Identification of a muscle factor related to Myod in a fish species

Pierre-Yves Rescan, Laurent Gauvry, Gilles Paboeuf, Benoit Fauconneau

ELSEVIER

Biochimica et Biophysica Acta 1218 (1994) 202–204

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

Identification of a muscle factor related to MyoD in a fish species

Pierre-Yves Rescan *, Laurent Gauvry, Gilles Paboeuf, Benoit Fauconneau

Laboratoire de Physiologie des Poissons, INRA, Campus de Beaulieu, 35042 Rennes, France

(Received 26 November 1993)

Abstract

We have isolated the cDNA encoding a myogenic factor expressed in embryonic trout muscle by hybridization with a *Xenopus* MyoD cDNA. Nucleotide sequence analysis and amino acid comparison showed that this cDNA called TMyoD encodes a polypeptide of 276 amino acids with 70% identity to the entire *Xenopus* MyoD protein and 92% identity within the basic and myc-like region. Results from Northern blotting showed that the corresponding transcript is expressed both in adult and embryonic skeletal musculature and in an in vitro myogenesis system, but is undetectable in cardiac and smooth muscles and in non muscle tissues.

Key words: MyoD; Myogenesis; Teleost; (Satellite cell)

The MyoD gene identified by substract cloning for myoblast specific RNA is the prototype of a family of master regulators of skeletal myogenesis which includes in vertebrates three other members namely myogenin, Myf5 and MRF4 identified subsequently [1–5]. All these genes encode proteins that share a highly conserved central region termed the basic/helix-loop-helix(B-HLH) domain related to the c-myc superfamily and contain sequences essential for both dimerisation and DNA binding [6]. In multipotential 10T1/2 cells, transfection experiments have shown that forced expression of these exogenous myogenic factors is sufficient to drive them down the muscle differentiation pathway suggesting their functions in myogenic lineage determination [1–5]. In contrast to vertebrates whose genome encodes multiple members of the MyoD family, invertebrates, including Sea urchin [7], *C. elegans* [8] and *Drosophila* [9], appear to contain only a single myogenic factor encoding gene. However, the myogenic factor for Sea urchin and *C. elegans* activates myogenesis in 10T1/2 cells indicating a highly conserved mechanism for muscle genes activation.

Myogenic factors have been studied in mammals, amphibians and birds [10], but nothing is known to date in fish. To analyse early developmental events leading to muscle formation in fish, we set out to isolate myogenic regulatory factors from Rainbow trout (*Oncorhynchus mykiss*) embryos (398 degree days) using a probe which spanned functional domains of the *Xenopus* MyoD cDNA. For this purpose, a λgt10 cDNA library was constructed from poly(A)$^+$ RNA of the trunk of rainbow trout embryos. The double strands cDNAs synthetized by the method of Gubler et al. [11] were size fractionated by gel filtration on a sepharose 4B column (Pharmacia) and the largest fractions were pooled, inserted into λgt10 vector (Stratagene) and encapsided using an in vitro packaging kit (Amersham). After amplification of the cDNA library, approximately $5 \cdot 10^5$ plaques were screened at low stringency with a fragment from the *Xenopus* MyoD cDNA encompassing the B-HLH domain [12]. From 9 positive clones, we identified a single cDNA of 1.5 kb which had two *Eco*RI fragments of aproximately 1.2 and 0.3 kb. Restriction analysis and sequencing showed that this internal *Eco*RI site was not situated near a *Not*1 site which is contained in the linker used for the ligation of the cDNAs in the λgt10 vector, so we did not think that the two fragments were inadvertently ligated during the construction of the library. The nucleotide sequence of our cDNA, determined by standard

* Corresponding author. Fax: +33 99 285020. E-mail: bf@beaulieu.rennes.inra.fr.

The sequence data reported in this paper have been submitted to the EMBL/Genbank/DDBJ Nucleotide sequence Databases under the accession number X75798.

SSDI 0167-4781(94)00003-L

```
   -30          -20          -10                          15
G  ATTTTGGCAG   AACAGAACAG   AAGAGTCACC   ATG GAG TTG CCG GAT ATT CCT TTC
                                          Met Glu Leu Pro Asp Ile Pro Phe

     30                   45                   60                   75
CCT ATA ACC TCT CCA GAT GAC TTC TAC GAC GAC CCT TGC TTC AAC ACC AGC GAC
Pro Ile Thr Ser Pro Asp Asp Phe Tyr Asp Asp Pro Cys Phe Asn Thr Ser Asp

             90                  105                  120
ATG CAT TTC TTT GAG GAC CTG GAC CCG AGA CTC GTT CAT GTG GGT CTC CTC AAG
Met His Phe Phe Glu Asp Leu Asp Pro Arg Leu Val His Val Gly Leu Leu Lys

135                  150                  165                  180
CCG GAC GAC CAC CAT CAC AAA GAG GAC GAG CAC ATC CGG GCA CCG AGT GGG CAC
Pro Asp Asp His His His Lys Glu Asp Glu His Ile Arg Ala Pro Ser Gly His

       195                  210                  225                  240
CAC CAG GCT GGC AGG TGC CTC CTG TGG GCC TGC AAA GCC TGC AAG AGG AAG ACC
His Gln Ala Gly Arg Cys Leu Leu Trp Ala Cys Lys Ala Cys Lys Arg Lys Thr

                 255                  270                  285
ACC AAT GCT GAT CGC AGG AAA GCG GCT ACC ATG CGG GAA AGA AGG CGA CTG AGC
Thr Asn Ala Asp Arg Arg Lys Ala Ala Thr Met Arg Glu Arg Arg Arg Leu Ser

    300                  315                  330                  345
AAG GTG AAC GAC GCC TTC GAG ACA CTG AAG AGA TGT ACG TCT ACT AAC CCT AAC
Lys Val Asn Asp Ala Phe Glu Thr Leu Lys Arg Cys Thr Ser Thr Asn Pro Asn

             360                  375                  390
CAG AGG CTG CCC AAA GTG GAT ATC CTG CGG AAT GCC ATC AGC TAC ATT GAG TCT
Gln Arg Leu Pro Lys Val Asp Ile Leu Arg Asn Ala Ile Ser Tyr Ile Glu Ser

405                  420                  435                  450
CTC CAA GGC CTG CTT CGT GGG GCC GGA CAG GAG GGC AAC TAT TAC CCG GTG ATG
Leu Gln Gly Leu Leu Arg Gly Ala Gly Gln Glu Gly Asn Tyr Tyr Pro Val Met

       465                  480                  495                  510
GAT CAC TAT AGC GGG GAC TCG GAT GCG TCC AGT CCC CGC TCC AAC TGC TCA GAC
Asp His Tyr Ser Gly Asp Ser Asp Ala Ser Ser Pro Arg Ser Asn Cys Ser Asp

                 525                  540                  555
GGC ATG ATG GAT TTC AAT GGT CAG TCT TGT CCA CCA AGA CGG AGA AAC AAG TAT
Gly Met Met Asp Phe Asn Gly Gln Ser Cys Pro Pro Arg Arg Arg Asn Lys Tyr

    570                  585                  600                  615
GAT AGC ACC TAC TTC AAC GAA GCA CCA AAT GAT TCC AGA CAC AAG AAG AAC TCT
Asp Ser Thr Tyr Phe Asn Glu Ala Pro Asn Asp Ser Arg His Lys Lys Asn Ser

             630                  645                  660
GTT ATT TCC AGT TTG GAC TGC CTG TCA AAC ATC GTG GAG CGA ATC ACC ACG GAT
Val Ile Ser Ser Leu Asp Cys Leu Ser Asn Ile Val Glu Arg Ile Thr Thr Asp

675                  690                  705                  720
ACC TCT GCC TGT CCC GCT GTT CAG GAC GGT TCC GAG GGT AGC AGC CCC TGT TCT
Thr Ser Ala Cys Pro Ala Val Gln Asp Gly Ser Glu Gly Ser Ser Pro Cys Ser

       735                  750                  765                  780
CCC GGG GAT GGT TCC ATA GCG AGT GAG AAC GGA GCC CCC ATC CCG TCC CCG ATC
Pro Gly Asp Gly Ser Ile Ala Ser Glu Asn Gly Ala Pro Ile Pro Ser Pro Ile

                 795                  810                  825
AAC TGC GTC CCC GCC TTA CAT GAC CCA AAC ACC ATC TAC CAG GTG TTG TGA
Asn Cys Val Pro Ala Leu His Asp Pro Asn Thr Ile Tyr Gln Val Leu  *

     840          850          860          870          880          890
AGTCGGGTC   GGTGGACTGC   ATACAGTAAT   TGTACATTCT   TCAAAACACA   ACTTATTCCT

      900          910          920          930          940          950
TATGGGGAGA   GAACATGCCA   AAGACTTGCC   TAAGGTCGCT   ACAAGGCTAC   ACACCAAAGA

      960          970          980          990        1,000        1,010
AGATCCGATA   CCGGCTTTGA   AAGACATTAA   AAAATGACGG   TGTCCAATTT   CTTAAAGAAC

    1,020        1,030        1,040        1,050        1,060        1,070
CCTTGGTCTA   CATTGGATAA   TGTGTGTGTG   TGTGTGTGTG   TGTGAATGTA   TGTGTGTGAA

    1,080        1,090        1,100        1,110        1,120        1,130
TGTATTTATA   TTGTAGTAAG   CCTATGCTAT   TCTAAGATAG   TACAAGTCTG   AATTCATATA

    1,140        1,150        1,160        1,170        1,180        1,190
AACGGATACC   ATTCTATTTG   TGATCGACAT   AATTTAATTC   AATGGATTAT   CTGTAATATG

    1,200        1,210        1,220        1,230        1,240        1,250
AACATTTCCA   GTTGGCAAGG   CGGAAACGCC   GGAGATAAAT   GAATCCGGAA   AGTGAGGACC

    1,260        1,270        1,280        1,290        1,300        1,310
ATTTTCTATA   TGTGTAAATA   AGAGCTGCTT   TGCAAAATAA   AGAAGAAGAA   GAAAAAAACG

    1,320        1,330        1,340        1,350        1,360        1,370
AACAACACAC   AGGAAGTGTT   TGTAATCATA   TTTAATGTTG   CTCTTGGATT   GTTTGTGTTA

    1,380        1,390        1,400        1,410        1,420        1,430
GATTTTTAAC   TTTATATTTA   TAATATCAAG   AACGGAGTGA   ATTACATTTT   AATAAATGTA

  1,438
TATTATAT
```

Fig. 1. Nucleotide and deduced amino-acid sequence of TMyoD. A *Not*1 restriction fragment containing the entire cDNA was subcloned into plasmid Bluescript and subjected to deoxysequencing with vector primers and additional primers specific to trout MyoD cDNA. The asterisk indicates the termination codon. The polyadenylation signal is underlined.

dideoxymethods [13] is shown in Fig. 1: the total cDNA contains 1469 bases with an open reading frame of 828 nucleotides. In the 3′ untranslated region, the internal *Eco*RI site is found at position 1119 and a putative polyadenylation site is apparent at position 1421.

```
Xen.MyoD  MELLPPPLRDMEVTEGSLCAFPTPDDFYDDPCFNTSDMSFFEDLDPRLVHVTLLKPEEPH 60
Tr.MyoD   ...PDI.FPI----------TS.............H.....E......G....DDH.

Xen.MyoD  HNEDEHVRAPSGHHQAGRCLLWACKACKRKTTNADRRKAATMRERRRLSKVNEAFETLKR 120
Tr.MyoD   .K....I.....................T.....................D.......

Xen.MyoD  YTSTNPNQRLPKVEILRNAIRYIESLQALLH--DQDEAFYPVLEHYSGDSDASSPRSNCS 178
Tr.MyoD   C...........D......S.....G..RGAG.EGNY...MD..............

Xen.MyoD  DGMMDYNSPPCGSRRRNSYDSSFYSDSPNDSRLGKSSVISSLDCLSSIVERISTQSPSCP 238
Tr.MyoD   .....F.GQS.PP....K...TYFNEA.....HK.N.........N.....T.DTSA..

Xen.MyoD  VPTAVDSGSEGS-PCSPLQGETLSERVITIPSPSNTCTQLSQDPSSTIYHVL          289
Tr.MyoD   ---..QD.....S....GD.SIA..NGAP....I.CVPA.-H..-N...Q..
```

Fig. 2. Comparison of the predicted protein sequences of the *Xenopus* (Xen) and Trout (Tr) MyoD cDNAs. The box surrounds the basic and myc-like regions. Note that there is extensive similarity even outside these regions.

The open reading frame encodes a polypeptide of 276 amino acids with 70% of identity with the *Xenopus* MyoD protein (Fig. 2). Within the B-HLH domain, the deduced polypeptide has higher identity with XMyoD (92%) than other *Xenopus* myogenic determinants as Myf 5 (82%), Myogenin (72%) and MRF4 (70%) [14,15]. These comparisons suggest that our clone encodes a teleost homologue of the myogenic factor MyoD. Moreover, outside the conserved basic and myc-like domain which is found in all myogenic factors of vertebrates and invertebrates, we identified four other blocks of conserved sequences which association is typical of the MyoD protein of all vertebrates studied so far [16]. Using XMyoD as the amino acid position reference (Fig 2), these conserved sequences include an acidic amino-terminal region (24–56), a histidine-cysteine rich domain (60–87) and two regions (163–183, 215–230) clustered in the carboxy-terminal half of the protein. On the whole, these data show that a teleost homologue of the myogenic factor MyoD does exist, its presence in fish indicates futhermore that this protein appeared early in vertebrate evolution as a distinct myogenic factor.

Although the teleost MyoD cDNA (TMyoD) was purified from an embryonic muscle cDNA library, it is conceivable that the corresponding transcript accumulates also in adult muscle tissue. To address this possibility we have used Northern blot analysis of RNA isolated from trout adult white muscle which constitutes the major part of the trunk musculature of fish and for comparison from just hatched larvae muscle. We observed one size of TMyoD transcript around 1.6 kb indicating that our cDNA is nearly full length. This transcript is clearly detected both in embryo and adult musculature in a similar amount (Fig. 3A). This result suggests that the corresponding protein may act as a developmental regulator of myogenesis in larvae and may also participate in the maintenance of the muscular phenotype in the adult.

To determine whether TMyoD represents a muscle specific gene product, we examined, by Northern blot, the steady state level of its mRNA in various tissues. As shown in Fig. 3B, TMyoD is present both in red and white fibers which compose the trunk musculature

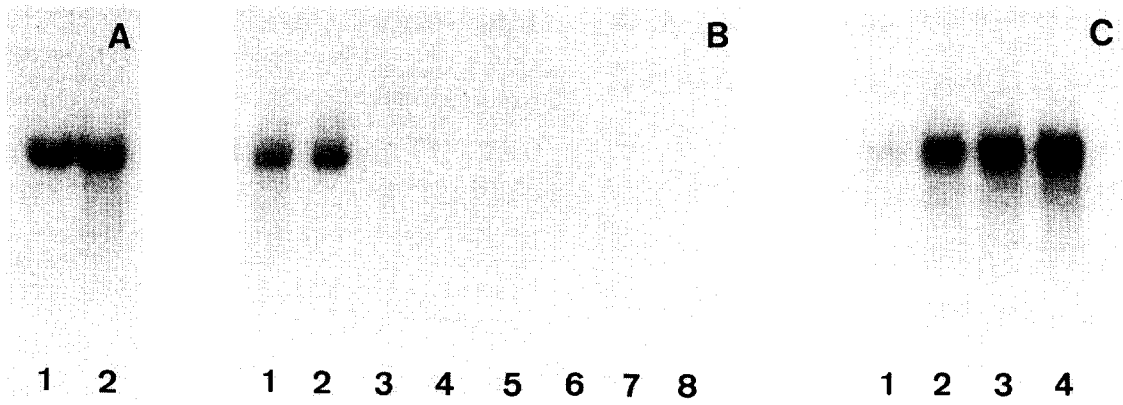

Fig. 3. TMyoD expression. Total RNA (10 μg) were resolved on a formaldehyde gel and transfered to hybond membranes. The blots were hybridized with a 1.1 kb *Eco*RI TMyoD insert radioactively labeled by random priming and washed at high stringency. (A) Northern blot of RNA from hatching embryos (lane 1) and adult (lane 2) trout muscles. (B) Northern blot of RNA from red muscle, white muscle, kidney, testis, gills, intestine, liver and heart (lanes 1 to 8). (C) Northern blot analysis of RNA from trout primary cultures of satellite. Total RNA was isolated from growing cells 4 h (lane 1) and 48 h (lane 2) after seeding, and in differentiated myotubes, 7 days (lane 3) and 11 days (lane 4) after seeding.

of teleosts. No detectable TMyoD mRNA is observed in cardiac (heart) and smooth (intestine) muscle tissues. TMyoD mRNA is undectable in non muscle tissue types including liver, kidney, gills and testis. Thus, our results show that TMyoD expression is restricted to a skeletal muscle lineage in adult animals.

We have also analysed the time course of expression of TMyoD during in vitro differentiation of myosatellite cells isolated from juvenile trouts and cultured at 18°C in DMEM medium supplemented with 10% calf serum. Fig. 3C shows that TMyoD mRNA is slightly detectable in 4-h cultures, its level is dramatically increased in growing myosatellite cell cultures (48 h after seeding) and remains constantly elevated in differentiated myotubes (7 and 11 days after seeding). These observations indicate that TMyoD acts as a major developmental regulator of myogenesis in fish.

We would like to thank Dr. J.B Gurdon for the generous gift of the XMyoD cDNA, Dr. B. Saulier and Dr. P. Thiebaud for critical reading of the manuscript.

References

[1] Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) Cell 51, 987–1000.
[2] Wright, W.E., Sasson, D.A. and Lin, V.K. (1989) Cell 56, 607–617.
[3] Edmonson, D.G. and Olson, E.N. (1989) Gene Dev. 3, 628–640
[4] Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. and Arnold, H.H. (1989) EMBO J. 8, 701–709.
[5] Rhodes, S.J. and Konieczny, S.F. (1989) Gene Dev. 3, 2050–2061.
[6] Edmonson, D.G. and Olson, E.N. (1993) J. Biol. Chem. 268, 755–758.
[7] Venuti, J.M., Goldberg, L., Chakraborty, T., Olson, E.N. and Klein, W.H. (1991) Proc. Nat. Acad. Sci. USA 88, 6219–6223.
[8] Krause, M., Fire, A., Harrisson, S.W., Priess, J. and Weintraub, H. (1990) Cell 63, 907–919.
[9] Paterson, B.M., Walldorf, U., Eldridge, J., Dübendorfer, A., Frasch, M. and Gehring, W.J. (1991) Proc. Natl. Acad. Sci. USA 88, 3782–3786.
[10] Emerson, C.P. (1993) Curr. Opin. Gen. Dev. 3, 265–274.
[11] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–269.
[12] Hopwood, N.D., Pluck, A. and Gurdon, J.B. (1989) EMBO J. 8, 3409–3417.
[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] Hopwood, N.D, Pluck, A. and Gurdon, J.B. (1991) Development 111, 551–560.
[15] Jennings, C.G.B. (1992) Dev. Biol. 150, 121–132.
[16] Charles de la Brousse, F. and Emerson, C.P. (1990) Gene Dev. 4, 567–581.