Volume 270, number 1,2, 57–61 FEBS 08842 September 1990

Molecular cloning of human cardiac troponin I using polymerase chain reaction

William J. Vallins[1], Nigel J. Brand[1], Nina Dabhade[1], Gillian Butler-Browne[2], Magdi H. Yacoub[1] and Paul J.R. Barton[1]

[1]*Department of Cardiothoracic Surgery, National Heart and Lung Institute, Dovehouse Street, London SW7 6LY, UK* and [2]*INSERM U262, Clinique Universitaire Baudeloque, 123 Port Royal, 75674 Paris, France*

Received 16 July 1990

We have used the polymerase chain reaction (PCR) to synthesise a cDNA encoding part of human cardiac troponin I. Amplification was achieved using fully degenerate sets of oligonucleotides corresponding to conserved regions of amino acid sequence identified in other troponin I isoforms. The cloned PCR fragment was subsequently used to isolate full-length cDNAs from a cardiac cDNA library. We describe the approach, as a general cloning strategy starting from limited amino-acid sequence data and report the cloning, and complete amino acid sequence of human cardiac troponin I. Analysis of human development using these clones demonstrates early expression of this gene in the heart.

Polymerase chain reaction; cDNA, Troponin I; Amino acid sequence

1. INTRODUCTION

Troponins I, T and C are a group of muscle-specific myofibrillar proteins which are involved in regulating the interaction of actin and myosin. Three isoforms of troponin I (TnI) have been identified and are present as the major isoforms in cardiac muscle (TnIc), slow skeletal muscle (TnIs) and fast skeletal muscle (TnIf). The available amino acid sequence data for these three isoforms, from bovine, rabbit and chick muscle [1–4], show that they are encoded by separate genes. The cardiac isoform differs from those found in skeletal muscle in that it has an extended N-terminal sequence. This N-terminal region contains a serine residue which is phosphorylated during the phosphorylation-mediated regulation of contraction by β-adrenergic agonists in the heart [5].

Antibody studies in rat and chick have demonstrated that there is a switch in TnI isoform expression during cardiac development, and suggest that the major isoform in early fetal heart is TnIs, and that this is replaced by TnIc late in fetal development [6,7]. We are interested in the regulation of troponin I gene expression in the heart, and have undertaken to clone human cardiac troponin I. To date, no amino-acid sequence data have been available for the human TnIc isoform. We have made use of amino acid sequence data from other species, and from other isoforms of TnI, in order to predict the human TnIc sequence, and have used the polymerase chain reaction (PCR) in order to amplify corresponding human cardiac cDNA sequences. In this way we generated a partial cDNA clone which was subsequently used to isolate full length cDNAs from a human cardiac library. Here we report the approach used for cloning the cDNA, of general use in cloning from limited amino-acid sequence data, and present the complete amino acid sequence of the human cardiac troponin I isoform.

Correspondence address: P.J.R. Barton, National Heart and Lung Institute, Dovehouse Street, London SW3 6LY, UK

Abbreviations: TnI, troponin I (TnIc, TnIf and TnIs; cardiac, fast skeletal and slow skeletal muscle isoforms, respectively); PCR, polymerase chain reaction; SDS, sodium dodecyl sulphate; BSA, bovine serum albumin, bp, base pairs

The nucleotide sequence reported here has been submitted to the GenBank/EMBL Data Bank with accession number X54163

2. MATERIALS AND METHODS

2.1. *Oligonucleotide primers*

Fully degenerate sets of oligonucleotide primers (oligo-sets), complementary to three regions that are 100% conserved between the published troponin I amino acid sequences [1–4] (see Fig. 2), were synthesized and used to prime PCR reactions. Set 1 (a sense strand 22-mer, with 128-fold degeneracy) corresponds to amino acid residues 136 to 140 (RGKFK). Set 2 (antisense strand 27-mer, with 64-fold degeneracy) corresponding to amino acids 179–184 (EDTEKE). Set 3 (antisense 22-mer, with 64-fold degeneracy) corresponding to amino acids 201–205 (MEGRK). Each set carries additional nucleotides (underlined) including *Pst*I or *Eco*RI restriction sites. Dgenerate positions are indicated in parentheses.

Set 1:5′-ATCTGCAG(C/A)G(T/C/G/A)GG(T/C/G/A)AA(G/A)-TT(T/C)AA-3′

Set 2:5′-CGCGGAATTCTC(T/C)TT(T/C)TC(T/C/G/A)GT(A/G)-TC(T/C)TC-3

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

```
                                                    5'-CTGAAGGTCACCCGGGCGGC

CCCCTCACTGACCCTCCAAACGCCCCTGTCCTCGCCCTGCCTCCTGCCATTCCCGGCCTG

                MetAlaAspGlySerSerAspAlaAlaArgGluProArgProAlaProAla
AGTCTCAGCATGGCGGATGGGAGCAGCGATGCGGCTAGGGAACCTCGCCCTGCACCAGCC

ProIleArgArgArgSerSerAsnTyrArgAlaTyrAlaThrGluProHisAlaLysLys
CCAATCAGACGCCGCTCCTCCAACTACCGCGCTTATGCCACGGAGCCGCACGCCAAGAAA

LysSerLysIleSerAlaSerArglysLeuGlnLeulysThrLeuLeuLeuGlnIleAla
AAATCTAAGATCTCCGCCTCGAGAAAATTGCAGCTGAAGACTCTGCTGCTGCAGATTGCA

LysGlnGluLeuGluArgGluAlaGluGluArgArgGlyGluLysGlyArgAlaLeuSer
AAGCAAGAGCTGGAGCGAGAGGCGGAGGAGCGGCGCGGAGAGAAGGGGCGCGCTCTGAGC

ThrArgCysGlnProLeuGluLeuThrGlyLeuGlyPheAlaGluLeuGlnAspLeuCys
ACCCGCTGCCAGCCGCTGGAGTTGACCGGGCTGGGCTTCGCGGAGCTGCAGGACTTGTGC

ArgGlnLeuHisAlaArgValAspLysValAspGluGluArgTyrAspIleGluAlaLys
CGACAGCTCCACGCCCGTGTGGACAAGGTGGATGAAGAGAGATACGACATAGAGGCAAAA

ValThrLysAsnIleThrGluIleAlaAspLeuThrGlnLysIlePheAspLeuArgGly
GTCACCAAGAACATCACGGAGATTGCAGATCTGACTCAGAAGATCTTTGACCTTCGAGGC

LysPheLysArgProThrLeuArgArgValArgIleSerAlaAspAlaMetMetGlnAla
AAGTTTAAGCGGCCCACCCTGCGGAGAGTGAGGATCTCTGCAGATGCCATGATGCAGGCG

LeuLeuGlyAlaArgAlaLysGluserLeuAspLeuArgAlaHisLeuLysGlnValLys
CTGCTGGGGGCCCGGGCTAAGGAGTCCCTGGACCTGCGGGCCCACCTCAAGCAGGTGAAG

LysGluAspThrGluLysGluAsnArgGluValGlyAspTrpArglysAsnIleAspala
AAGGAGGACACCGAGAAGGAAAACCGGGAGGTGGGAGACTGGCGGAAGAACATCGATGCA

LeuSerGlyMetGluGlyArgLysLysLysPheGluSer***
CTGAGTGGAATGGAGGGCCGCAAGAAAAAGTTTGAGAGCTGAGCCTTCCTGCCTACTGCC

CCTGCCCTGAGGAGGGCCACTGAGGAATAAAGCTTCTCTCTGAGCTGAAAAAAAAAAAAA

AAAAAAAAAAAAAAAAAAAA-3'
```

Fig. 1. Nucleotide sequence of clone pCTI-2 and deduced amino acid sequence of human cardiac troponin I. Underlined (from 5′ to 3′) are the translation initiation consensus sequence [14], the positions of oligo-sets 1, 2 and 3 (see text), and the poly-adenylation signal sequence [12].

Set 3:5′-ATGAATTCTT(T/C/G/A)C(T/G)(T/C/G/A)CC(C/T)-TCCAT-3′

2.2. *cDNA synthesis and polymerase chain reactions*

cDNA synthesis was carried out as follows: 50 pmol of oligo-set 3 was annealed with 1 μg of human ventricular muscle RNA in 20 μl of H_2O by heating to 65°C and cooling slowly to 20°C. Extension was carried out in 10 mM Tris-HCl (pH 8.4) 50 mM KCl, 1.5 mM $MgCl_2$, 100 μg/ml BSA, 150 mM each of dATP dCTP dGTP and dTTP, 10 mM DTT, 0.2 units/μl RNase inhibitor (Amersham) and 15 units AMV reverse transcriptase (Pharmacia) at 42°C for 1 h. Polymerase chain reactions were carried out in the same buffer by the addition of appropriate oligo-sets (50 pmol) and 1 unit Taq polymerase (United States Biochemical). 40 cycles were carried out for 1.5 min at 94°C, 2 min at 50°C and 2 min at 72°C, followed by a final 10 min extension at 72°C.

```
[H]  M A D G S S D A * A R E P R P A P A P I R R R S S * N Y R A Y A   30
[B]    - - R - G G S T - G D T V - - - P - V - - - - - A - - - - - - -
[R]    - - E - R - - * - G - A - - - - - - * V - - * * - * D * - - - -

[H]  T E P H A K K K S K I S A S R K L Q L K T L L L Q I A K Q E L E   62
[B]  - - - - - - - - - - - - - - - - - - - - - - - - M - - - - - - - -
[R]  - - - - - - - S - K - - - - - - - - - - - - - - M - - - - - - - -
[RS]     - E V E R - - - - - T - - - - - - * - - S - M - A K - - E C Q Q
[RF]   G D E E - R N R A - T - * - R Q H - - S V M - - - - A T - - -

[H]  R E A E E R R G E K G R A L S T R C Q P L E L T G L G F A E L Q   94
[B]  - - - - - - - - - - - - - - - - - - - - - - - - - - A - - - - - - -
[R]  - - - - - - - - - - - - - - - - - - - - - - - - - - A - - - - - - -
[RS] E H E A * - E A - - V - Y - A E - I P A - Q T R - - S L S A - -
[RF] K - E G R - E A - - Q N Y - A E H - P - - S - P - * S M - - V -

[H]  D L C R Q L H A R V D K V D E E R Y D I E A K V T K N I T E I A  126
[B]  - - - - - - - - - - - - - - - - - - - - V - - - - - - - - - - -
[R]  - - - - - - - - - - - - - - - - - - - - V - - - - - - - - - - -
[RS] - - - - - - - - K - E V - - - - - - - - - - C L H - T R - - K
[RF] E - - K - - - - K I - A A E - - K - - M - I - - Q - S S K - L E

[H]  D L T Q K I F D L R G K F K R P T L R R V R I S A D A M M Q A L  158
[B]  - - N - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
[R]  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
[RS] - - K L - V L - - - - - - - - - P - - - - - V - - - - - L R - -
[RF] - M N - - L - - - - - - - - - - P - - - - - M - - - - - L K - -

[H]  L G A R A K E S L D L R A H L K Q V K K E D T E K E N R * E V G  189
[B]  - - - - - - - - T - - - - - - - - - - - - - - - - - - - * - - -
[R]  - - T - - - - - T - - - - - - - - - - - - - - - - - - - * - - -
[RS] - - S K H - V - M - - - - N - - S - - - - - - - - - - R P V - - -
[RF] - - S K H - V C M - - - - N - - - - - - - - - - - - - R * * D - -

[H]  D W R K N I D A L S G M E G R K K K F E S                        210
[B]  - - - - - - - - - - - - - - - - - - - - - G
[R]  - - - - - - - - L - - - - - - - - - - - - G
[RS] - - - - - V E - M - - - - - - - - - M - D A A K S P T S Q
[RF] - - - - - - - E E K - - - - - - - - - M - - - E S
```

Fig. 2. Alignment of the human cardiac troponin I sequence with other published sequences [see 1]. Single letter code has been used; (*N*) = identical amino acid, (*) = no corresponding amino acid. Underlined are the amino acid sequences used in generating oligo-sets 1, 2 and 3 (see text). Numbering refers to the human sequence only.

2.3. *Isolation of full length cDNAs and sequence analysis*

cDNA clones encoding cardiac troponin I were isolated from a human ventricular cDNA library cloned in Lambda ZAPII (Stratagene Inc) using insert DNA from a PCR-derived clone (pCTI-1), and standard screening methods [8]. Isolated phages were converted to pBluescript plasmid by Lambda ZAPII excision in vivo. 7 independent clones were analysed for insert size and two were selected for sequence analysis. DNA sequencing was carried on both DNA strands by the dideoxy chain-termination method using additional oligonucleotide primers synthesised for internal sequences.

2.4. *RNA preparation and Northern blot analysis*

RNA was prepared from cardiac muscle obtained from therapeutic abortions, autopsy samples and transplant material, using isothiocyanate and CsCl centrifugation [9]. Northern blots were prepared using 10 μg of total RNA run on 1% agarose-formaldehyde gels [8] and transferred in alkali to Hybond N+ membranes (Amersham UK). ^{32}P labelled probes were prepared by random priming [8] of insert DNA from clone pCTI-3.

3. RESULTS AND DISCUSSION

PCR amplification of single stranded cDNA (prepared from human ventricular muscle RNA) with primers 1 and 3 resulted in a single 225bp DNA fragment - the size predicted for TnIc (data not shown). The identity of this fragment was further confirmed by re-amplification of the DNA fragment using primers 1 and 2. This resulted in the expected size change from 225 bp to 164 bp (data not shown). This fragment was subcloned into pBluescript (Stratagene Inc) after digestion with *Eco*RI and *Pst*I, and the sequence of the resulting clone (pCTI-1) confirmed its identity with TnIc. Due to an internal *Pst*1 site in the PCR fragment, clone pCTI-1 in fact contains a 113 bp insert including sequence corresponding to amino acids 151–184.

The use of the polymerase chain reaction to clone cDNAs from poorly defined amino-acid sequence data is clearly of major importance [10]. We have taken the approach of using fully degenerate oligonucleotide sets (e.g. [11]) covering all possible coding sequences corresponding to conserved protein domains. One potential disadvantage of this method is that it increases the chances of amplifying unrelated sequences. In our experience this is overruled by the certainty of including the exact oligonucleotides required for successful amplification. In addition, the use of internal primers represents a rapid and convenient approach to confirming the identity of amplified products.

Screening of a human cardiac cDNA library with the DNA insert from pCTI-1 identified 7 independent hybridizing clones, two of which were fully sequenced. Fig. 1 shows the sequence of clone pCTI-2 which contains the entire coding sequence of human TnIc. The identity of this clone is confirmed both by comparison with amino acid sequences of other troponin I isoforms (Fig. 2) and by Nothern blot hybridization (Fig. 3). This clone contains the poly-adenylation signal (AATAAA [12]) and a poly-A tail of 33 nucleotides indicating that it contains the entire 3′ non-coding sequence. Clone pCTI-3 was found to contain sequences from amino

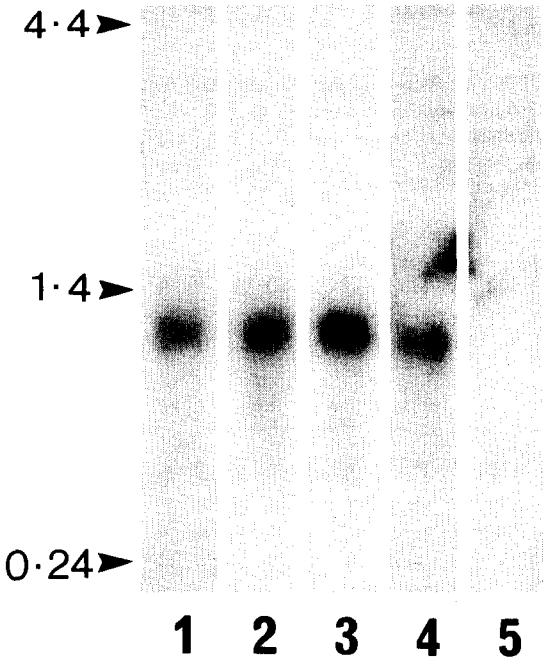

Fig. 3. Northern blot hybridization of ventricular muscle RNA from 20 week fetal heart (track 1), 28 week fetal heart (track 2), 9 month postnatal heart (track 3), adult ventricular muscle (track 4), and adult skeletal (quadriceps) muscle (track 5). Following hybridization, filters were washed in 0.1 × SSC, 0.1% SDS at 65°C. The positions of RNA size markers are indicated in thousands of nucleotides.

acid 77 to the end of the 3′ non-coding sequence but not including any poly-A tail. The amino acid sequence derived from these clones is shown in Fig. 1. The three published cardiac sequences were aligned as shown in Fig. 2, where they are contrasted to the fast and slow skeletal isoforms of rabbit [4]. Alignment of the cardiac sequences shows that they are highly conserved throughout their length with the exception of the N-terminal sequence. This is of particular interest considering that the N-terminal sequence is involved in the phosphorylation-mediated response to β-adrenergic agonists [5]. Manipulation of the troponin I sequence by the use of site directed mutagenesis should allow a more detailed analysis of the functional importance of this and other domains.

During development of the rat and chick heart, there is a transition from expression of TnIs to that of TnIc late in fetal life [6,7]. We have used our cloned cDNA to examine the expression of TnIc in the developing human heart. Fig. 3 shows that the adult cardiac isoform already represents a major transcript by 20 weeks gestation. No cross-hybridisation with skeletal muscle troponin I mRNA, present in RNA from quadriceps, is detected in this experiment (track 5, Fig. 3) indicating the specificity of the probe used. This result is consistent with previous reports using antibodies which detected TnIc as the major isoform at 20 weeks [14], but contrasts with the rat and chick where

transition from TnIs to TnIc occurs around the time of birth. Our data imply that if there is a transition TnI expression in the human myocardium which is comparable to that seen in rat and chick, it must occur before 20 weeks. Further developmental studies using cloned probes should allow us to determine the exact timing of activation of this gene.

Acknowledgements: This work was supported by the British Heart Foundation. We thank S. Cotterill and C. Lichtenstein for help in oligonucleotide synthesis.

REFERENCES

[1] Leszyk, J., Dumaswala, R., Potter, J.D. and Collins, J.H. (1988) Biochemistry 27, 2821–2827.
[2] Koppe, R.I., Hallauer, P.L., Karpati, G. and Hastings, K.E.M. (1989) J. Biol. Chem. 264, 14327–14333.
[3] Baldwin, A.S., Kittler, E.L.W. and Emerson, C.P. (1985) Proc. Natl. Acad. Sci. USA 82, 8080–8084.
[4] Wilkinson, J.M. and Grand, R.J.A. (1978) Nature (Lond.) 271, 31–35.
[5] Solaro, R.J., Moir, A.J.G. and Perry, S.V. (1976) Nature (Lond.) 262, 615–617.
[6] Saggin, L., Gorza, L., Ausoni, S. and Schiaffino, S. (1989) J. Biol. Chem. 264, 16299–16302.
[7] Sabry, M.A. and Dhoot, G.K. (1989) J. Mucle. Res. Cell Motil. 10, 85–91.
[8] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning - A laboratory Manual, 2nd edn, Cold Spring Harbour Lab. Press, New York.
[9] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[10] Erlich, H.A. (ed) (1989) PCR Technology, Principals and Applications for DNA Amplification, Stockton, New York.
[11] Mack, D.H. and Sninsky, J.J. (1988) Proc. Natl. Acad. Sci. USA 85, 6977–6981.
[12] Proudfoot, N.J. and Brownlee, G.G. (1976) Nature (Lond.) 263, 211–214.
[13] Humphreys, J.E. and Cummins, P.J. (1984) J. Mol. Cell. Cardiol. 16, 643–657.
[14] Kozak, M. (1989) J Cell Biol. 108, 229–241.