217**, 452–453.
8. Asatoor, A. M. & Armstrong, M. E. (1967) *Biochem. Biophys. Res. Commun.* **26**, 168–174.
9. Huszar, G. & Elzinga, M. (1972) *J. Biol. Chem.* **247**, 745–753.
10. Barany, M., Bailin, G. & Barany, K. (1969) *J. Biol. Chem.* **244**, 648–657.
11. Schaub, M. C. & Watterson, J. G. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 153–156.
12. Seidel, J. C. (1973) *Arch. Biochem. Biophys.* **157**, 588–596.
13. Elzinga, M. (1970) *Biochemistry* **9**, 1365–1374.
14. Swank, R. T. & Munkres, K. D. (1971) *Anal. Biochem.* **39**, 462–477.
15. Laursen, R. A. (1971) *Eur. J. Biochem.* **20**, 89–102.
16. Horn, M. J. & Laursen, R. A. (1973) *FEBS Lett.* **36**, 285–288.
17. Bridgen, J. (1975) *FEBS Lett.* **50**, 159–162.
18. Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. & Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912–4921.
19. Kielley, W. W. & Bradley, L. B. (1956) *J. Biol. Chem.* **218**, 653–659.
20. Sekine, R. & Kielley, W. W. (1964) *Biochim. Biophys. Acta* **81**, 336–345.
21. Yamaguchi, M. & Sekine, R. (1966) *J. Biochem.* **59**, 24–33.
22. Sekine, R., Barnett, L. M. & Kielley, W. W. (1962) *J. Biol. Chem.* **237**, 2569–2772.
23. Gröschel-Stewart, U. & Turba, F. (1963) *Biochem. Z.* **337**, 109–114.
24. Gröschel-Stewart, U., Rudiger, H. & Turba, F. (1964) *Biochem. Z.* **339**, 539–542.
25. Yamashita, R., Soma, Y., Kobayashi, S., Sekine, R., Titani, K. & Narita, K. (1964) *J. Biochem. (Tokyo)* **55**, 576–577.
26. Yamashita, T., Soma, Y., Kobayashi, S. & Sekine, R. (1965) *J. Biochem. (Tokyo)* **57**, 460–461.
27. Kimura, M. & Kielley, W. W. (1966) *Biochem. Z.* **345**, 188–201.
28. Yamashita, T., Soma, Y., Kobayashi, S. & Sekine, T. (1974) *J. Biochem. (Tokyo)* **75**, 447–453.
29. Mannherz, H. G. & Goody, R. S. (1976) *Annu. Rev. Biochem.* **45**, 427–465.
30. Riordan, J. F., McElvany, K. D. & Borders, C. L., Jr. (1977) *Science* **195**, 884–885.
31. Takamori, K., Kato, K. A. & Sekine, T. (1976) *J. Biochem. (Tokyo)* **30**, 101–110.
32. Schliselfeld, L. H. (1977) *Biophys. J.* **17**, 37a.
33. Reisler, E., Burke, M., Himmelfarb, S. & Harrington, W. F. (1974) *Biochemistry* **13**, 3837–3840.
34. Reisler, E., Burke, M. & Harrington, W. F. (1974) *Biochemistry* **13**, 2014–2022.
35. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 222–245.
36. Weeds, A. G. & Hartley, B. S. (1968) *Biochem. J.* **107**, 531–547.
37. Weeds, A. G. (1967) *Biochem. J.* **104**, 44.
38. Starr, R. & Offer, G. (1973) *J. Mol. Biol.* **81**, 17–31.
39. Hardy, M. F. & Perry, S. V. (1969) *Nature* **223**, 300–302.
40. Huszar, G. (1975) *J. Mol. Biol.* **94**, 311–326.
41. Kubo, S., Tokuyama, H. & Tonomura, Y. (1965) *Biochim. Biophys. Acta* **100**, 459–470.
42. Barany, G. & Merrifield, B. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 121–125.