Correction. In the article "An assay of ribonuclease H, endoribonucleases, and phosphatases" by Jannis G. Stavrianopoulos and Erwin Chargaff, which appeared in the May 1976 issue of the Proc. Natl. Acad. Sci. USA 73, 1556-1558, the authors have requested the following change. In Table 1, line 5 of footnote* should read "centrifuged in a horizontal rotor (2 min, 12,000 × g)."

Correction. In the article "Amino-acid sequence of tropomyosin-binding component of rabbit skeletal muscle troponin" by J. R. Pearlstone, M. R. Carpenter, P. Johnson, and L. B. Smillie, which appeared in the June 1976 issue of the Proc. Natl. Acad. Sci. USA 73, 1902-1906, the authors have requested the following change. On page 1904, in Fig. 2, the second amino-acid residue should be Asp and not Asn.
Amino-acid sequence of tropomyosin-binding component of rabbit skeletal muscle troponin

ABSTRACT The complete amino-acid sequence of rabbit skeletal troponin-T is reported. The protein consists of a single polypeptide chain of 259 amino acids; it has an acetylated amino terminus and a molecular weight of 30,503. The sequence was determined by manual and/or automated Edman degradation techniques on the six fragments obtained after cleavage with cyanogen bromide. The larger fragments were further digested with trypsin, chymotrypsin, a-lytic protease, thermolysin, or pepsin to obtain smaller fragments suitable for manual sequencing. About 50% of the residues are charged at neutral pH with highly acidic amino-terminal (residues 1-39) and highly basic carboxyl-terminal regions (residues 221-259). Predictions of secondary structure indicate 37% helical content with two long sections (residues 90-105 and 122-148) in that portion of the molecule implicated in binding to tropomyosin. Two of the three phosphorylated sites in the molecule are located at serine-149 or -150. The sequence about the latter site resembles somewhat the phosphorylase kinase phosphorylation sites in phosphorylase a and troponin-I.

The troponin complex of skeletal muscle tissue in conjunction with tropomyosin is known to be intimately involved in the regulation of contraction by calcium ions (see review by Weber and Murray in ref. 1). Rabbit troponin is a complex of three different proteins: troponin-C (Tn-C, 17,900 in molecular weight), which binds calcium, troponin-I (Tn-I, 20,900), which inhibits actin-myosin interaction in the presence and absence of calcium, and troponin-T (Tn-T, 30,500), which binds to tropomyosin. All three components, present in a 1:1:1 molar ratio, are required for full calcium sensitization of the actomyosin ATPase (2-4). A full understanding at the molecular level of the ways in which these proteins interact with each other and with the actomyosin complex during contraction and relaxation will require a detailed knowledge of their three-dimensional structures. As a step towards this end it is necessary to know their amino-acid sequences. This information is now available for rabbit skeletal a-tropomyosin (5), Tn-C (6), and Tn-I (7) and for bovine cardiac Tn-C (8). We now report the complete primary structure of rabbit skeletal muscle Tn-T. While this work was in progress, the isolation and characterization of the cyanogen bromide fragments of Tn-T were reported by Collins (9) and by Jackson et al. (10). The latter workers also ordered the fragments by the isolation of methionine overlap peptides. Their results are in essential agreement with those reported herein.

MATERIALS AND METHODS

Troponin-T was isolated from the troponin complex by the method of Greaser and Gregely (11). The protein was cleaved with cyanogen bromide, and the resulting mixture of fragments (CB, cyanogen bromide fragment) separated by gel filtration on Sephadex G-75 equilibrated with 10% formic acid (Fig. 1). Mixtures of CB2 + CB3 and CB5 + CB6 were further purified on DEAE-cellulose and phosphocellulose columns, respectively. The purity and approximate size of fragments CB1, CB2, CB3, and CB4 were determined by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate.

The sequences of the cyanogen bromide fragments were determined by a combination of automatic and manual methods. The former were performed on a Beckman Sequencer (model 890B) with identification of the residues as their phenylthiohydantoins by gas-liquid chromatography (12) and/or as the free amino acids after hydrolysis of the thiazoliones (13). The fragments were also digested with trypsin, chymotrypsin, thermolysin, a-lytic protease, or pepsin and the resulting peptides separated by high-voltage electrophoresis on paper at pH 6.5 and 1.8. The digestion conditions for a-lytic protease, prepared in the laboratory of Dr. D. R. Whitaker, University of Ottawa, have already been described (14). Manual sequencing of these peptides was by the dansyl-Edman procedure (15, 16). Amides were assigned from the mobilities of peptides relative to aspartic acid at pH 6.5 (17). Penicillocarboxypeptidase-S (gift of Dr. T. Hofmann, University of Toronto) was used to confirm the sequence of the COOH-terminus (18) of some peptides. The NH2-terminal acetyl group was identified by acid hydrolysis of CB3, followed by gas-liquid chromatography using a Hewlett-Packard (model 5700A) instrument fitted with a Supelco SP1200 column (19). Amino-acid analyses were performed on either a Durrum (model D500) or Beckman (model 120C) instrument.

The order of the cyanogen bromide fragments was determined by the isolation of methionine peptides from a peptic digest of the whole protein. These were purified by ion-exchange chromatography and high-voltage paper electrophoresis by methods previously described (20, 21). Methionine-containing peptides were located, and their purity was established by amino-acid analyses. Sequences were determined manually or automatically.

RESULTS AND DISCUSSION

Troponin-T is a single polypeptide of 259 amino acid residues with a molecular weight of 30,503, substantially less than the value of 37,000 previously evaluated by sodium dodecyl sulfate gel electrophoresis. Like the other myofibrillar proteins of rabbit skeletal muscle, its NH2-terminus is acetylated. Since Tn-T has five methionines, the protein was initially cleaved by cyanogen bromide and the resultant lyophilized digest applied to a Sephadex G-75 column equilibrated with 10% formic acid (Fig. 1). Fragments CB1 and CB4 were sufficiently pure for direct sequence analysis. The mixture of CB2 and CB3 was pooled and further fractionated on DEAE-cellulose. The mixture of CB5
and CB6 was separated on a phosphocellulose column. CB7 was further purified by high-voltage paper electrophoresis. The last peak off the column contained an insignificant amount of peptide material, as indicated by amino-acid analysis.

The amino-acid compositions of the seven cyanogen bromide fragments and their NH$_2$-terminal residues are shown in Table 1. These analyses are in good agreement with those published recently by others (9, 10). In each case, the composition corresponds to that derived from the sequence (given in parentheses). The sum of CB2-CB7 inclusive is 259 residues, with a molecular weight of 30,560. This result agrees closely with the composition of whole Tn-T (last column). A comparison of the first two columns indicates strongly that CB1 represents the same amino-acid composition of CB2 and CB3 and has resulted from incomplete cleavage of the methionyl peptide bond between these two fragments. An analysis of the NH$_2$-terminus of each fragment by the dansyl method showed that CB1 and CB3 were blocked. When automatic sequence analysis was applied to CB1, the first cycle was blank, but steps 2-30 gave the same NH$_2$-terminal sequence as CB2. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of the products after one cycle through the sequencer showed two additional bands with mobilities identical to CB2 and CB3. From these results it is evident that the cyanogen bromide cleavage at methionine-70 (see Fig. 2) was incomplete but that this bond in CB1 was partially cleaved under the conditions of the first Edman cycle. Since CB3 was blocked, the subsequent cycles involved determination of the sequence of CB2. Further evidence that CB1 arose from incomplete cleavage at methionine-70 was provided by the isolation of a picptic fragment from CB1 which, when its sequence was determined, was found to correspond to residues 58-72, thus providing an overlap for the COOH-terminus of CB3 and the NH$_2$-terminus of CB2.

The isolation of methionine peptides from a picptic digest of the whole protein provided overlaps for each of the cyanogen bromide fragments and permitted their ordering as CB3-CB2-CB5-CB4-CB7-CB6. This alignment is in essential agreement with that recently published by Jackson et al. (10). However, we find that residue 68 is aspartic acid rather than glutamic acid, as reported in the methionine overlap sequence for CB3 and CB2 by these workers. A detailed report of the sequence determination will be published elsewhere.

The most remarkable feature of the sequence of this protein (Fig. 2) is its high content of charged residues. Out of a total of 259 residues, 61 are aspartates and glutamates, 64 are arginines and lysines, and 6 are histidines. Thus, about 50% of the amino acid residues of Tn-T bear a positive or negative charge at neutral pH. Inspection shows that the distribution of these residues is clearly not random. The NH$_2$-terminal region (residues 1-39) is highly acidic (18 aspartates plus glutamates) and devoid of basic residues except for four histidines. The COOH-terminal sequence (residues 221-259), however, is

Table 1. Amino-acid compositions, yields, and molecular weights of the cyanogen bromide fragments from troponin-T

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>CB1</th>
<th>CB2</th>
<th>CB3</th>
<th>CB4</th>
<th>CB5</th>
<th>CB6</th>
<th>CB7</th>
<th>CB2-CB7</th>
<th>Troponin-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>12.0</td>
<td>10.7(10)</td>
<td>4.5(4)</td>
<td>6.2(6)</td>
<td>7.2(7)</td>
<td>2.0(2)</td>
<td>1.0(1)</td>
<td>20.9(20)</td>
<td>20.7</td>
</tr>
<tr>
<td>Thr</td>
<td>1.8</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
<td>2.1(2)</td>
<td>1.1(1)</td>
<td>1.9(2)</td>
<td>6.1(6)</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>4.2</td>
<td>4.5(5)</td>
<td>2.8(3)</td>
<td>1.7(2)</td>
<td>1.1(1)</td>
<td>1.8(2)</td>
<td>0.9(1)</td>
<td>8.3(9)</td>
<td>8.7</td>
</tr>
<tr>
<td>Glu</td>
<td>38.4</td>
<td>43.9(45)</td>
<td>24.4(24)</td>
<td>19.5(21)</td>
<td>8.7(8)</td>
<td>3.0(3)</td>
<td>1.0(1)</td>
<td>56.5(57)</td>
<td>55.7</td>
</tr>
<tr>
<td>Pro</td>
<td>8.3</td>
<td>7.1(8)</td>
<td>7.1(8)</td>
<td>1.0(1)</td>
<td>8.1(9)</td>
<td>9.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>2.0</td>
<td>1.2(1)</td>
<td>1.2(1)</td>
<td>2.3(2)</td>
<td>3.9(4)</td>
<td>8.6(8)</td>
<td>9.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>17.5</td>
<td>16.9(16)</td>
<td>11.7(11)</td>
<td>5.2(5)</td>
<td>2.4(2)</td>
<td>4.1(4)</td>
<td>3.1(3)</td>
<td>1.0(1)</td>
<td>27.4(26)</td>
</tr>
<tr>
<td>Val</td>
<td>7.5</td>
<td>7.3(8)</td>
<td>0.9(1)</td>
<td>6.4(7)</td>
<td>7.5(8)</td>
<td>8.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>4.9</td>
<td>4.6(5)</td>
<td>2.8(3)</td>
<td>1.8(2)</td>
<td>2.9(3)</td>
<td>7.5(8)</td>
<td>8.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>10.2</td>
<td>9.3(9)</td>
<td>7.2(7)</td>
<td>2.1(2)</td>
<td>7.7(8)</td>
<td>0.9(1)</td>
<td>1.0(1)</td>
<td>18.9(19)</td>
<td>11.7</td>
</tr>
<tr>
<td>Tyr</td>
<td>1.9(2)</td>
<td>1.8(2)</td>
<td>1.8(2)</td>
<td>1.0(1)</td>
<td>0.9(1)</td>
<td>1.0(1)</td>
<td>18.9(19)</td>
<td>11.7</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>1.8</td>
<td>2.0(2)</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
<td>2.0(2)</td>
<td>1.0(1)</td>
<td>5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>4.9</td>
<td>4.3(5)</td>
<td>0.9(1)</td>
<td>3.4(4)</td>
<td>1.0(1)</td>
<td>5.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>19.7</td>
<td>17.9(18)</td>
<td>10.8(11)</td>
<td>7.1(7)</td>
<td>10.0(11)</td>
<td>4.1(4)</td>
<td>5.7(6)</td>
<td>37.7(39)</td>
<td>36.5</td>
</tr>
<tr>
<td>Trp</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
<td>+1(1)†</td>
</tr>
<tr>
<td>Arg</td>
<td>17.2</td>
<td>16.3(16)</td>
<td>14.2(14)</td>
<td>2.1(2)</td>
<td>3.9(4)</td>
<td>2.0(2)</td>
<td>1.0(1)</td>
<td>1.9(2)</td>
<td>25.1(25)</td>
</tr>
<tr>
<td>Hse</td>
<td>1.5</td>
<td>2.0(2)</td>
<td>1.0(1)</td>
<td>1.0(1)</td>
<td>0.9(1)</td>
<td>0.8(1)</td>
<td>1.0(1)</td>
<td>4.7(5)</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>0.4</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>152.3</td>
<td>(151)</td>
<td>(81)</td>
<td>(70)</td>
<td>(55)</td>
<td>(24)</td>
<td>(21)</td>
<td>(8)</td>
<td>(259)</td>
</tr>
<tr>
<td>Yield (%)</td>
<td>40.0</td>
<td>39</td>
<td>57</td>
<td>57</td>
<td>48</td>
<td>48</td>
<td>34</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>$M$,§</td>
<td>17,899</td>
<td>17,893</td>
<td>9,852</td>
<td>8,059</td>
<td>6,789</td>
<td>2,702</td>
<td>2,217</td>
<td>974</td>
<td>30,503</td>
</tr>
<tr>
<td>NH$_2$-terminus</td>
<td>Glx</td>
<td>Lys</td>
<td>Gly</td>
<td>Leu</td>
<td>Axx</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
and CB3 positively charged, though there are two stretches of 147-160.

FIG. 2. Amino-acid sequence of troponin-T. The cyanogen bromide fragments (CB2 to CB7) are indicated. Fragment CB1 is the sum of CB2 and CB3 and arises from incomplete cleavage at methionine-70.

highly positively charged, with a total of 13 basic residues and only two acidic residues. The acidic and basic residues in the remainder of the structure are more evenly distributed, although there is a distinct tendency for the charged amino acids to be grouped in clusters up to seven residues in length. A further notable feature is the almost total absence of extensive stretches of nonpolar residues in the sequence. The only such region longer than five amino acid residues in length is from residues 147-160. These observations would indicate that there is very little in the way of a nonpolar core to this molecule and suggest that its structure is an open one permitting ready interactions with solvent, ions, or other protein molecules. Such an open structure is consistent with the susceptibility of Tn-T to proteolytic degradation during its isolation, an observation made in a number of laboratories.

An analysis of the sequence by the method of Chou and Fasman for predicting secondary structure (23) has indicated $\alpha$-helical segments between residues 1-8, 26-31, 68-76, 80-102, 122-146, 170-184, and 216-225, totalling 37% $\alpha$-helix. Four sections of $\beta$-sheet structure were predicted at residues 107-112, 204-211, 229-236, and 238-242, totalling 10% of the protein. In addition, seven $\beta$-turns were located in the areas 16-19, 49-52, 64-67, 118-121, 152-155, 155-158, and 165-168. Of the nine prolines found in Tn-T, eight are located between residues 16 and 51, predicted as having little helical content. These secondary structural features of Tn-T are shown schematically in Fig. 3.

Jackson et al. (10) have examined the binding properties of the cyanogen bromide fragments of Tn-T with tropomyosin. Fragment CB1 (residues 1-151), when added to tropomyosin, showed a sharp increase in viscosity similar to that observed with whole Tn-T. Fragment CB3 (residues 1-70) gave no such increase, while the viscosity change with CB2 (residues 71-151) was intermediate. These authors have concluded that the tropomyosin binding site on Tn-T is located between residues 71 and 151, although one cannot exclude the possibility of a role for the COOH-terminal region of CB3 in this binding as well as other parts of the structure.

In considering the nature of the structural features of CB2 that may be responsible for tropomyosin binding, it is probably significant that this region of the molecule is predicted to have
**Biochemistry:** Pearlstone et al. 

**Phosphorylase a** Ser-Asp-Gln-Glu-Lys-Arg-Glu-Gln-Ile-Ser(F)-Val-Arg-Gly-Leu

**Troponin-I** Asp-Glu-Glu-Lys-Arg-Arg-Arg-Ala-Ile-Thr(F)-Ala-Arg-Arg-Gln

**Troponin-T** Glu-Asp-Leu-Lys-Lys-Lys-Leu-Leu-Ser(F)-Ser-Met-Cly-Ala

**FIG. 4.** Comparison of the amino-acid sequences about the sites phosphorylated by phosphorylase kinase in phosphorylase a (33), troponin-I (32), and troponin-T (ref. 31 and present work).

by far the largest percentage of α-helical content. In particular, the two long segments (residues 80–102 and 125–146) found in this part of the molecule are highly charged, and it is tempting to speculate that it is one or both of these segments which interact over a limited surface region of the two-stranded coiled-coil structure of tropomyosin. It would be premature to speculate on the exact nature of this interaction without a more precise knowledge of the binding site of Tn-T on the tropomyosin molecule. However, such interaction must involve a considerable number of salt linkages.

The interaction of Tn-T and Tn-C has been well documented in a number of laboratories (3, 24–28), and there seems little question that these two proteins are intimately associated in the troponin complex. Although the evidence for Tn-T to Tn-I interaction is much less extensive, the recent studies of Hitchcock (29), using crosslinking with dimethylimido esters, have shown that these two components must be within 0.6 nm or less of each other. Because of the highly acidic character of Tn-C (6) and the basic properties of Tn-I (7), it is tempting to speculate that the highly basic COOH-terminal region of Tn-T (residues 221–259) may serve as a binding site for Tn-C and the highly acidic NH2-terminal residue (residues 1–59) as an interaction site for Tn-I. It should be possible to test these predictions by measuring the binding properties of suitably prepared fragments of Tn-T with Tn-C and Tn-I.

The phosphorylation of myofibrillar proteins has been studied in several laboratories in recent years. In particular, three phosphorylation sites on Tn-T have been reported (30). Amino-acid compositions or sequences about two of these sites have been described (31), and these may now be placed in the completed primary structure of the protein. One of these sites is the NH2-terminal serine residue and is found in the phosphorylated state in the protein isolated by the usual preparative procedures. It is not readily dephosphorylated by phosphatases present in the muscle extracts, and the enzyme responsible for the phosphorylation is not known. The second known site is at either serine-149 or -150. The phosphorylation of this site is catalyzed by phosphorylase kinase. A similar site (threonine-11) in Tn-I is also known to be phosphorylated by this enzyme (32), and with the complete sequence of Tn-T now known, it is possible to compare the sequences about the sites of phosphorylation by this enzyme in phosphorylase a, Tn-I, and Tn-T. This comparison (Fig. 4) shows some degree of isology and indicates the importance of the hydrophobic amino acid on the immediate NH2-terminal side of the phosphorylated site and of a cluster of basic residues three to five amino acids removed from the serine or threonine which is phosphorylated. The significance of these phosphorylation sites in the functional role of Tn-T in the calcium regulatory system is presently unknown. However, as Perry et al. (10, 31) have pointed out, it is probably significant that the serine-149 site is close to a region of the polypeptide chain implicated in the binding of this protein to tropomyosin. Similarly the serine-1 site of phosphorylation is in an already highly acidic region and may be involved in a site of interaction with the basic protein Tn-I. It will be of considerable interest to elucidate the third phosphorylation site in this protein.

A preliminary examination of the sequence of Tn-T for internal homologies or for homologies with other myofibrillar proteins has been inconclusive. Short regions of similar sequence in Tn-T, for example, are observed between residues 62–69, 83–90, and 135–142. In addition, there is a resemblance in sequence between the sites phosphorylated by phosphorylase kinase in Tn-T and Tn-I (see Fig. 4) as well as some indication of similarity between residues 41–46 of Tn-I and residues 119–124 of Tn-T. Whether these and other possible similarities are statistically significant and are indicative of gene duplication must await a computer analysis of these sequences.

We thank Dr. J. H. Collins and Prof. S. V. Perry for copies of manuscripts prior to publication. We acknowledge the expert technical assistance of Mr. M. Nattris and Mrs. Lois Serink. We are grateful to Dr. R. McElhaney and Mr. S. Cook for their help in determining the NH2-terminal acetyl group. This work was supported by a Medical Research Council of Canada postdoctoral fellowship to J.P. and a grant to the M.R.C. Group of Protein Structure and Function in this department.