Vol. 183, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 28, 1992 Pages 176-183

ISOLATION AND CHARACTERIZATION OF A RAT VENTRICULAR cDNA EXPRESSED SPECIFICALLY IN CARDIAC AND SKELETAL MUSCLES[1]

Madhu Gupta[*α], Steven P. Smeekens[+], Mahesh P. Gupta[*β], and Radovan Zak[*#@]

[*]Departments of Medicine, [#]Organismal Biology and Anatomy, [@]Pharmacological & Physiological Sciences, and [+]Howard Hughes Medical Institute, The University of Chicago, Chicago, Illinois 60637

Received January 13, 1992

SUMMARY: We describe the isolation of a novel cDNA named Myomy and show that its transcripts are present in skeletal and cardiac muscles as well as in differentiated Sol 8 skeletal muscle cell line. Sequence analysis revealed that neither nucleotides nor deduced protein product have any significant homology to those previously described. The encoded protein of Myomy cDNA consists of 76 amino acids and has a molecular weight of 8,000 dalton. Based on its muscle specific expression, low abundance and a higher occurrence of SP(T)XX, S(T)S(T)XX motifs, we suggest that Myomy encodes a new muscle specific transcription factor.

The regulation of gene transcription involves specific interactions between regulatory nuclear proteins and target gene sequence elements. The expression or activity of these sequence specific transcription factors may be regulated in a cell type, tissue specific or cell cycle dependent manner (1-4). In recent years several eukaryotic DNA binding transcription factors have been purified. They possess well conserved specific DNA binding domain and play an important role during development. Foremost among these are the homeodomain and paired box gene families that encode Hox, Pou, and Pax proteins suspected of specifying

[1]This work was supported by the Natl. Inst. of Health grant HL20592.

[α]Recipient of Research Fellowship of American Heart Association (Metropolitan Chicago).

[β]Senior Research Fellow of American Heart Association (Metropolitan Chicago).

[#]To whom all correspondence should be addressed.

0006-291X/92 $1.50

positional regulatory information during embryogenesis (5-8) as well as helix loop helix (HLH) genes. The HLH genes share a common sequence motif that encodes a basic region followed by helix loop helix, the structures essential for DNA binding and dimerisation. Some of these factors (myoD, myogenin, MRF4) are found exclusively in skeletal muscle cells and their transfection into non muscle cells leads to the expression of various muscle specific markers (9-11). While we are beginning to understand the role of these factors in skeletal muscle myogenesis, little is known about the regulatory aspect of cardiac myogenesis. In an attempt to search for a cardiac analog of MyoD class of factors, we have isolated a novel cDNA which shows muscle restricted expression.

MATERIALS AND METHODS

Materials

Oligonucleotides were synthesized on an Applied Biosystems model 380 B DNA synthesizer. Taq DNA Polymerase was obtained from Perkin-Elmer Cetus Corp., Sequanase from United States Biochemicals, other DNA modifying enzymes, and restriction endonucleases were obtained from Bethesda Research Laboratories. Radiolabelled nucleotides were obtained from Amersham Corporation.

Amplification of DNA fragments

cDNA templates for polymerase chain reaction (PCR) were synthesized using 5 μg of rat skeletal and cardiac muscle total RNA and Moloney murine leukemia virus reverse transcriptase as described by supplier (BRL),and 1/10th of the cDNA reaction mixture was used for each PCR. Reaction mixture for PCR (0.1 ml) contained cDNA template, 100 pmol each of the degenerate primers myo(+) and myo(-) (Fig.1), and 2.5 units of Taq DNA polymerase in 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$ containing dNTPs at 0.2 mM each and 0.01% gelatin. Reactions were carried out in a Perkin-Elmer Cetus thermal cycler for 30 cycles of denaturation (94°C, 1 min), annealing (55°C, 1 min), and extention (72°C, 2 min). After fractionation on 6% polyacrylamide gels the 150 bp PCR product was eluted, purified, blunt ended with T4 DNA polymerase and ligated into the EcoRV site of pBluescript (Stratagene).

Screening of cDNA library

A rat (7 days old neonatal) heart cDNA library in a lambda gt10 vector (kindly provided by Dr.R. Rogart,Univ. of Chicago) was screened (12,13) with a 150 bp PCR product of rat heart cDNA. Approximately $1x10^6$ recombinant phage were screened with the cDNA probe. Briefly, the phage were plated onto twenty five 15 cm L-broth agar plates with a lawn of E. Coli (BB4, Stratagene) in 0.7% top agarose, grown overnight at 37°C, and then the nylon filters were overlaid (Amersham Hybond). Duplicate filters were made from each plate. The filters were immersed immediately in 0.5 M NaOH, 1.5 M NaCl for 10 min and then neutralized in 0.5 M Tris(pH 7.4), 1.5 M NaCl for an additional 10 min. After air drying, filters were baked at 80°C for 2 hrs in a vacuum oven and hybridized [2xSSC(1xSSC is 150 mM NaCl/15 mM sodium citrate, pH 7.0), 50% formamide, 45°C for 18 hrs] with random primed cDNA probe and washed(2xSSC, 0.1% sodium dodecyl sulphate, 65°C for 1 hr). After

secondary and tertiary screening, the phage inserts were subcloned into the EcoR1 site of pBluescript and sequenced in both directions by dideoxy chain termination method, using the Sequanase kit (United States Biochemicals).

Northern Blot Analysis

Total RNA was prepared from tissue culture cell lines or adult rat tissue using the guanidium isothiocyanate/CsCl procedure (14). Poly(A^+) RNA was selected from total RNA by using a commercial kit (5 Prime-3 Prime, Inc.). Thirty micrograms of total RNA or 5 μg of Poly (A^+) RNA was dissolved in RNA loading buffer [20 mM MOPS(pH 7.8), 5 mM Sodium acetate, 0.1 mM EDTA, 6.7% formaldehyde, 50% deionized formamide]. The samples were heated to 65°C for 15 min, chilled on ice and electrophoresed for 4 hr at 4 V/Cm in a 1.5% agarose gel containing 6.7% formaldehyde. The gel was rinsed with water, treated with 50 mM NaOH for 20 min, neutralized in 20x SSC for 10 min before being transferred to a nylon membrane (GeneScreen, New England Nuclear) for 16 hrs. The filter was baked for 2 hrs at 80°C in a vacuum oven and hybridized [2x SSPE (1x SSPE is 180 mM NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA), 50% formamide, 45°C, 18 hrs], washed (0.5x SSPE/0.1% sodium dodecyl sulphate, 60°C, 1 hr) and autoradiographed with intensifying screens. The probe used in this assay was the BamH1-EcoR1 fragment of Myomy cDNA that had been purified and random primed.

In Vitro Transcription and Translation

The plasmid used to make RNA contained the entire cDNA insert cloned into EcoR1 site of pBluescript vector. It was linearised with Sma1 and used in preprative T7 RNA polymerase reaction. Generally, 10-20 μg of RNA was synthesized in a 50 μl reaction using 100 U of T7 RNA polymerase. In vitro translation was performed using approximately 0.5 μg of RNA in a 50 μl reaction containing 35 μl of rabbit reticulocyte lysate (Promega). Radioactively labelled protein was generated by incorporating L-[^{35}S]cysteine (>1000 Ci/mmole; ICN). Reaction was performed at 30°C for 90 min and the products were run on a 10-20% SDS polyacrylamide gradient gel. Parallel reaction with no added RNA served as a control.

RESULTS AND DISCUSSION

To identify cardiac analog of MyoD family of transcription factors we designed degenerate PCR primers that were complementary to the stretches of amino acid sequences highly conserved among MyoD, MRF4, myogenin and Myf-5. Alignment of these proteins indicated that sequences located amino terminally to helix I and helix II contained the most useful sequences for this purpose (Fig.1). After PCR amplification of cDNA prepared from rat cardiac and skeletal muscle total RNA, the expected 150 base pair (bp) fragment was subcloned into pBluescript and individual plasmid DNA were sequenced. Of 6 DNAs analyzed, 4 showed similar sequences and data bank search revealed them to be distinct from those previously characterized cDNAs. This newly identified cDNA clone was designated as Myomy.

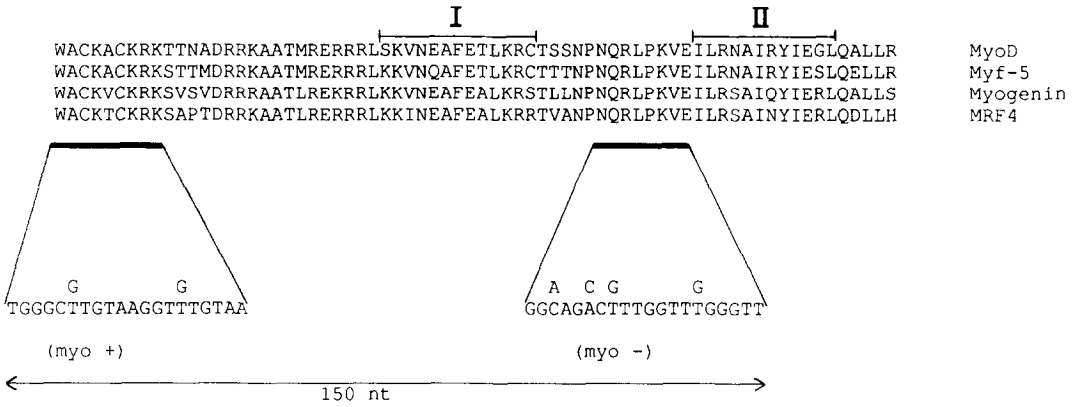

Fig.1. Schematic representation of the cloning strategy of Myomy. Amino acid similarities in MyoD, Myogenin, Myf-5, and MRF-4 are aligned. Top lines (I and II) mark the amphipathic helices. The sequences of degenerate forward (myo+) and reverse (myo-) primers used in polymerase chain reaction (PCR) are shown. Bold lines represent the site of selection of these primers. The length of the expected PCR product is indicated.

In order to isolate full length cDNA clone, a cDNA library prepared from poly(A^+) RNA of 7 day neonatal rat heart was screened using 150 bp fragment (generated by PCR) as a probe. Upon screening $1X10^6$ plaques, six clones were identified and both strands of one of these clones which had an insert of 1.049 kilobase (kb) pairs was sequenced. The complete sequence of Myomy cDNA is shown in Fig.2. Sequence analysis showed that the probe (generated by PCR) used to screen the library was located in the 5' noncoding region of the cDNA clone. The clone contains a poly A tail, a polyadenylation signal (AATAAA) 12 bases upstream from the start of poly A tail and an open reading frame (ORF) of 249 bp which shows the existence of two AUG (502) and (526) initiator codons. Only the 2nd (526) codon is in good context with Kozak consensus sequence (15) and is predicted to encode a protein of 76 amino acids. Following transcription / translation in vitro the synthesized protein from Myomy cDNA migrated as an approximately 8 kd polypeptide (n=2) which is the expected size from the predicted ORF (Fig. 3B). Comparison of these sequences with most recently available protein and DNA sequence databases (Biovax; TFastA and FastA) failed to identify related gene products of significant homology.

To determine the tissue distribution of this clone Northern blot analysis was carried out on total RNA and poly(A+) RNA isolated from a variety of tissues and cultured cell lines. 30

```
CCCACCCACGCCACAGCAGAGGGACTGCACAGGAGGTCAGCGTGGACTCCCAGTTTGCTT   60
CTCAATGTGCAGCATTCTTACACTGGGTTTGTAAGACGTCCTTTCAGGAAGAGATGATGA  120
TGATAAGAGCACTGGTTGCTCTTCTGAGGACCTGGGCTCAATTTCCAGCGCCCCACGTGG  180
CAGTTCACGCCTGTCTGTAACCCCAGCTCCAGGGCTTCGGCACCCTCACACAGACATCC  240
ATGTGGATAAAACACCTATGAACATAAGCTAAAAAATAAAATAAAAAAGCCAGCGATGTC  300
ACATAGGATATCCCCAATGACATGGTTCAGCTGAGCACCGTGGGAAGCGGTGAGATTCAG  360
CCAAGTCTGAGCCCTGGCTTCACATTTTAAAAACTCAGGGCAGGCGGTAACCCTAACGCG  420
CTTGTGGCAGGGGCCGAGCATCTGCCGCTGTGTGAGTGTTAAGCAGAGGACTGTGCAACA  480
CTCCAGACAGCAGGATCCTCCATGCCGCGTGTCAGACCTAAAGAGATGGGAAACAAACGC  540
                      M  P  R  V  R  P  K  E  M  G  N  K  R
CCGGACCATTTAGTTGGGTGTGGCAGTGAATGCTTGCGGGAGCCTACAGCAAGAGCTACA  600
 P  D  H  L  V  G  C  G  S  E  C  L  R  E  P  T  A  R  A  T
ATTCAGGATCAGGCTGGGCTGTCTAGTGACACTGCCTCGCCTGCCCCACACTGCTGGTGC  660
 I  Q  D  Q  A  G  L  S  S  D  T  A  S  P  A  P  H  C  W  C
TCTCCAGCATTTAGCCACTCTAGACATGGTCACCAGTGTCTGCACATGTCGTTTAGAAGA  720
 S  P  A  F  S  H  S  R  H  G  H  Q  C  L  H  M  S  F  R  R
GTGGATCTGCTGAGGTGCAGAAACTTGGGTTGATTTAGAGACTCCAGGACTAAAAGTAAA  780
 V  D  L  L  R  C  R  N  L  G
GAGGTAAAACAGCCATGTTGCAAGTGGAACGAGAGGTAAGAAAGCCAGGGGTGGGGATGC  840
CTTGCCCACCCTGTGCTGTGCGCCAGGACGGTGTCTCCATGGCAGTTGGTAGCACCTGGC  900
AACAGGAGAGGGTGGCCATTGTGGGGAAAACAACCAACCTTGGCCGTGTATGTGGAGGTT  960
GTGGCTACTGTGTGTTCCTAGAACTGAACAGCCTGATCTACGTACATCCCCAAATAAAAA 1020
CCGCCAAGGGAAAAAAAAAAAAAAAAAAA
```

Fig.2. Nucleotide and predicted amino acid sequence of Myomy. The nucleotide number is indicated at the end of each line. The initiation codon which is in good context with Kozak consensus sequence and the stop codon are highlighted in bold letters. The signal for polyadenylation is underlined.

μg of total RNA did not show any signal in any of the tissues studied (data not shown). 5 μg of poly(A^+) RNA showed a strong signal in skeletal and cardiac muscles and did not hybridize to mRNA of smooth muscle, liver, brain and kidney (fig. 4). These data indicate the muscle specificity and the low abundance of this transcript. Interestingly, skeletal muscle showed an additional signal at 2.3 kb position, may be due to differential splicing of Myomy in this particular tissue. In order to determine whether this muscle specific cDNA is expressed only in differentiated muscle or is also expressed in myoblast, total RNA from Sol 8 myoblast (4 days in culture) and myotubes (48 hrs in differentiation media) was probed with the same cDNA. We have found the signal to be present in myotubes only (data not shown).

Differentiation of muscle cell lineage is a complex phenomenon involving withdrawal of myoblast from cell cycle and expression of a battery of contractile proteins. Although cloning of MyoD cDNA was the first necessary step in understanding the

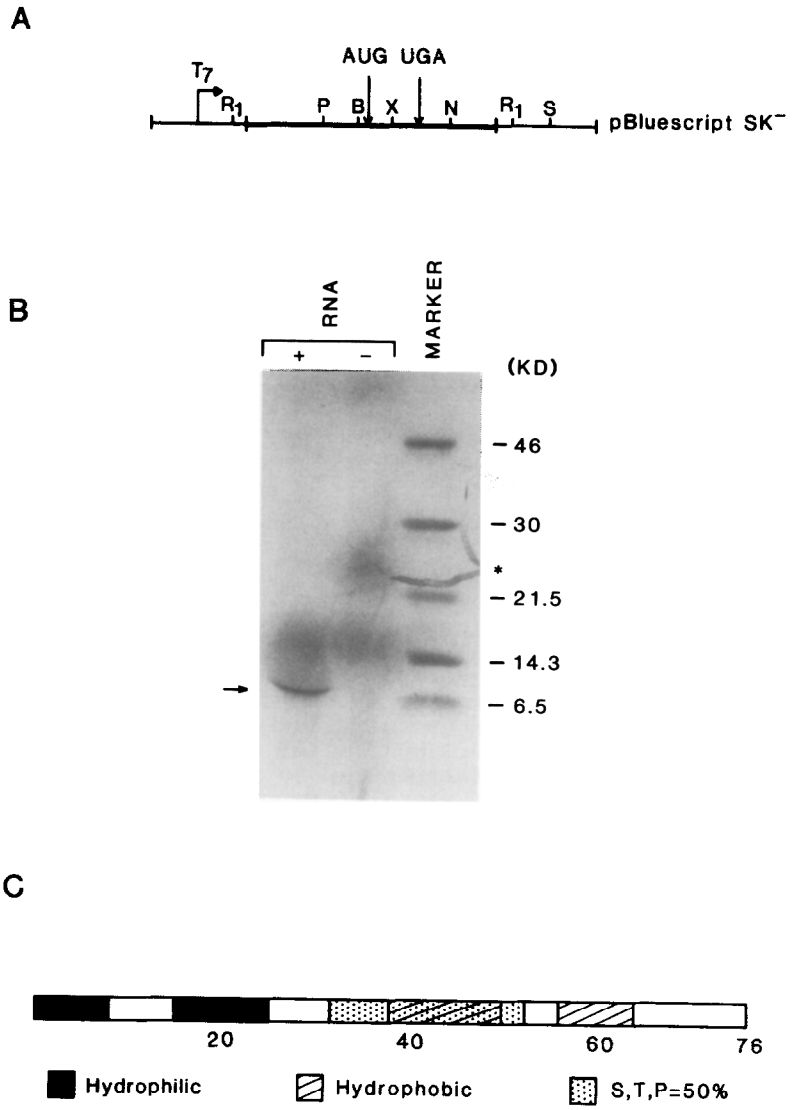

Fig.3. In vitro translation of the Myomy cDNA.
(A) The structure of the cDNA-containing plasmid pBluescript used for in vitro transcription by T7 polymerase is shown schematically. P=PvuII, B=Bam H1, N=Nco1, and R1=EcoR1. The solid line flanked by two vertical bars (indicating the R1 site) represent the cDNA insert. The position of T7 promoter, the potential initiator (AUG) and stop (UGA) codons are indicated. (B)In vitro transcription/ translation of Myomy ORF. Myomy RNA was synthesized using T7 polymerase. Translation reactions were performed in vitro with rabbit reticulocyte lysate in presence(+) and absence(-) of synthesized RNA. 4 μl of the translation products were analyzed on a denaturing SDS-polyacrylamide gel.The arrow indicates the position of the major 8 Kd product, asterics indicates an artifact from gel drying. (C) Some primary structural features of Myomy protein. The amino acid sequences are schematically represented. Dot filled box refers to the region containing a preponderance of amino acids indicated at the bottom using single letter code.

regulation of myogenesis, the diversity of muscle phenotypes can not be explained by the existence of MyoD and related members of transcription factors. For example, several subtypes of fast and

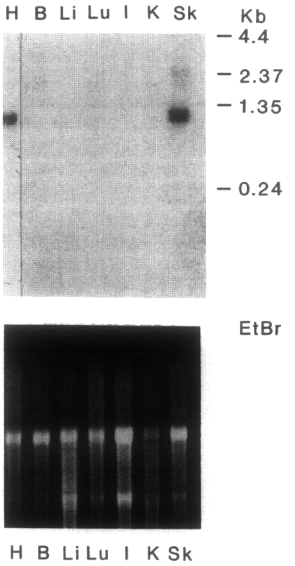

Fig.4. Tissue distribution study of Myomy expression. Northern blot analysis was performed on RNA obtained from heart(H),brain(B), liver(L),lung(Lu), small intestine(I), kidney(K) and skeletal muscle(Sk). Five microgram of Poly A+ RNA was fractionated on formaldehyde-agarose gel and probed as described in Materials and Methods, film was exposed to the blot for 48 hr. The signal shown in heart required 96 hr of exposure, no signal was observed in other tissues even after 2 weeks of exposure (data not shown). Bottom panel shows the Ethidium bromide staining of the gel. The size of RNA in kilobases(kb) is indicated.

slow contracting fibers are present in fetal and adult skeletal muscles and this diversity in some instances is thought to be a product of developmental history (16) implying the existence of a correspondingly diverse and subtle system of muscle gene regulation. On the other hand, there are contractile proteins which are expressed both in cardiac and at very early stage of skeletal muscle development (primordial phenotype) (17), indicating the presence of an as yet unidentified pathway of muscle gene regulation that is common to both types of muscle. Since Myomy is expressed in cardiac as well as in skeletal muscle cells, it is tempting to speculate that this particular factor may be a part of the pathway that regulates the expression of the primordial phenotype.

Analysis of the amino acid sequences of this clone revealed a higher occurrence of Ser, Pro, Thr amino acids (Fig.3C). Since Ser-Pro-x-x and Thr-Pro-x-x motifs are found more commonly in genes specifying regulatory than in structural proteins, we calculated the frequency of occurrence of SP(T)XX, S(T)S(T)XX sequences based

on the parameters described by Suzuki 1989 (18). Our calculated value ($6.82x10^{-3}$) for the occurrence of these motifs is well above that described for general proteins (2.89 to $3.76x10^{-3}$) suggesting that this newly cloned cDNA may fall within the category of genes encoding regulatory proteins. Direct demonstration of Myomy as a transcription factor awaits functional analysis including DNA binding experiments with muscle specific promoters. In addition we are attempting to raise antibodies against Myomy protein which will help us in carrying out some of these studies and in identifying its subcellular localization.

REFERENCES

1. Fletcher,C., Heintz,N., and Roeder,R.G. (1987) Cell 51, 773-781.
2. Scheidereit,C., Heguy,A. and Roeder,R.G. (1987) Cell 51, 783-793.
3. Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) Science 236, 1237-1244.
4. Monaci,P., Nicocia,A. and Cortese,R. (1988) EMBO J. 7, 2075-2087.
5. Affolter,M., Shier,A. and Gehring,W.J. (1990) Curr. Opin. Cell Biol. 2, 485-495.
6. He,X., Treacy,M.N., Simmons,D.M., Ingraham,H.A., Swanson,L.W. and Rosenfeld,M.G. (1989) Nature 340, 35-42.
7. Akam,M. (1989) Cell 57, 347-349.
8. Kessel,M. and Gruss,P. (1990) Science 249, 374-379.
9. Tapscott,S.J., Davies,R.L., Thayer,M.J., Chang,P., Weintraub,H. and Lassar,A.B. (1988) Science 242, 405-411.
10. Murre,C., McCaw,P.S., Vaessin,H., Caudy,M., Jan,L.Y., Jan,Y.N., Cabrare,C.V., Buskin,J.N., Hauschka,S.D., Lassar,A.B., Weintraub,H. and Baltimore,D. (1989) Cell 58, 537-544.
11. Davis,R,L., Weintraub,H. and Lassar,A.B. (1987) Cell 51, 987-1000.
12. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular cloning, A Laboratory Mannual. Cold Spring Harbor Lab. Cold Spring Harbor, NY.
13. Ausubel,F.M., Brent,R., Kinston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., and Struhl,K. (1989) Current Protocols in Molecular Biology, Vol. 1, Greene Publishing Associtates and Wiley-Interscience, NY.
14. Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J., and Rutter,W.J. (1979) Biochemistry 18, 5294-5299.
15. Kozak,M. (1989) The scanning model for translation: an update. J. Cell. Biol. 108, 229-241.
16. Miller,J.B., and Stockdale F.E. (1986) Proc. Natl. Acad. Sci. USA 83, 3860-3864.
17. Sweeney,L.J., Kennedy,J.M., Zak,R.Z., Kokjohn,K., and Kelly,S.W.(1988) Dev. Biol. 133, 361-374.
18. Suzuki,M. (1989) J. Mol. Biol. 207, 61-84.