The Amino Acid Sequence of Rabbit Cardiac Troponin I

By ROGER J. A. GRAND and J. MICHAEL WILKINSON
Department of Biochemistry, University of Birmingham, P.O. Box 363,
Birmingham B15 2TT, U.K.

and LAWRENCE E. MOLE
Medical Research Council Immunochemistry Research Unit, Department of Biochemistry,
University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

(Received 25 June 1976)

The complete amino acid sequence of troponin I from rabbit cardiac muscle was determined by the isolation of four unique CNBr fragments, together with overlapping tryptic peptides containing radioactive methionine residues. Overlap data for residues 35–36, 93–94 and 140–145 are incomplete, the sequence at these positions being based on homology with the sequence of the fast-skeletal-muscle protein. Cardiac troponin I is a single polypeptide chain of 206 residues with mol.wt. 23550 and an extinction coefficient, $E_{280}^{1\%,1\text{cm}}$, of 4.37. The protein has a net positive charge of 14 and is thus somewhat more basic than troponin I from fast-skeletal muscle. Comparison of the sequences of troponin I from cardiac and fast skeletal muscle show that the cardiac protein has 26 extra residues at the *N*-terminus which account for the larger size of the protein. In the remainder of sequence there is a considerable degree of homology, this being greater in the *C*-terminal two-thirds of the molecule. The region in the cardiac protein corresponding to the peptide with inhibitory activity from the fast-skeletal-muscle protein is very similar and it seems unlikely that this is the cause of the difference in inhibitory activity between the two proteins. The region responsible for binding troponin C, however, possesses a lower degree of homology. Detailed evidence on which the sequence is based has been deposited as Supplementary Publication SUP 50072 (20 pages), at the British Library Lending Division, Boston Spa, Wetherby, West Yorkshire LS23 7QB, U.K., from whom copies may be obtained on the terms given in *Biochem. J.* (1976) 153, 5.

The troponin complex, consisting of the proteins troponin C, troponin I and troponin T, together with tropomyosin is responsible for the control of actomyosin ATPase* activity, and hence of contraction, in striated muscle. Greaser *et al.* (1972) and Tsukui & Ebashi (1973) have shown that the components of troponin in cardiac muscle are different from those in fast skeletal muscle on the basis of their molecular weights, and Syska *et al.* (1974) have shown that the three troponin I species isolated from fast and slow skeletal and cardiac muscles differ not only in their molecular size, but also in their inhibitory activity and in the electrophoretic mobility of the complex formed with troponin C. Amino acid-sequence studies (Collins *et al.*, 1973; van Eerd & Takahashi, 1975) confirm that differences exist between troponin C from fast skeletal and cardiac muscle, and it appears

possible that one of the four proposed Ca^{2+} -binding sites in fast-skeletal-muscle troponin C is absent from the cardiac protein. It thus seems likely that each muscle type contains a different set of troponin components which alter the characteristics of the regulation exercised and hence modify contractile activity.

We have published the amino acid sequence of troponin I from rabbit fast skeletal muscle (Wilkinson & Grand, 1975) and this information has been used to define those regions of the molecule responsible for the inhibitory activity and for binding troponin C (Syska *et al.*, 1976). In view of the higher molecular weight and lower inhibitory activity of cardiac troponin I (Syska *et al.*, 1974), we have now determined the amino acid sequence of the cardiac protein in order to investigate the relationship between the two proteins. The present paper presents the evidence for this sequence obtained by cleavage of the protein into four fragments by CNBr treatment together with the isolation of appropriate overlapping peptides.

* Abbreviations: ATPase, adenosine triphosphatase; EGTA, ethanedioxybis(ethylamine)tetra-acetate; Dnp-, 2,4-dinitrophenyl-.

Materials and Methods

Cardiac troponin I was prepared from frozen rabbit hearts (purchased from the Buxted Rabbit Co., Buxted, Sussex, U.K.) by the affinity-chromatographic method of Syska *et al.* (1974). Frozen hearts (50g) were homogenized with 500ml of 9M-urea/75mM-Tris/50mM-HCl/15mM-2-mercaptoethanol/1mM-CaCl₂, pH 8.0, and then dialysed for 6h against the same buffer. The homogenate was stirred for 16h with 60g of troponin C-Sepharose (containing 2.5–3.0mg of troponin C/g of Sepharose). The Sepharose affinity absorbent was collected by centrifugation at 1500g for 15 min, resuspended in the above buffer and again centrifuged and finally packed into a column (3cm×10cm). The column was washed with buffer until the E_{280} of the eluate was the same as that of the starting buffer and the troponin I was then eluted with the same buffer containing 10mM-EGTA. The protein was desalted on a column (2cm×100cm) of Sephadex G-25 eluted with 10mM-HCl, and finally freeze-dried. The yield of troponin I was approx. 25mg from 50g of hearts.

Troponin C-Sepharose

Troponin C, prepared from rabbit fast skeletal muscle by the method of Perry & Cole (1974), was linked to Sepharose 4B as described by Syska *et al.* (1974).

CNBr cleavage

Cardiac troponin I was dissolved (15mg/ml) in 70% (v/v) formic acid, and 100-fold molar excess of CNBr over methionine residues was added. After 24h at room temperature (20°C) the digest was diluted 10-fold with water and freeze-dried.

Radioactive labelling

Cysteine residues were carboxymethylated with iodo[¹⁴C]acetic acid in 6M-guanidine hydrochloride/0.4M-Tris/HCl, pH 8.2, as described by Wilkinson *et al.* (1972). CNBr-cleaved peptides were alkylated in this way after CNBr cleavage.

Methionine residues were alkylated with iodo[¹⁴C]-acetic acid by the method of Wilkinson (1968) after alkylation of the cysteine residues with iodoacetic acid.

Molecular-weight determination

The method of Mann & Fish (1972), using gel filtration on Sepharose 6B in the presence of 6M-guanidine hydrochloride, was used. A column (90cm×1.5cm), packed with Sepharose 6B, was eluted with 6M-guanidine hydrochloride/50mM-sodium acetate, pH 4.8. Guanidine hydrochloride (Sigma grade I) was obtained from Sigma (London) Chemical Co., Kingston-upon-Thames, Surrey KT2 7BH, U.K. and was decolorized with charcoal and filtered before use.

Before chromatography all proteins were reduced with 0.1M-2-mercaptoethanol in 6M-guanidine hydrochloride/0.4M-Tris/HCl, pH 8.0, and alkylated with 0.1M-iodoacetate. Protein samples, containing 20% (w/v) sucrose, 0.2% Dextran Blue and 0.1% Dnp-glycine, were applied to the column in a volume of 0.5ml. A calibration curve was constructed by using transferrin (mol.wt. 77000), bovine serum albumin (mol.wt. 66500), ovalbumin (mol.wt. 44000), chymotrypsinogen (mol.wt. 25700), haemoglobin (mol.wt. 15500) and lysozyme (mol.wt. 14300).

Analytical methods

Amino acid analysis was carried out as described by Wilkinson *et al.* (1972), cysteine was determined as cysteic acid by the method of Moore (1963) and tryptophan was determined spectrophotometrically as described by Crumpton & Wilkinson (1963). Extinction coefficients were determined by the method of Babul & Stellwagen (1969). Dansylation of proteins for N-terminal amino acid determination was carried out as described previously (Wilkinson, 1974). Polyacrylamide-gel electrophoresis in sodium dodecyl sulphate and in urea at pH 3.2 were performed as described by Wilkinson *et al.* (1972).

Dephosphorylation of CNBr fragment CCN3

Fragment CCN3 (0.6μmol) was incubated overnight at 37°C in 1ml of 50mM-Tris/HCl/10mM-CaCl₂, pH 8.3, containing 1mg of alkaline phosphatase (Worthington) (obtained from Cambrian Chemicals Ltd., Beddington Farm Road, Croydon CR0 4XB, Surrey, U.K.). CaCl₂ was included to remove, by precipitation, any phosphate formed during the reaction, as this acts as a competitive inhibitor of the enzyme. The digest was chromatographed on a column (2.5cm×100cm) of Sephadex G-50, eluted with 10mM-HCl, to separate the enzyme from the CNBr fragment. The fragment was freeze-dried before digestion with chymotrypsin.

Enzymic digests

For tryptic digestion of the whole protein, solid NH₄HCO₃ was added to a solution (10mg/ml) of troponin I in water to give a final concentration of 1%. The protein precipitated at this point and trypsin was added to give an enzyme/substrate ratio of 1:50 (w/w). The digest was incubated for 3h at 37°C, during which time the precipitate disappeared, and was then applied directly to a column (2.2cm×110cm) of Sephadex G-50 eluted with 50mM-NH₄HCO₃.

Peptides were digested with trypsin, thermolysin and chymotrypsin in 1% NH₄HCO₃ and with pepsin in 10mM-HCl at an enzyme/substrate ratio of 1:50 (w/w) at 37°C for 4h. The nomenclature used for peptides is as follows: Tr, Ch, Th and P refer to those obtained from digests with trypsin, chymotrypsin, thermolysin and pepsin respectively. Subsequent

letters and numerals refer to the purification steps used.

Amino acid-sequence determination

The complete amino acid sequence was obtained from the sequence of the CNBr fragments, determined by subdigestion with trypsin, chymotrypsin, thermolysin and pepsin. The CNBr fragments were overlapped by isolation of methionine-containing tryptic peptides. Peptides were separated, after enzymic digestion, either by chromatography on columns of Sephadex G-50 (2.2 cm × 110 cm) or G-25, (2.5 cm × 100 cm), eluted with 50 mM-NH₄HCO₃, followed by high-voltage paper electrophoresis, or directly by high-voltage paper electrophoresis on Whatman no. 1 or 3MM paper (Wilkinson, 1974). Radioactive peptides were detected by radioautography by using Kodak Blue Brand X-ray film.

The methods of specific staining for tyrosine-, tryptophan- and histidine-containing peptides, for the use of the dansyl-Edman technique, hydrazinolysis and the assignment of amide groups have been described previously (Wilkinson & Grand, 1975). Peptides with blocked *N*-terminal residues were detected, after paper electrophoresis, by the method of Pan & Dutcher (1956).

The first 40 residues of CNBr fragment CCN2 were determined with a Beckman 890C automatic sequencer by using the dimethylalkylamine fast peptide programme (Beckman Instruments, Glenrothes, Fife, Scotland, U.K.). The phenylthiohydantoin derivatives released on sequence analysis were identified by g.l.c. (Hewlett-Packard 5830A automatic gas chromatograph), by t.l.c. on polyamide plates (Summers *et al.*, 1973) and by back-hydrolysis with HI followed by amino acid analysis with a Rank Hilger analyser (Rank Precision Industries, Margate, Kent, U.K.).

Results

Properties of cardiac troponin I

Cardiac troponin I, prepared by the single-step affinity-chromatographic method of Syska *et al.* (1974), was shown to be homogeneous by polyacrylamide-gel electrophoresis both in sodium dodecyl sulphate and in 6M-urea, pH 3.2. An average yield of 25 mg of protein was obtained from 50 g of rabbit hearts. If the protein eluted from the affinity column was dialysed before freeze-drying, a brownish product was obtained which absorbed strongly in the 260 nm region; this was not so if urea and buffer salts were removed by desalting on Sephadex G-25, and this technique was adopted for routine use.

Molecular weights of 24200 and 21800 for cardiac and skeletal-muscle troponin I were determined by chromatography on Sepharose 6B in the presence of 6M-guanidine hydrochloride. The accuracy of the

method is estimated by Mann & Fish (1972) to be ±7%.

The amino acid composition of the protein was determined on duplicate samples hydrolysed for 24 and 72 h. The results are given in Table 1. The extinction coefficient, $E_{280}^{1\%,1\text{cm}}$, was 4.37. No free *N*-terminal amino acid could be detected by the dansyl procedure.

Separation of CNBr fragments

From the analysis of cardiac troponin I, four or five peptides would be expected to result from CNBr cleavage. The CNBr digest was chromatographed on Sephadex G-75 after radioactive labelling of the cysteine residues, and the elution profile is shown in Fig. 1. Fractions indicated were pooled, desalted on Sephadex G-10 in 10 mM-HCl and freeze-dried. Polyacrylamide-gel electrophoresis at pH 3.2 gave a single band for fractions 1, 2 and 4 and two bands for fraction 3. Fraction 3 was rechromatographed on a column (3.0 cm × 195 cm) of Sephadex G-50 eluted with 10 mM-HCl. Two peaks were obtained, which were pooled and freeze-dried. Gel electrophoresis of these two fractions showed them to be essentially homogeneous. Fraction 4 gave one major band on high-voltage paper electrophoresis at pH 6.5. This band had a mobility of 0.52 relative to aspartic acid = -1 and was purified by preparative electrophoresis.

The composition of the fractions is given in Table 2. Fractions 1, 2, 3a, 3b and 4 correspond to fragments CCN1, 2, 3, 4 and 5 respectively.

Table 1. Amino acid composition of troponin I from cardiac and fast skeletal muscle

	Cardiac troponin I (mol/24000 g)	Fast-skeletal-muscle troponin I* (mol/21000 g)
Asp	17.7	17.1
Thr	10.1	3.3
Ser	8.9	8.9
Glu	32.2	33.2
Pro	6.2	6.3
Gly	11.2	7.8
Ala	24.5	15.2
Val	8.5	7.1
Met	3.2	7.4
Ile	6.7	4.9
Leu	22.5	17.9
Tyr	3.3	2.0
Phe	4.1	2.8
His	3.1	3.6
Lys	23.5	25.8
Arg	25.0	14.3
Cys	2.0	2.7
Trp	1.0	0.9
Total	213.7	181.2

* Values taken from Wilkinson (1974).

Fig. 1. Chromatography of a CNBr digest of troponin I

The digest was applied to a column (2.2cm×140cm) of Sephadex G-75 in 6M-urea/0.2M-sodium formate, pH 3.5. —, E_{280} ; ----, radioactivity. Horizontal bars indicate the fractions that were pooled.

Amino acid sequence of CNBr fragments

The complete amino acid sequence of cardiac troponin I is shown in Fig. 2, which also shows the position of each of the CNBr fragments. All references to residue numbers refer to their position in the complete sequence.

Fragment CCN2 was submitted to automatic Edman degradation and unambiguous identification of the first 40 residues in its sequence was obtained, except for cysteine at position 75. The next three residues were tentatively identified, but it was not possible to assign acidic or amide side chains to residues 89 or 90.

Fragment CCN2 was digested with trypsin, chymotrypsin and thermolysin and the resulting peptides were purified and sequenced by manual procedures; these peptides provide evidence for the remainder of the sequence and also confirm the sequenator data.

Evidence for the sequence of fragment CCN3 was obtained from peptides from digests of the fragment with trypsin, chymotrypsin and thermolysin, for fragment CCN4 from digests with trypsin and thermolysin and for fragment CCN5 with trypsin alone.

Tryptic digestion of the CNBr fragment CCN1 gave rise to the peptides X-Ala-Asp-Glu-Ser-Arg and Ile-Ser-Ala-Asp-Ala-Hse, which are the N- and C-terminal peptides of fragments CCN3 and CCN2

Table 2. Amino acid composition of CNBr fragments from cardiac troponin I

The first column of values for each fragment gives the composition as determined by analysis in mol/mol of peptide. The second column gives the composition calculated from the amino acid sequence. The final column gives the total composition of the protein calculated from the sequence. + indicates the presence of Cys, Trp or Hse in a peptide identified by radioactivity, u.v. absorption or partial resolution on the amino acid analyser respectively.

Fragment	...	CCN1	CCN2		CCN3		CCN4		CCN5		
		Analysis	Analysis	Sequence	Analysis	Sequence	Analysis	Sequence	Analysis	Sequence	Total
Asp		12.0	8.8	7	3.4	3	6.3	4	—	—	14
Asn			1	1	—	—	2	2	—	—	3
Thr		7.4	5.2	5	2.0	2	2.6	3	—	—	10
Ser		6.5	2.6	2	4.0	5	0.9	1	—	—	8
Glu		21.2	16.1	12	4.2	3	5	5	2.0	2	22
Gln			6	6	1	1	2	2	—	—	9
Pro		5.5	2.2	2	2.6	3	—	—	—	—	5
Gly		8.4	5.5	5	1.6	1	2.9	3	2.0	2	11
Ala		18.9	10.0	10	8.7	10	3.2	3	—	—	23
Val		7.1	4.6	5	1.3	1	1.7	2	—	—	8
Ile		6.1	5.0	5	1.0	1	1.0	1	—	—	7
Leu		16.8	12.2	13	3.4	3	6.7	7	—	—	23
Tyr		2.6	1.0	1	1.1	1	—	—	—	—	2
Phe		3.1	3.0	3	—	—	—	—	0.9	1	4
His		2.5	1.1	1	0.9	1	1.0	1	—	—	3
Lys		14.3	8.7	8	5.8	6	6.6	6	3.0	3	23
Arg		17.6	12.1	13	5.9	6	4.0	4	1.1	1	24
Cys		+	+	2	—	—	—	—	—	—	2
Trp		—	—	—	—	—	+	1	—	—	1
Hse		+	+	1	+	1	+	1	—	—	3

Fig. 2. Amino acid sequence of rabbit cardiac troponin I

The position of the N-terminus of each CNBr fragment is indicated.

respectively. In addition several other peptides characteristic of these fragments were obtained. It is therefore concluded that fragment CCN1 is a partial cleavage product comprising fragments CCN3 and CCN2 in which cleavage at methionine-48 has not occurred.

Sequencing of all the small peptides was straightforward except for two peptides from fragment CCN3. The *N*-terminal tryptic peptide of the fragment, CCN3 Tr2, was ninhydrin negative and possessed no free *N*-terminal amino acid. It was therefore taken to be the *N*-terminal peptide of the whole molecule and hence fragment CCN3 to be the *N*-terminal CNBr fragment. Digestion of this peptide with pepsin for 16h followed by chromatography of the digest on Zeo-Karb 225 (H^+ form) gave rise to one peptide of composition (Asp,Ala), which was not retarded by the column, and one peptide of composition (Ser,Glu,-Arg), which was eluted with $1M-NH_3$. Hydrazinolysis of the dipeptide showed the *C*-terminal residue to be aspartic acid and hence the sequence to be X-Ala-Asp, and the sequence of the tripeptide was shown to be Glu-Ser-Arg by the Dansyl-Edman procedure. Thus the sequence of the whole peptide is X-Ala-Asp-Glu-Ser-Arg, where X is a blocking group.

Fig. 3. Chromatography of the methionine-containing tryptic peptides of troponin I

A tryptic digest of troponin I, with the methionine residues alkylated with iodo[^{14}C]acetic acid, was applied to a column ($1.8cm \times 140cm$) of Sephadex G-50 in $50mM-NH_4HCO_3$. —, E_{215} ; ····, E_{280} ; ----, radioactivity. Horizontal bars indicate the fractions which were pooled.

Peptides CCN3 Th9 (residues 17–24) and CCN3 Tr4 (residues 19–23) presented a rather different problem. In both cases it was not possible to sequence after residue 19 by the dansyl-Edman procedure, probably owing to phosphorylation of serine-20 (Solaro *et al.*, 1976).

In fact the mobilities of both peptides were consistent with the serine residue bearing a negative charge. Therefore before digestion of fragment CCN3 with chymotrypsin, the fragment was treated with alkaline phosphatase. The peptide CCN3 Ch5a (residues 18–24) was then isolated with a mobility consistent with the serine residue being uncharged. Sequencing of the dephosphorylated peptide was then possible by the normal procedures.

Isolation of methionine-containing peptides

To provide overlap information for the CNBr fragments, troponin I in which the methionine residues had been labelled with iodo[^{14}C]acetic acid was digested with trypsin. The digest was chromatographed on Sephadex G-25, eluted with $50mM-NH_4HCO_3$ as shown in Fig. 3 and fractions were pooled as indicated. Only two fractions, 3 and 4, were radioactive, and from these, three radioactive peptides were purified by high-voltage electrophoresis. One of these peptides, T3-1 (residues 145–158) had approximately twice the specific radioactivity of the other two and also gave twice the amount of methionine on analysis. It was shown to contain a Met-Met sequence (residues 150 and 151), thus giving a total of four methionine residues in the protein. These methionine sequences are shown in Fig. 4.

These data, in conjunction with the fact that the *N*-terminus of fragment CCN3 is blocked and that fragment CCN5 contains no homoserine, allow the CNBr fragments to be aligned in the order: CCN3-CCN2-CCN4-CCN5.

Cysteine sequences

As shown in Fig. 1, fragments CCN1 and CCN2 were the only CNBr fragments that contained radioactive *S*-carboxymethylcysteine after labelling of the cysteine residues of the digest. As fragment CCN1 is a partial cleavage product, all the cysteine must be located in fragment CCN2. A tryptic digest of troponin I whose cysteine residues had been labelled with iodo[^{14}C]acetic acid, followed by chromatography on Sephadex G-25 in $50mM-NH_4HCO_3$, gave rise to a profile virtually identical with that in Fig. 3, except that all the radioactivity was in a single peak corresponding to fraction 2. From this fraction a single radioactive peptide Tr2 (residues 75–93) was isolated which had *N*-terminal *S*-carboxymethylcysteine. Digestion of this peptide with chymotrypsin gave rise to two radioactive peptides Tr2C1 (residues 75–80) and Tr2C2 (residues 86–93), thus locating the two cysteine residues in the sequence at residues 75 and 92.

T3-1	Ile-Ser-Ala-Asp-Ala-Met-Met-Gln-Ala-Leu-Leu-Gly-Thr-Arg
T4-1	Asn-Ile-Asp-Leu-Leu-Ser-Gly-Met-Glu-Gly-Arg
T4-2	Thr-Leu-Met-Leu-Gln-Ile-Ala-Lys

Fig. 4. Amino acid sequence of the methionine-containing tryptic peptides of cardiac troponin I

Supplementary publication

The detailed data on which the sequence presented here is based have been deposited as Supplementary Publication SUP 50072 at the British Library (Lending Division). This publication gives details of the composition and electrophoretic mobility of the peptides isolated from enzymic digests of cardiac troponin I and its CNBr fragments. Details of both automatic and manual Edman degradations are given together with a discussion of the deduction of the sequence from the evidence provided.

Discussion

Despite the difficulty of obtaining a sufficient quantity of protein for sequencing, troponin I from rabbit heart was chosen for study so that the comparison of cardiac and fast-skeletal-muscle sequences should not be complicated by interspecies differences. To this end the affinity-chromatographic method of Syska *et al.* (1974) was used for preparation, for although it was not possible to obtain more than an average of 25 mg of protein in any one batch, the protein obtained by this single-step procedure was of very high purity.

From the evidence presented here it has been possible to deduce a unique amino acid sequence of 206 amino acids. This gives a calculated mol.wt. of 23 550, which is in good agreement with the value of 24 200 determined by chromatography on Sepharose 6B, but is considerably less than the previously determined values of 30 000, 28 000 and 29 000, obtained by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis by Greaser *et al.* (1972), Tsukui & Ebashi (1973) and Syska *et al.* (1974) respectively. The protein is 27 residues longer than the analogous protein from fast skeletal muscle, and the relationship between these two sequences is discussed below. Cardiac troponin I has an overall net positive charge of 14, including histidine residues, and is thus a somewhat more basic protein than its fast-skeletal-muscle counterpart. In common with other myofibrillar proteins, cardiac troponin I has no free α -amino group. The nature of the blocking group has not been determined, but it seems most likely that it is an acetyl group in view of the results obtained from other myofibrillar proteins.

Validity of the proposed sequence

The supplementary publication to this paper (SUP 50072) presents the detailed evidence on which the sequence is based and with a few exceptions these data contain adequate overlaps to confirm the ordering of the peptides. There are three positions at which it has not been possible to obtain overlap data; these are between residues 35 and 36, between residues 93 and 94 and between residues 140 and 145.

The sequence around residue 35 is Lys-Lys-Lys-Leu. Both trypsin and thermolysin cleaved after the third lysine residue, whereas chymotrypsin cleaved after the first lysine, but no subsequent peptide could be isolated from this digest. The sequence is thus particularly difficult to establish unambiguously and relies on the positioning of all the other tryptic peptides which account for the total composition of the fragment.

A similar form of reasoning applies to the other two cases, but for these positions the argument is strengthened by the homology of the sequence with that of the fast-skeletal-muscle troponin I as shown in Fig. 5. The sequence from residues 140 to 145 has been taken to be Arg-Leu-Arg-Val-Arg-Ile; this rests basically on homology with the fast-skeletal-muscle sequence, but the order of Leu-Arg and Val-Arg could clearly be reversed.

Relationship with fast-skeletal-muscle troponin I

A comparison of the sequences of cardiac and fast-skeletal-muscle troponin I is shown in Fig. 5. It is necessary to introduce three deletions in the fast-skeletal-muscle sequence, at residues 39, 83 and 183, in order to obtain the best fit between the two sequences.

The main differences to be observed between the two proteins is that the cardiac protein has 26 extra residues at the *N*-terminus, and at the *C*-terminus the fast-skeletal-muscle protein has two extra residues. The extra *N*-terminal sequence contains the unusual sequence Pro-Ala-Pro-Ala (residues 13–16), which is reminiscent of the sequence of the A1 light chain of myosin (Frank & Weeds, 1974). Solaro *et al.* (1976) have shown that serine at residue 20 is phosphorylated during perfusion and that this phosphate can be readily exchanged. In view of their findings it seems probable that the extra length of polypeptide chain is

Cardiac	[A D E S R D A A G E A K P A P A V R R S D R A Y A	25
Fast		
Cardiac	T E P H A <u>K</u> S K <u>K</u> K <u>I</u> <u>S</u> <u>A</u> S <u>R</u> <u>K</u> L Q <u>L</u> <u>K</u> <u>T</u> <u>L</u> <u>M</u> <u>L</u> <u>Q</u>	50
Fast	[G D E E K R N R A I T A - R R Q H L K S V M L Q	
Cardiac	<u>I</u> <u>A</u> K Q <u>E</u> <u>L</u> <u>E</u> <u>R</u> <u>E</u> A E E <u>R</u> R <u>G</u> <u>E</u> <u>K</u> G R A <u>L</u> S T R <u>C</u>	75
Fast	I A A T E L E K E E G R R E A E K Q N Y L A E H C	
Cardiac	Q <u>P</u> <u>L</u> E <u>L</u> A <u>G</u> L G <u>F</u> <u>A</u> <u>E</u> <u>L</u> <u>Q</u> <u>D</u> <u>L</u> <u>C</u> <u>R</u> <u>Q</u> <u>L</u> <u>H</u> <u>A</u> <u>R</u> <u>V</u> <u>D</u>	100
Fast	P P L S L P G - S M A E V Q E L C K Q L H A K I D	
Cardiac	K V <u>D</u> <u>E</u> <u>E</u> <u>R</u> <u>Y</u> <u>D</u> <u>V</u> <u>E</u> A <u>K</u> <u>V</u> T <u>K</u> N I T <u>E</u> <u>I</u> A <u>D</u> <u>L</u> <u>T</u> <u>Q</u>	125
Fast	A A E E E K Y D M E I K V Q K S S K E L E D M N Q	
Cardiac	<u>K</u> <u>I</u> <u>F</u> <u>D</u> <u>L</u> <u>R</u> <u>G</u> <u>K</u> <u>F</u> <u>K</u> <u>R</u> <u>P</u> <u>T</u> <u>L</u> <u>R</u> <u>L</u> <u>R</u> <u>V</u> <u>R</u> <u>I</u> <u>S</u> <u>A</u> <u>D</u> <u>A</u> <u>M</u>	150
Fast	K L F D L R G K F K R P P L R R R V R M S A D A M	
Cardiac	<u>M</u> <u>Q</u> <u>A</u> <u>L</u> <u>L</u> <u>G</u> <u>T</u> <u>R</u> A <u>K</u> E T <u>L</u> <u>D</u> <u>L</u> <u>R</u> <u>A</u> <u>H</u> <u>L</u> <u>K</u> <u>Q</u> <u>V</u> <u>K</u> <u>K</u> <u>E</u>	175
Fast	L K A L L G S K H K V C M D L R A N L K Q V K K E	
Cardiac	<u>D</u> <u>T</u> <u>E</u> <u>K</u> <u>E</u> N <u>R</u> <u>E</u> <u>V</u> <u>G</u> <u>D</u> <u>W</u> <u>R</u> <u>K</u> <u>N</u> <u>I</u> <u>D</u> <u>L</u> <u>L</u> <u>S</u> <u>G</u> <u>M</u> <u>E</u> <u>G</u> <u>R</u>	200
Fast	D T E K E - R D V G D W R K N I E E K S G M E G R	
208		
Cardiac	<u>K</u> <u>K</u> <u>K</u> <u>F</u> <u>E</u> G	
Fast	K K M F E S E S	

Fig. 5. Comparison of the amino acid sequence of cardiac and fast-skeletal-muscle troponin I

Residues underlined with a solid line are identical; those with a dashed line are conservative replacements. Single-letter code for amino acids: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; W, Trp; Y, Tyr.

situated on the surface of the molecule and has a functional significance in the heart but that it is not required in the fast-skeletal-muscle protein.

If the remainder of the two proteins are compared, 105 residues out of a possible 179 are identical in both sequences, that is a difference of 41%. If consideration

is also given to positions that require only a single base change giving rise to a structurally similar amino acid, these values become 133 and 26% respectively. Thus apart from the *N*-terminal region of cardiac troponin, there is a strong overall homology between the two proteins.

It is noticeable that the sequence homology is considerably stronger in the C-terminal part of the sequence. This appears to begin at residue 85 and with some interesting exceptions continues to the C-terminus of both proteins. There is one area in particular in this region, residues 114–118, where homology breaks down. This may be a linking region unimportant for the structural integrity of the molecules, or it may impose an important structural difference which affects the biological activity. In the N-terminal part of the sequence the homology is much less strong except in the region residues 44–59. There is a single proline at residue 77, in contrast with the Pro-Pro sequence in the fast-skeletal-muscle protein; however, this will probably have a similar effect on the structure at this point.

There are only two cysteine residues in the cardiac sequence, at residues 75 and 92, these being homologous with cysteines in the fast-skeletal-muscle sequence. The third cysteine in the fast-skeletal-muscle sequence, at residue 162, has been replaced by threonine in the cardiac protein.

Four sites in fast-skeletal-muscle troponin I have been shown to be phosphorylated under various conditions (Huang *et al.*, 1974; Moir *et al.*, 1974); these are at positions corresponding to residues 37, 46, 116 and 146 in the cardiac sequence. Of these residues, 116 is asparagine in the cardiac sequence and thus cannot be phosphorylated. There is evidence that serine-146 can be phosphorylated by a cyclic AMP-dependent protein kinase (A. J. G. Moir, personal communication), but there is as yet no evidence that serine-37 or threonine-46 are phosphorylated, although the sequences around these residues are very similar to the corresponding sequences in the fast-skeletal-muscle protein, as indeed is that around serine-146.

The two important interactions of troponin I are with actin to produce an inhibition of the actomyosin ATPase, and with troponin C, in the presence of Ca^{2+} , to relieve this inhibition. It has been shown previously (Syska *et al.*, 1976) that the inhibitory activity of fast-skeletal-muscle troponin I can be localized in a CNBr peptide corresponding to residues 124–145 in the cardiac sequence, whereas the region responsible for binding troponin C is located in the N-terminal peptide corresponding to cardiac residues 27–74. As cardiac troponin I shares these two activities with the fast-skeletal-muscle protein it is of interest to compare these regions in particular. The inhibitory region is almost identical in the two proteins, the major differences being a threonine-for-proline substitution at residue 138 and a leucine for arginine at residue 141. Syska *et al.* (1974) have reported a considerable difference in the inhibitory activity of the parent proteins, but it is not clear if these differences in the primary structure could be responsible for this

effect. The situation with the troponin C-binding region is somewhat different, in that out of 48 residues only 19 are identical, and on the basis of homology data it appears that the most likely troponin C-binding region is between residues 36 and 59, spanning the methionine at residue 48. No quantitative data on differences in the binding of troponin I to troponin C are available, but on the basis of the sequence data it would be reasonable to expect that such differences exist.

We thank Professor S. V. Perry for his support and encouragement during the course of this work and Miss Susan Brewer for excellent technical assistance. This work was supported in part by grants from the Medical Research Council and the Muscular Dystrophy Association of America Inc.

References

- Babul, J. & Stellwagen, E. (1969) *Anal. Biochem.* **28**, 216–221
- Collins, J. H., Potter, J. D., Horn, M. J., Wilshire, G. & Jackman, N. (1973) *FEBS Lett.* **36**, 268–272
- Crumpton, M. J. & Wilkinson, J. M. (1963) *Biochem. J.* **88**, 228–234
- Frank, G. & Weeds, A. G. (1974) *Eur. J. Biochem.* **44**, 317–334
- Greaser, M. L., Yamaguchi, M., Brekke, C., Potter, J. & Gergely, J. (1972) *Cold Spring Harbor Symp. Quant. Biol.* **37**, 235–244
- Huang, T. S., Bylund, D. B., Stull, J. T. & Krebs, E. G. (1974) *FEBS Lett.* **42**, 249–252
- Mann, K. G. & Fish, W. W. (1972) *Methods Enzymol.* **26**, 28–42
- Moir, A. J. G., Wilkinson, J. M. & Perry, S. V. (1974) *FEBS Lett.* **42**, 253–256
- Moore, S. (1963) *J. Biol. Chem.* **238**, 235–237
- Pan, S. C. & Dutcher, J. D. (1956) *Anal. Chem.* **28**, 836–838
- Perry, S. V. & Cole, H. A. (1974) *Biochem. J.* **141**, 733–743
- Solaro, R. J., Moir, A. J. G. & Perry, S. V. (1976) *Nature (London)* **262**, 615–617
- Summers, M. R., Smythers, G. W. & Oroszlan, S. (1973) *Anal. Biochem.* **53**, 624–628
- Syska, H., Perry, S. V. & Trayer, I. P. (1974) *FEBS Lett.* **40**, 253–257
- Syska, H., Wilkinson, J. M., Grand, R. J. A. & Perry, S. V. (1976) *Biochem. J.* **153**, 375–387
- Tsukui, R. & Ebashi, S. (1973) *J. Biochem. (Tokyo)* **73**, 1119–1121
- van Eerd, J.-P. & Takahashi, K. (1975) *Biochem. Biophys. Res. Commun.* **64**, 122–127
- Wilkinson, J. M. (1968) *FEBS Lett.* **4**, 170–172
- Wilkinson, J. M. (1974) *Biochim. Biophys. Acta* **359**, 379–388
- Wilkinson, J. M. & Grand, R. J. A. (1975) *Biochem. J.* **149**, 493–496
- Wilkinson, J. M., Perry, S. V., Cole, H. A. & Trayer, I. P. (1972) *Biochem. J.* **127**, 215–228