Volume 228, number 1, 22-26 FEB 05527 February 1988

Cloning, sequencing and expression of a full-length rabbit fast skeletal troponin-C cDNA

Qian Chen*, Janos Taljanidisz‡, Satyapriya Sarkar†*, Terence Tao*+ and John Gergely*°

*Department of Muscle Research, ‡Department of Metabolic Regulation, Boston Biomedical Research Institute, + Department of Neurology, °Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School and Neurology Service, Massachusetts General Hospital and †Department of Anatomy and Cellular Biology, Tufts University School of Medicine, Boston, Massachusetts, USA

Received 13 November 1987

A full-length cDNA coding for troponin-C isolated from an adult rabbit fast skeletal muscle library has been sequenced and expressed in an *E. coli* host. The amino acid sequence derived from the coding region agrees with the published protein sequence except that the first two residues are reversed. The expressed protein was found to be identical to purified rabbit skeletal troponin-C based on electrophoretic mobilities in polyacrylamide gels containing either SDS or urea, and on immunoblotting. These results establish that our troponin-C cDNA clone is suitable for site-directed mutagenesis studies on the structure and function of troponin-C

Troponin-C; cDNA; Nucleotide sequence; cDNA expression; *(E. coli)*

1. INTRODUCTION

Contraction of mammalian striated muscle is regulated by Ca^{2+} and requires the regulatory proteins troponin and tropomyosin [1]. Binding of Ca^{2+} to the low-affinity metal binding sites of troponin-C (TnC), the calcium-binding subunit of troponin, propagates conformational transitions from TnC itself to the other subunits of troponin, and ultimately to the entire thin filament so as to activate contraction [2,3]. Site-directed mutagenesis provides a means for elucidating the mechanism of the Ca^{2+}-regulation process at the molecular level. As a first step to this end, we have isolated and sequenced a full-length cDNA coding for rabbit fast skeletal TnC. We have also expressed this cDNA in an *E. coli* host, and identified the expressed protein as TnC based on electrophoretic mobilities in either SDS or urea containing polyacrylamide gels, and on an immunoblot assay.

Correspondence address: T. Tao, Dept. of Muscle Research, Boston Biomedical Research Institute, 20 Staniford St., Boston, MA 02114, USA

Abbreviations: TnC, rabbit fast skeletal troponin-C; PAGE, polyacrylamide gel electrophoresis

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00760

2. MATERIALS AND METHODS

2.1. *Materials*

Materials for library construction were from Pharmacia. Restriction enzymes and materials for sequencing were from New England Biolabs. AMV reverse transcriptase was from Life Sciences. Radioisotopes were from New England Nuclear. Materials for electrophoresis were from BioRad. TnC was purified from rabbit skeletal muscle ether powder according to [4].

2.2. *Cloning and sequencing*

For library construction, total RNA was isolated from adult rabbit fast skeletal muscle according to [5]. Poly(A)$^+$ RNA was isolated from total RNA according to [6]. The cDNA library was constructed from 10 μg of Poly(A)$^+$ RNA according to [7] using oligo(dT) tailed PSV7186 and oligo(dG) tailed PSV1932 as the vector primer and linker DNA, respectively.

The library was screened with a rabbit fast skeletal TnC

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

cDNA probe containing 94 nucleotides in the coding region (nucleotides 218–311 in fig.2) [8]. The probe was subcloned into M13 and radiolabelled by primer extension [9] in the presence of [α-^{32}P]ATP. The recombinants were screened by colony hybridization [10] using about 2–3 × 10^6 cpm of ^{32}P-labelled probe per filter, each filter containing about 1000 recombinants.

For Southern blot analyses [11] plasmid DNA isolated from clones of individual positive recombinant bacteria was cut with the restriction enzymes *Hin*dIII and *Pvu*II or *Hin*dIII and *Acc*I [7], electrophoresed in 1% agarose gels, and transferred to nitrocellulose paper. The blots were then hybridized with the radiolabelled probe. For Northern blot analyses total RNA isolated from rabbit fast skeletal muscle was electrophoresed in 1% formaldehyde agarose gels, transferred to nitrocellulose paper and the resulting blots were hybridized [12] with ^{32}P-labelled probes prepared by nick-translation [12,13] of plasmid DNA isolated from positive recombinants.

Sequencing of the TnC cDNA was performed in both double-stranded DNA of the PSV cloning vector and single-stranded DNA of the M13mp18 and M13mp19 sequencing vectors by the dideoxy chain termination method [14] following the New England Biolabs protocol. Sequencing was initiated within the coding region by the use of synthesized (Cyclone DNA synthesizer, Biosearch) oligonucleotide primers I and II (see fig.2), corresponding to the 3′- and 5′-ends of the screening probe, and extended using primers III and IV (fig.2). Universal and other synthesized primers were used to obtain overlapping sequences on both strands.

2.3. *Expression and characterization of the product*

A vector containing the T_7 bacteriophage RNA polymerase/promoter system [15] was used for the expression of the TnC cDNA. The insert containing the complete TnC cDNA was ligated into the polylinker *Hin*dIII and *Sma*I restriction sites of the plasmid pT7-6, which contains the T_7 RNA polymerase promoter ϕ10. A second plasmid, pGP1-2, containing the T_7 RNA polymerase gene which is controlled by the temperature inducible λP_L promoter was then introduced into a competent strain of *E. coli* (BK9MDG) [16] together with the first plasmid. Transformed cells were grown at 30°C until mid-log phase was reached. The temperature was then raised to 42°C to initiate expression of the insert, while rifampicin was added to inhibit further expression of bacterial proteins. To radiolabel the expressed protein, [^{35}S]methionine was present in the M9 growth medium. To maximize expression, enriched medium (2% tryptone, 1% yeast extract, 0.5% NaCl, 0.2% glycerol, 50 mM K_2HPO_4, pH 7.2) was used.

For partial purification of the expressed material, cells were lysed in the French pressure cell press at 12000 lb/inch^2, 4°C, and solid $(NH_4)_2SO_4$ was added to the bacterial lysate to yield a 55% saturated solution. The pellet was discarded, and the pH of the supernatant was adjusted to 4.1. After stirring for 30 min at 4°C, the pellet was collected and resuspended in 20 mM Hepes, 2 mM EDTA, 10 mM dithiothreitol, pH 7.5, followed by dialysis against the same buffer. SDS and urea-EDTA PAGE were carried out according to [17] and [18], respectively.

Anti-TnC was obtained by immunizing a goat with purified rabbit fast skeletal TnC. The bleeds were pooled and the antiserum was subjected to DEAE column chromatography, yielding a purified IgG fraction. Immunoblotting [19] was carried out essentially according to the protocol provided by Millipore. After transfer of proteins from gels to PVDF membranes (Immobilon, Millipore) [20], the blots were blocked with bovine serum albumin, incubated with the goat anti-TnC, then with horseradish peroxidase-conjugated rabbit anti-goat IgG (Sigma). Color was developed with diaminobenzidine (0.8 mg/ml) and H_2O_2 (0.03%).

3. RESULTS AND DISCUSSION

The PSV cloning vector was chosen for construction of the library because Okayama and Berg [7] reported that cloning in this vector favors full-length or near full-length cDNAs. Out of ~5000 recombinants that were screened with the 94 nucleotide TnC cDNA probe, 35 positive colonies

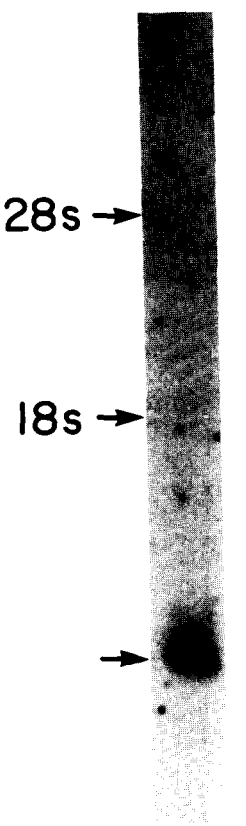

Fig.1. Hybridization of total rabbit fast skeletal RNA (15 μg) with ^{32}P-labelled nick-translated G7G TnC plasmid DNA (Northern blot). Approx. 2 × 10^6 cpm of the probe was hybridized in 6 × SSC, 1 × Denhardt solution, 0.5% SDS, 100 μg/ml ssDNA, 0.05% $Na_4P_2O_7$, pH 6.4, at 68°C for 12 h. A single band (arrow) corresponding to 741 nucleotides was obtained.

ranging in size from 350–750 nucleotides were obtained. This suggests that TnC mRNA amounts to about 0.7% of total mRNA, consistent with the known TnC content of mammalian skeletal muscle [21]. Southern blot analysis of the plasmid DNA from the recombinant with the longest insert (designated as G7G) confirmed that its size is about 750 nucleotides (not shown). Northern blot analysis of total mRNA using radiolabelled nick-translated G7G plasmid DNA yielded a single band of hybridized RNA whose size was estimated to be 740 nucleotides (fig.1). These results indicate that the G7G insert is very likely a full-length TnC cDNA.

The nucleotide sequence of the G7G insert (fig.2) contains the complete coding region of 483 nucleotides (including the initiation and termination codons), the 5′- and 3′-untranslated regions of 65 and 102 nucleotides, respectively, terminated with a poly(A) tail of about 100 nucleotides. This sequence is similar, but not identical to that reported by Zot et al. [22] for a TnC cDNA isolated from a neonatal rabbit skeletal muscle library. The derived amino acid sequence is identical to the published amino acid sequence [23] of rabbit fast skeletal muscle TnC except that the first two residues at the N-terminus are reversed. Since the N-terminus of TnC is acetylated, we believe that the determined amino acid sequence is more likely to be in error owing to the difficulty in sequencing N-terminal peptides of blocked proteins. A search for secondary structures revealed the presence of a hair-pin structure in the 3′-untranslated region. This hair-pin contains 22 nucleotides in the stem and 13 bases in the loop, flanked by 20 and 47 nucleotides in the proximal and distal ends,

```
-65  CGTCT GTGCGTGGTG GAGTTCAGGG GAGGCGGCTC ACGGCAGGGG TGGCCCAGTG ACCAGCAACC

  1  ATG ACG GAC CAG CAG GCT GAG GCC CGG TCC TAC CTC AGC GAG GAG ATG ATC GCT
     Met Thr Asp Gln Gln Ala Glu Ala Arg Ser Tyr Leu Ser Glu Glu Met Ile Ala
     ---------- III
 55  GAG TTC AAG GCC GCC TTC GAC ATG TTT GAT GCT GAT GGC GGT GGT GAC ATC AGC
     Glu Phe Lys Ala Ala Phe Asp Met Phe Asp Ala Asp Gly Gly Gly Asp Ile Ser

109  GTC AAG GAG TTG GGC ACA GTC ATG AGG ATG CTG GGC CAG ACA CCC ACC AAA GAG
     Val Lys Glu Leu Gly Thr Val Met Arg Met Leu Gly Gln Thr Pro Thr Lys Glu

163  GAG CTG GAC GCC ATC ATC GAG GAG GTG GAT GAG GAT GGC AGC GGC ACC ATC GAC
     Glu Leu Asp Ala Ile Ile Glu Glu Val Asp Glu Asp Gly Ser Gly Thr Ile Asp
          <---------------------I
217  TTC GAG GAG TTC TTG GTC ATG ATG GTG CGC CAG ATG AAA GAG GAC GCC AAG GGC
     Phe Glu Glu Phe Leu Val Met Met Val Arg Gln Met Lys Glu Asp Ala Lys Gly
                              II ---------->
271  AAG AGC GAG GAG GAG CTG GCC GAG TGC TTC CGC ATC TTC GAC AGG AAC GCA GAC
     Lys Ser Glu Glu Glu Leu Ala Glu Cys Phe Arg Ile Phe Asp Arg Asn Ala Asp

325  GGC TAC ATC GAC GCC GAG GAG CTG GCT GAG ATT TTC AGG GCC TCT GGG GAG CAC
     Gly Tyr Ile Asp Ala Glu Glu Leu Ala Glu Ile Phe Arg Ala Ser Gly Glu His
                                                                  IV----
379  GTG ACG GAC GAG GAG ATC GAA TCC CTG ATG AAG GAC GGG GAC AAG AAC AAC GAT
     Val Thr Asp Glu Glu Ile Glu Ser Leu Met Lys Asp Gly Asp Lys Asn Asn Asp
     ---------->
433  GGC CGC ATT GAC TTC GAC GAG TTC CTG AAG ATG ATG GAG GGC GTG CAG TAA
     Gly Arg Ile Asp Phe Asp Glu Phe Leu Lys Met Met Glu Gly Val Gln  *

484  GGCTGGGGTC TCCGCCTCCG GCCCCGCGCG GCTCGTTCCC TGCCCCGCGC GGGGCGGCCC CTTCT
                           .....
549  GGCTC CGTGTCTGGA AGGAATAAAA GCAAACGTTG CCAAAAAAA.....
```

Fig.2. Complete nucleotide sequence of rabbit fast skeletal TnC cDNA. Lines with arrows indicate synthesized sequencing primers and direction of sequencing. The consensus polyadenylation sequence AATAAA (dotted line) is found 14 residues from the poly(A) tail. A stem (solid line)-loop (broken line)-stem (solid line) structure is found as residues 504–539.

respectively. Interestingly, such stem-loop structures have been deduced for the 3′-untranslated regions of many prokaryotic and eukaryotic mRNAs [24].

The cloned TnC cDNA was expressed in *E. coli* using the T_7 polymerase/promoter expression vector [15] in the presence of [^{35}S]methionine and the bacterial lysate subjected to SDS-PAGE. Autoradiography of the gel (fig.3) reveals a single major band whose mobility is the same as that of purified rabbit skeletal TnC. This band does not appear in the lysate of cells that were transformed with the expression vector not containing the insert. Using molecular mass markers, the size of the expressed protein was estimated to be 18 kDa. Thus, the expressed protein appears to be identical to rabbit skeletal TnC with respect to its molecular mass.

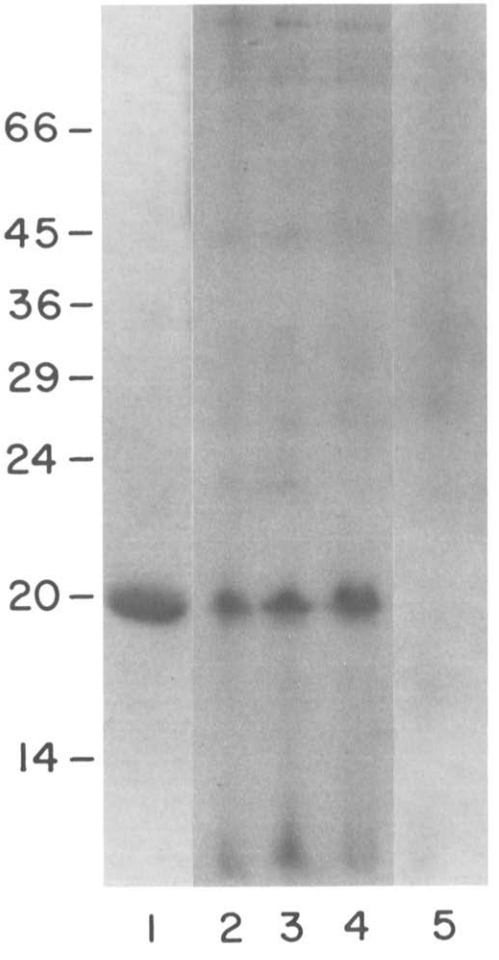

Fig.3. Autoradiograph of [^{35}S]methionine labelled expressed protein run in 12.5% SDS PAGE. Lanes: 1, purified TnC (Coomassie blue stained); 2–4, lysate of cells transformed with expression vector containing TnC cDNA, expression induced in the presence of [^{35}S]methionine for 10, 20 and 60 min, respectively; 5, control experiment, using lysate of cells transformed with expression vector that does not contain TnC cDNA. 10 μl lysate was loaded in each lane (2–5).

To enrich the expressed protein, bacterial lysate was subjected to partial purification (section 2.3), then analyzed by urea-EDTA PAGE, which separates proteins according to charge [18]. Fig.4a shows that a band of high mobility is present in the enriched lysate of cells that has been transformed with insert-containing vector. This band is barely detectable in the lysate before enrichment, is not present in the lysate of cells transformed with insert-free vector (either before or after enrichment), and has the same mobility as that of purified TnC in the same gel system. Thus, the expressed protein is indistinguishable from TnC with respect to its overall charge.

A urea gel of the partially purified expressed

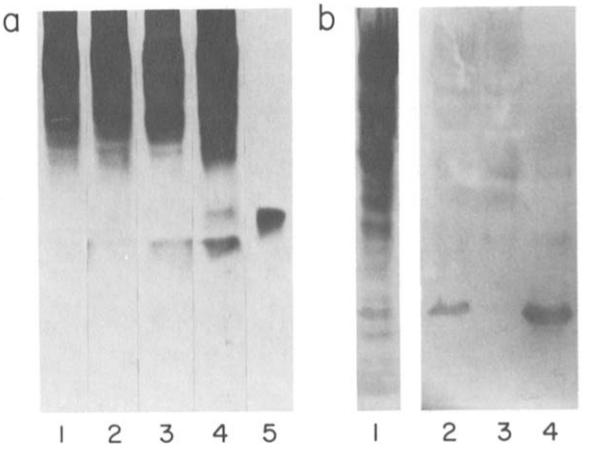

Fig.4. Analysis of expressed TnC by urea-EDTA PAGE (10% polyacrylamide in 5 M urea, 2 mM EDTA, 0.08 M Tris-glycine, pH 8.6) and immunoblotting. (A) Lanes: 1 and 3, lysate (~ 120 μg total protein) of *E. coli* cells transformed with expression vector that does not contain TnC cDNA, before and after partial purification, respectively; 2 and 4, lysate (~ 120 μg total protein) of cells transformed with expression vector that contains TnC cDNA, before and after partial purification, respectively; 5, purified TnC (10 μg). (B) Lanes: 1, Coomassie blue stained blot of bacterial lysate containing the expressed protein; 2 and 3, immunoblots of lysate from cells transformed with expression vector that does and does not contain TnC cDNA, respectively; 4, immunoblot of purified TnC.

protein was also analyzed by immunoblotting. A single band in the lysate containing the expressed protein reacts with anti-TnC (fig.4b). This band has the same mobility as that of purified TnC, and does not appear in the lysate of bacteria transformed with insert-free vector. Thus, the expressed protein is immunologically indistinguishable from TnC.

In conclusion, we have isolated and characterized a full-length cDNA clone of adult rabbit fast skeletal TnC, which when inserted into an expression vector gave rise to synthesis in *E. coli* of a protein that can be identified as TnC by several biochemical and immunological criteria. This expression of a full-length mammalian skeletal muscle TnC cDNA as reported here represents the first step towards overproduction of mutagenized TnC in order to systematically analyze the structure and function of TnC. Work along these lines is in progress in our laboratory.

Acknowledgements: We thank Drs N. Sarkar and J. Hong, in whose laboratories some of this work was conducted; Drs Z. Grabarek and M. Rosemblatt for the anti-TnC and for helpful discussions; Dr S. Putney for the TnC cDNA probe; Drs S. Tabor and C. Richardson for the pT7-6 and pGP1-2 plasmids; and Drs A. Zot, J. Potter and W. Strauss for making available a preprint of their work. Supported by grants from NIH (AR-21673, AR-13238, R37-HL-05949), and MDA.

REFERENCES

[1] Ebashi, S. and Endo, M. (1968) Prog. Biophys. Mol. Biol. 18, 123–183.
[2] Leavis, P.C. and Gergely, J. (1984) CRC Crit. Rev. Biochem. 16, 235–305.
[3] Zot, A.S. and Potter, J.D. (1987) Annu. Rev. Biophys. Biophys. Chem. 16, 535–559.
[4] Greaser, M.L. and Gergely, J. (1971) J. Biol. Chem. 246, 4223–4226.
[5] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5301.
[6] Sarkar, S., Mukherjee, A.K. and Guha, C. (1981) J. Biol. Chem. 256, 5077–5086.
[7] Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161–170.
[8] Putney, S.D., Herlihy, W.C. and Schimmel, P. (1983) Nature 302, 718–721.
[9] Messing, J. (1983) Methods Enzymol. 101, 20–78.
[10] Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961–3965.
[11] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY.
[13] Rigby, P.W.J., Diekmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237–251.
[14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[15] Tabor, S. and Richardson, C.C. (1985) Proc. Natl. Acad. Sci. USA 82, 1074–1078.
[16] Nohno, T., Saito, T. and Hong, J. (1986) Mol. Gen. Genet. 205, 260–269.
[17] Laemmli, U.K. (1970) Nature 227, 680–685.
[18] Drabikowski, W., Kuznicki, J. and Grabarek, Z. (1978) Comp. Biochem. Physiol. 60C, 1–6.
[19] Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA 76, 4350–4354.
[20] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[21] Potter, J.D. (1974) Arch. Biochem. Biophys. 162, 436–441.
[22] Zot, A.S., Potter, J.D. and Strauss, W.L. (1987) J. Biol. Chem., in press.
[23] Collins, J.H., Greaser, M.L., Potter, J.D. and Horn, M.J. (1977) J. Biol. Chem. 252, 6356–6362.
[24] Platt, T. (1986) Annu. Rev. Biochem. 55, 339–372.