Determination of the Amino Acid Sequence of Troponin C from Rabbit Skeletal Muscle*

(Received for publication, May 10, 1976, and in revised form, April 12, 1977)

JOHN H. COLLINS,†, ‡ MARION L. GREASER,§ JAMES D. POTTER,‡, § AND MARCUS J. HORN

From the †Department of Muscle Research, Boston Biomedical Research Institute, Boston, Massachusetts 02114, the Departments of Neurology and Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02114, and the §Muscle Biology Laboratory, University of Wisconsin, Madison, Wisconsin 53706

The complete amino acid sequence of rabbit skeletal muscle troponin C (TnC) was determined from studies on overlapping peptides isolated from cyanogen bromide and tryptic digests. TnC was found to be a single polypeptide chain of 159 amino acid residues, including 2 residues of tyrosine and 1 each of cysteine, histidine, and proline. The amino end is acetylated, the calculated net charge at pH 7.0 is -29, and the calculated molecular weight is 17,965. There was no evidence for heterogeneity in the sequence. The previously proposed four apparent Ca$^{2+}$ binding sites are located at residues 27 to 38, 63 to 74, 103 to 114, and 139 to 150. TnC is the Ca$^{2+}$-binding subunit of troponin. Troponin, a complex containing two additional subunits, is bound to tropomyosin in the thin filaments of striated muscle. The troponin-tropomyosin system confers Ca$^{2+}$ sensitivity to the interaction of actin and myosin that occurs during muscle contraction (for reviews, see Refs. 1 and 2). TnC of rabbit skeletal muscle possesses two Ca$^{2+}$-Mg$^{2+}$ sites and two Ca$^{2+}$-specific sites (3,4). The Ca$^{2+}$-Mg$^{2+}$ sites are always occupied by Ca$^{2+}$ or Mg$^{2+}$ under physiological conditions and the regulation of actin-myosin interaction possibly occurs by reversible binding to the Ca$^{2+}$-specific sites (4,5). This report is a brief description of the determination of the complete amino acid sequence of TnC from rabbit skeletal muscle. This information provides a basis for carrying out structure-function studies and for the eventual interpretation of the three-dimensional structure of TnC, which recently has been crystallized (6).

RESULTS AND DISCUSSION

The amino acid sequence of TnC was reconstructed from studies on peptides produced from cleavage with CNBr and trypsin. First, the single cysteine of TnC was S-aminoethylated and the protein cleaved at its 10 methionines with CNBr. Nine peptides (designated CB1 to CB9), ranging in size from 2 to 52 amino acid residues, were isolated. These peptides, plus 2 mol/mol of free homoserine, accounted for all 159 residues of the polypeptide chain (see Fig. 1). The partial sequence of CB8 and the complete sequences of the other CNBr peptides were determined (Fig. 2). Next, 11 peptides (designated R1 to R8 in Fig. 1), which also accounted for the entire protein, were isolated from a tryptic digest of performic acid-oxidized and cyanogenylated TnC. Partial sequence determination of "R3" (actually a mixture of R3B and R3A-R3B) completed the sequence of CB8 and established the alignment CB1-CB8-Met-CB3. The remaining alignments of the CNBr fragments were deduced from the amino acid compositions of R2A, R2B, R2C, R4, R7, and R8. Confirming evidence for the sequence of TnC was provided by isolating and analyzing 18 peptides (designated T1 to T18B in Fig. 1) from a tryptic digest of aminoethylated TnC. These peptides accounted for all residues except 79 to 84 of TnC.

Each of the three digests of TnC was separated into acid-soluble and acid-insoluble fractions prior to chromatographic separation of the peptides on columns of Sephadex G-50, Bio-Gels P-2, P-4, and P-6, sulfopropyl (SP)-Sephadex, quaternary aminomethyl (QAE)-Sephadex, and AG 50W-X8. During the course of the sequencing studies, several of the peptides from TnC digests were further cleaved into subpeptides (see Fig. 2) with trypsin, chymotrypsin, pepsin, thermolysin, and pronase. The subpeptides were isolated by the mentioned chro-

* This work was supported by Grants AM-14728, AM-14776, HL-05811, HL-05949, and HL-17464 from the National Institutes of Health, Grants GM-15094 and CH-35070 from the National Science Foundation, and Grants 1133, 1181, and 1296 from The American Heart Association, and grants from The American Heart Association, Massachusetts Affiliate, Inc., the Department of Health, Education, and Welfare, the National Institutes of Health, and the Muscular Dystrophy Association of America, Inc., an Established Investigatorship (to J. D. P.) of the Muscular Dystrophy Association of America, Inc., an Established Investigatorship (to J. D. P.) of the American Heart Association, and the College of Physicians and Surgeons, Harvard Medical School, Boston, Massachusetts 02114. Requests for reprints should be sent to Dr. J. D. Potter.

† Present address, Department of Pharmacology and Cell Biophysics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267.

‡ Present address, Department of Neurology and Biological Chemistry, Harvard Medical School, Boston, Massachusetts 02114.

§ Present address, Department of Neurology, Harvard Medical School, Boston, Massachusetts 02114.

¶ Present address, Department of Pharmacology, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267.

# Present address, Department of Neurology, Harvard Medical School, Boston, Massachusetts 02114.

* The abbreviations used are: TnC, troponin C; CNBr, cyanogen bromide; Ac, acetyl group; CB, cyanogen bromide peptide; R, peptide from tryptic digest of cyanogenylated TnC; T, other tryptic peptide; C, chymotryptic peptide; P, pepsin peptide; M, thermolysin peptide; S, Pronase peptide.
peptides were isolated from the acid-soluble fraction which and reprecipitating by acidification. All of the other CNBr peptides and the sequence of CB4 was reconstructed from the homogenous state by repeatedly dissolving in dilute pyridine protein with CNBr (CB) and trypsin (R,T). Residues predicted to be cle TnC and placement of peptides obtained by digestion of the intact
tion, by manual subtractive Edman degradation, and by hy-
drolysis with aminopeptidase M and carboxypeptidase A. Amino acid sequences of peptides aspartic acid precedes threonine at the NH₂ terminus of T1. bonds of T1, except the Asx-Thr bond, were hydrolyzed by played endopeptidase activity. It appeared that all the peptide sequences of T1 is AC-(Asp, Thr, Gln, Gin)-Ala-Glu-Ala-Arg. A carboxypeptidase A digest of P1 provided further evidence that aspartic acid precedes threonine at the NH₂ terminus of T1.

CNBr Digest – The acid-insoluble fraction of this digest consisted mainly of CB8, which could be brought to a highly homogenous state by repeatedly dissolving in dilute pyridine and reprecipitating by acidification. All of the other CNBr peptides were isolated from the acid-soluble fraction which also contained 1.7 mol/mol of free homoserine. The sequences of CB1 to CB3 were determined from studies on the intact peptides and the sequence of CB4 was reconstructed from the sequences of its two tryptic subpeptides. CB2 was presumed to be the COOH-terminal peptide of TnC since it was the only CNBr peptide not containing homoserine.

The sequence of CB5 was also reconstructed from its two tryptic subpeptides (T1 and T2), but because T1 was acetylated at its α-NH₂ group, its sequence could not be determined by Edman degradation. From an aminopeptidase M digestion of T1 we obtained evidence that T1 contains 2 residues of glutamate and that its NH₂-terminal sequence is Ac-Asx-Thr-. (The carboxypeptidase A digest of Pl provided further evidence that its NH₂-terminal sequence is Ac-Asp-Thr-Gln-Gln-Ala-Glu-Ala-Arg-Ser-Tyr-Leu-Ser-Glu-Glu-Met-Leu-Ala-Glu-Phe-Lys.)

We concluded that the NH₂-terminal sequence of T1 (and also of TnC) is Asp-Thr-Gln-Gln-Ala-Glu-Ala-Arg.

CB6 was cleaved into the subpeptides T1 and T2 with trypsin and T1 was further cleaved into P1 and P2 with pepsin. The sequence of CB6 was then reconstructed from the NH₂-terminal sequence of T1 and the complete sequences of P2 and T2. Tryptic digestion of CB7 yielded three subpeptides (T1 to T3) and 1 mol/mol of free homoserine. The sequences of T2 and T3 were determined and a solid phase Edman degradation of intact CB7 established the sequence of T1 and the order of the three tryptic subpeptides within CB7.

The NH₂-terminal 16 residues of CB8 were identified by solid phase Edman degradation. Digestion of CB8 with carboxypeptidase A at pH 8 released the presumed COOH-terminal homoserine plus 1 mol/mol each of phenylalanine, leucine, and valine. When the carboxypeptidase A digestion was repeated at pH 5, these same 4 residues, plus the second phenyl-
alaine, the only remaining isoleucine, a glutamic acid, and two aspartic acids, were released in yields of 0.74 to 1.02 mol/mol. The carboxypeptidase A digestion at pH 5 also released 0.40 threonine, 0.20 serine, and 0.20 glycine. The partial sequence of CB8 shown in Fig. 2 is based on the results of amino acid analysis, solid phase Edman degradation, and a tentative interpretation of the carboxypeptidase A digestions.

Two tryptic digestions of CB9 were carried out. In one of these, cleavage was restricted to the 3 arginines by prior blocking of amino groups with citraconic anhydride. The subpeptide T2 to T4 were isolated from both tryptic digests, while T1 was obtained only from citraconylated CB9 and T1A to T1D were obtained only from CB9 that was not citraconylated. Partial sequence data on T1 and on T1A to T1D were used to reconstruct the complete sequence of T1. The sequences of T2 and T4 were determined from studies on the intact subpeptides. T3 was shown to have NH2-terminal Asx, and then was further digested with thermolysin to yield the subpeptides M1 to M4. The sequences of M1, M3, and M4 were determined and the partial sequence of T3 was then deduced to be (Asn-Ala-Asp-Gly-Tyr/Ile, Asx, Ala, Glx, Glx/Leu-Ala-Glu)-Ile-Phe-Arg. CB9 was then digested with chymotrypsin and the subpeptides C1 to C4 were isolated. The sequence of C2 was determined and from their amino acid compositions C1 was assumed to be T1 less the COOH-terminal arginine and C4 was assumed to be T4 plus an additional arginine at the NH2 terminus. (Homoserine was presumed to be the COOH-terminal residue of CB9.) C3 (which obviously overlaps T3, since both subpeptides contain the only tyrosine of CB9) was digested with pepsin and the subpeptides P1 to P6 were isolated. The NH2-terminal sequence of P1 and the complete sequences of P3, P4, and P6 were determined and from this information the complete sequences of both T3 and C3 were reconstructed. The sequence of C3 established the order of T2 and T3 within CB9. It then followed that the NH2-terminal arginine of C2 could only come from T1, thus establishing the order of T1 and T2 within CB9. Since T4 must be at the COOH terminus of CB9, the order T1-T2-T3-T4 was established, thus completing the sequence of CB9.

Tryptic Digest of Citraconylated TnC—Since TnC contains 7 arginines, we expected to find eight peptides (R1 to R8) in the tryptic digest of citraconylated TnC. Peptides R1 and R4 to R8 were isolated from the acid-soluble fraction of this digest, while R2 (residues 9 to 44) and R3 (residues 45 to 81) appeared to be in the acid-insoluble fraction. The amino acid compositions of R4, R7, and R8 established the alignment CB3-CB8-CB7-Met-CB2 at the COOH terminus of TnC. The acid-insoluble fraction (from which the citraconyl groups had been removed by acidification) was solubilized by suspending in H2O and adjusting to pH 8.4 with KOH. This treatment appeared to reactivated the trypsin that was present, causing complete peptide bond cleavage at the 2 lysines of R2 and partial cleavage at the single lysine of R3. As a result, R2A to R2C (but not R2), R3A, and a mixture of R3B and uncleaved R3A-R3B were the only peptides isolated from this fraction. The amino acid compositions of R2A to R2C (all of which were acid-soluble) established the alignment CB5-CB4-CB6-CB1 at the NH2 terminus of TnC. R3A was also acid-soluble while R3B and R3A-R3B remained soluble only at pH above 4.

The mixture of R3B and R3A-R3B (listed as "R3" in Fig. 2) was digested with chymotrypsin and the subpeptides C1 (58% yield) and C2 (78% yield) were isolated and their sequence determined. Nine subpeptides (designated M1 to M6 in Fig. 2) were isolated from thermolysin digests of "R3." Studies on the subpeptides of "R3" were used to reconstruct the sequence of the COOH-terminal 17 residues of CB8 and also to establish the alignment CB1-CB8-Met-CB3 within the sequence of TnC. From its amino acid composition, it was obvious that M1 was equivalent to the NH2-terminal 9 residues of CB8 plus a methionine which can only be at the NH2 terminus of M1. M1A, M2, and M3A were presumed to represent residues 8 to 16 of CB8. The sequence of M3A was determined and found to be identical with that of residues 13 to 16 of CB8. Since arginine must be the COOH-terminal residue of "R3," a comparison of the amino acid compositions of M5 and M6 with the sequences of C1 and C2 established the COOH-terminal sequence of "R3" to be -Leu-Val-Met-Met-Arg. This established the sequence of the COOH-terminal 3 residues of CB8 and the alignment CB8-Met-CB3 within the sequence of TnC. The alignment CB1-CB8 could then be assumed by difference, since all the other CNBr fragments had been aligned. This is confirmed by the (presumed) NH2-terminal methionine of M1 which is very likely to be derived from cleavage at the arginine contained in CB1. The complete sequences of M3B and M4 were determined and these subpeptides were identified from comparison with the partial sequence of CB8 as equivalent to residues 17 to 30 of CB8.

It should be mentioned that the assignment of threonine as the COOH-terminal residue of M3B is based solely on the result that carboxypeptidase A digestion of this subpeptide released a rather low amount of threonine (0.09 mol/mol) and no other amino acid. The NH2-terminal sequence of M3B was established as Val-Asx-Glx-Asx-Gly- by subtractive Edman degradation. It was concluded that the COOH-terminal sequence of M3B is either -Gly-Ser-Thr or -Ser-Gly-Thr. Aminopeptidase M digestion of M3B (which showed that M3B contains a glutamic acid and two aspartic acids) released (in addition to other residues) 0.50 mol/mol of serine, but only 0.12 mol/mol of threonine. This showed that the COOH-terminal sequence of M3B cannot be -Gly-Ser-Thr and we, therefore, concluded that it must be -Ser-Gly-Thr. Further support for the sequence of M3B may be derived from comparison of the sequence of rabbit skeletal muscle TnC with the recently reported sequences of bovine cardiac muscle TnC (7) and chicken skeletal muscle TnC (8). All three TnCs have identical sequences in the region represented by peptide M3B (residues 62 to 69 in Fig. 1).

Tryptic Digest of Aminoethylated TnC—This digest would be expected to yield 18 peptides if complete and specific tryptic cleavage occurred at the 9 lysines, 7 arginines, and 1 cysteine of S-aminoethylated TnC. From the peptides isolated (designated T1 to T18B in Fig. 1), all of which could be placed in the sequence of TnC, it was apparent that complete and specific tryptic cleavage was indeed achieved with two exceptions. The Met-Met bond of residues 78 to 79 was completely cleaved and the Lys-Glu bond of residues 52 to 53 was only partially cleaved. As a result, the acid-insoluble fraction of this digest was a mixture of T6A and uncleaved T5-T6A. The peptides T1 to T9 and T8 to T17 were isolated from the acid-soluble fraction. Peptide T6B (Met-Val-Arg, residues 79 to 81) was not isolated, but did appear (from amino acid analysis) to be present in a fraction which yielded T8, T9, T15, and T16. Another fraction which appeared to be a mixture of T7 (Gln-Met-Lys, residues 82 to 84) and T13 (Met-Met-Glu-Gly-Val-Gln, residues 154 to 159) did not yield pure peptides. The peptides T18A and T18B were obtained from a thermolysin digest of this mixture. Comparison of the amino acid compositions of T18A and T18B with the sequences of the CNBr...
Amino Acid Sequence of Rabbit Skeletal Muscle Troponin C peptides allowed us to unequivocally place T18A and T18B at the COOH terminus of TnC.

Evolution and Predicted Three-dimensional Structure—In our preliminary reports (9, 10) we presented the nearly complete sequence of rabbit skeletal muscle TnC and confirmed its predicted (11) homology with Ca\(^{2+}\)-binding parvalbumin. By analogy with the known three-dimensional structure of a parvalbumin from carp muscle (12), we proposed that the three-dimensional structure of TnC consists of four similar domains, each of which contains a 12-residue Ca\(^{2+}\)-binding segment flanked on either side by a pair of α helices. It has subsequently been shown that TnC is also homologous to both the alkali and sulfhydryl-reactive light chains of rabbit skeletal muscle myosin (13, 14) and to a cyclic nucleotide phosphodiesterase activator protein isolated from bovine brain (15, 16). All of these proteins, which have different functional properties, appear to have evolved from a common ancestor which was produced by reduplication of a gene for a small (approximately 35 amino acid residues) precursor Ca\(^{2+}\) binding protein. Numerous detailed analyses (17-23) of the structure and evolution of TnC and related proteins have been presented elsewhere (see Ref. 23 for a review). Fig. 1 shows the residues predicted to form four Ca\(^{2+}\)-binding sites and eight α helices in TnC.

Acknowledgments—The major part of this work was greatly facilitated by the excellent technical assistance of Nora Jackman and Gedelita Wilshire. The very capable assistance of Erlinda Capuno, Ellen Kraft, and Christine Martin is also acknowledged. The authors are grateful to Drs. Marshall Elzinga and John Gergely for their support.

REFERENCES

Amino Acid Sequence of Rabbit Skeletal Muscle Troponin C

The complete amino acid sequence of rabbit skeletal muscle troponin C (TnC) has been determined. This protein is composed of 351 amino acids and has a molecular weight of approximately 39,500. The sequence was obtained by automated Edman degradation and amino acid analysis. The sequence shows significant homology to other troponin C sequences from different species, suggesting a conserved role in muscle function. Further studies on the function of TnC are ongoing.
Amino Acid Sequence of Rabbit Skeletal Muscle Troponin C
Determination of the amino acid sequence of troponon C from rabbit skeletal muscle.
J H Collins, M L Greaser, J D Potter and M J Horn


Access the most updated version of this article at http://www.jbc.org/content/252/18/6356

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/252/18/6356.full.html#ref-list-1