Volume 335, number 1, 41–46 FEBS 13286 November 1993

The MyoD binding site is dispensable for cardiac actin gene expression in the somites of later stage *Xenopus* embryos

Xi-Lin Su**, Hugh R. Woodland*

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK

Received 27 September 1993

We have studied the transcription directed by the promoter of a *Xenopus borealis* cardiac actin gene fused to a globin reporter sequence after injection into *X. laevis* embryos. This promoter is known to be effective specifically in somites of neurula stage embryos, and work on a very similar *X. laevis* gene has indicated the importance of an upstream MyoD binding site (E box) in this process. We show that, although in the absence of the MyoD site there is a small amount of tissue-specific expression in neurulae, transcription is abundant in the tailbud embryo a few hours later. Thus, although the E box is initially essential, other sequences must adopt the same role soon afterwards.

Actin promoter; Muscle development; MyoD; *Xenopus*

1. INTRODUCTION

The α-cardiac actin gene is one of the first tissue-specific genes to be activated in the *Xenopus* embryo, appearing at the late gastrula stage in the regions shortly to form the somites [1]. The cloned promoter region has been shown to retain its tissue specificity when injected as part of a linear DNA molecule into the cytoplasm of the fertilized *Xenopus* embryo [2,3]. This is also true for the α3 skeletal actin gene [4,5]. Detailed analysis of both promoters has shown that essential components are the CArgG boxes, 10 base sequences (consensus $CC[A/T]_6GG$) that bind members of the serum response factor (SRF) family. In the case of cardiac actin it is only the proximal CArG box that is required [6], but two boxes are needed for the skeletal α3 gene [5]. However, the CArG boxes seem to play only a permissive role in transcription, and other sequences are needed to activate the gene in muscle. For the *X. laevis* cardiac gene, Taylor et al. [7] found that a MyoD binding site, located upstream of the distal CArG box, was sufficient to produce stage- and tissue-specific activaiton of expression. Their experiments were conducted on neurula stage embryos. We have now extended our original analysis of an almost identical *X. borealis* cardiac actin gene [3] to show that the MyoD site is either inoperative or redundant at later stages.

*Corresponding author.

***Present address:* Department of Obstetrics and Gynecology, University of Wisconsin, Madison, WI 53706, USA.

2. MATERIALS AND METHODS

2.1. *Biological materials*

Embryos were obtained, dissected and microinjected at the 2-cell stage as previously described [3]. The DNA for micro-injection (25 μg/ml) was linearised with *Pst*I.

2.2. *DNA manipulation*

The cardiac actin clone used was that of *X. borealis* described by Wilson et al. [3]. The first *Eco*RI fragment of this clone contains the first exon and its flanking regions (790 bp upstream to 489 bp downstream of the start of transcription). The fragment was digested or partially digested with *Sst*I and *Acc*I, appropriate restriction fragments were subcloned inot M13 mp18 and sequenced by the dideoxy chain-termination method [8].

The 5′ regions, from an *Acc*I site 331 bp upstream of the start of transcription (Fig. 1A), to a *Bst*NI site in the first, untranslated exon of the actin gene, were excised from the cloned *Eco*RI fragment mentioned above and the ends infilled. The upstream regions of a *X. laevis* β-globin gene [9], from an *Eco*RI site at the insert/vector boundary to an *Nco*I site at the first translated ATG in exon 1, were excised and the ends filled in and dephosphorylated. The actin upstream regions were then fused into this globin-carrying vector to give an actin/globin fusion gene, pAG1, fused in the two first exons (Fig. 1A).

Further deletions of the actin upstream regions were made by exonuclease III digestion as follows. pAG1 was digested with *Eco*RI followed by *Exo*III (18 U/μg DNA). The reaction was terminated by adding 0.1 vol of 10 × S1 nuclease buffer (0.5 M sodium acetate, pH 4.0, 60 mM $ZnSO_4$). The ends were made blunt by treating with S1 nuclease (10 U/μg DNA, 23°C, 15 min). The reaction was terminated with 0.1 vol. 0.1 M EDTA, 1.5 M sodium acetate and the solution extracted with phenol/chloroform. Any remaining overhangs were eliminated by infilling with the Klenow fragment of DNA polymerase. The constructs were re-circularised in the presence of *Eco*RI linker.

2.3. *RNA analysis*

The total nucleic acid from 4–10 whole embryos or dissected embryo fragments was extracted by a proteinase K/phenol extraction procedure [3]. Correct transcripts from the actin/globin fusion genes were detected by primer extension from an end-labelled 27 base synthetic oligonucleotide complementary to residues 13–39 of the first

exon of the β-globin gene. Methods were as described previously [3] and the products were analysed on 10% sequencing gels. Standard Southern blotting methods were used to analyse distribution of the injected cloned DNA, using the total nucleic acid extracts described above after *Bam*HI digestion [3]. The nick-translated cloned plasmid was used as a probe.

3. RESULTS

We have previously described the transcription of an *X. borealis* cardiac actin/globin fusion gene injected into two-cell *X. laevis* embryos [3]. In this paper we examine the transcription of a series of 5′ deletions of this gene (the clone differs in that the reporter part of the construct was previously from mouse β-globin, but in this case it is from *X. laevis*; see Lakin et al. [5] for similar use of this reporter).

First we sequenced the entire *Eco*RI fragment which encompasses exon 1 of the *X. borealis* cardiac actin gene. The part of this fragment relevant to the experiments described below is shown in Fig. 1. The *X. laevis* and *X. borealis* sequences are extremely similar as far as 410 bp upstream of the start of transcription, but thereafter diverge almost completely. Between −320 bp and −350 bp the upstream region of the promoter contains three E boxes, or binding sites for the MyoD group of helix-loop-helix transcription factors (consensus CANNTG). This region is also present in the *X. laevis* cardiac actin gene, a region called by Taylor et al. [7] the "M-region". At the neurula stage they found it to be essential for transcription. The longest clone studied here, pAG1, was truncated at an *Acc*I site 331 bp upstream of the start of transcription and it contains only the proximal E box (Fig. 2). The wild-type gene also contains an E box consensus sequence at the 3′ end of the first exon, but this was destroyed in fusing the actin fragment to the globin gene. The clone deleted to −258 bp, called $pAG1_{\Delta 258}$, lacks all of the E boxes (Fig. 2). Further downstream are the four CArG boxes, also seen in the *X. laevis* gene. These are progressively removed in the series of deletions shown in Fig. 2.

As in previous studies the cloned genes were injected into 2-cell embryos and, in order to study transcription at the neurula stage (stage 19) the embryos were dissected into animal, equatorial and vegetal fragments at stage 9 (blastula; Fig. 3D), and these fragments were incubated until undissected embryos had reached stage 19. Muscle develops only in the cultured equatorial fraction. Fig. 3B shows that at this stage the transcription of pAG1, the longest construct, was only just detectable and, as expected, it was in the equatorial fraction. In the $pAG1_{\Delta 258}$ deletion as extremely faint band was visible to the eye in the axis fragment. For neither stage were there detectable signals in the intact embryos. This is in line with the results of Taylor et al. [7] and indicates that the incubated embryo fragments are very slightly different from intact embryos, perhaps in timing events. At

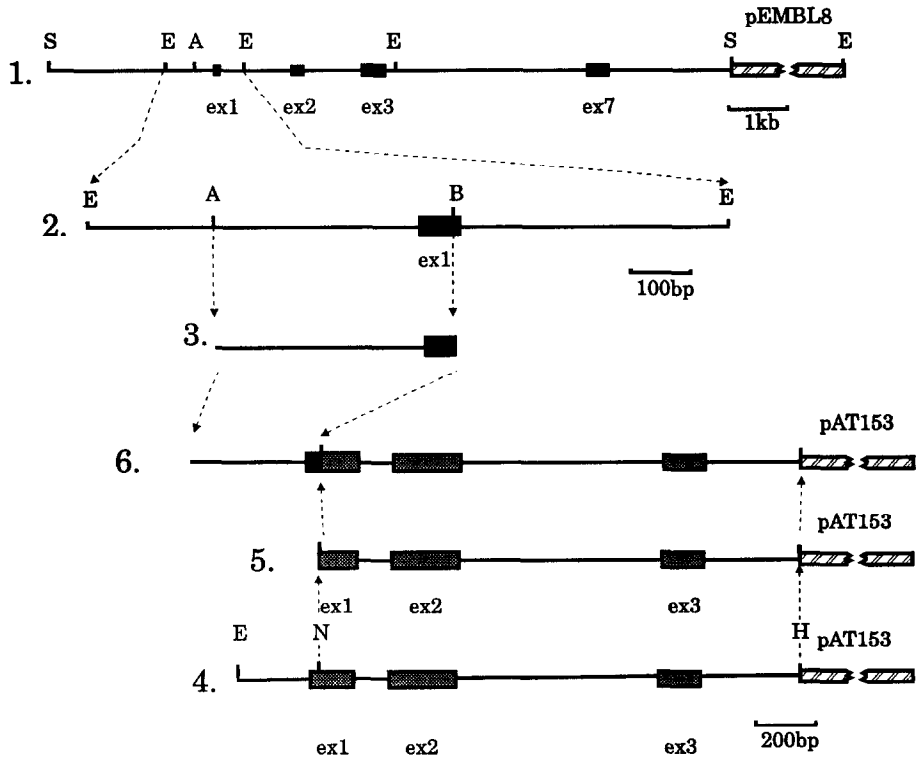

Fig. 1. Construction of the actin/globin fusion clone. (1) Structure of the *X. borealis* cardiac actin gene in the vector, pEMBL8. The position of only four of the exons was determined by sequencing. (2) An *Eco*RI fragment was subcloned from this and (3) used as a source of an *Acc*I–*Bst*NI fragment. (4) Structure of the *X. laevis* β-globin clone [9]. (5) An *Nco*I–*Eco*RI fragment of the latter was isolated. (6) It was fused to the *Acc*I–*Bst*NI fragment described above. Abbreviations for restriction enzyme cutting sites: A, *Acc*I; B, *Bst*NI; E, *Eco*RI; N, *Nco*I.

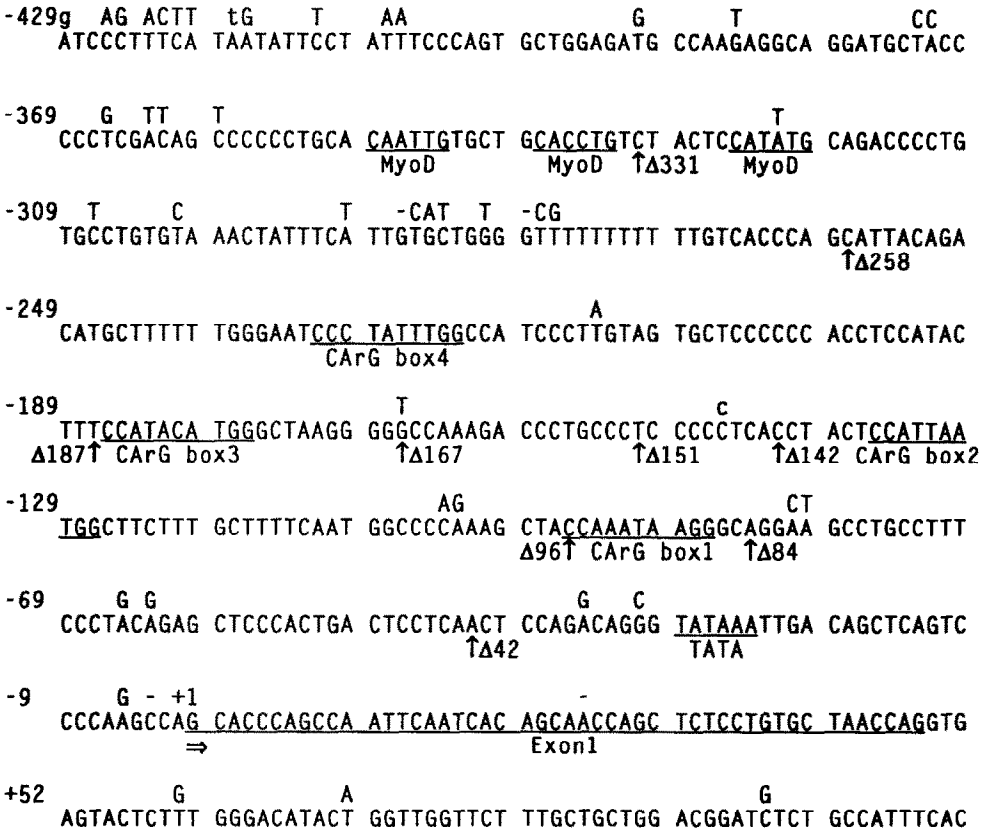

Fig. 2. Sequence of part of the *Eco*RI fragment shown in Fig. 1(2). The 5′ start points of all the clones is shown below the sequence (Δ331–Δ42). CArG boxes and MyoD binding sites are also shown below the sequence. The most downstream MyoD1 site is within the *Bst*NI site used for fusion to globin and was destroyed in this process. Base differences with the *X. laevis* cardiac actin gene [2] are marked above the sequence; extra bases in *X. laevis* are marked in lower case, and are in the space after the residue marked; extra bases in *X. borealis* are marked with a hyphen.

the later tailbud tadpole stage, dissections were performed at the actual stage tested (Fig. 3C). A strong tissue-specific signal was obtained with pAG1 and this was seen only in the muscle-containing axial fragment. Since the injected DNA was found in all fractions of the embryo tested, the reason for an absence of expression in non-muscle-containing regions was not that the DNA was unevenly distributed (Fig. 4). This Southern blot expeirment shows variability of levels of the injected cloned sequence, but there is sometimes strong expression from axis fragments where hybridisation to the injected DNA level is relatively weak (e.g. $pAG_{\Delta 258}$), and absent when it is strong (e.g. $pAG_{\Delta 151}$).

Although $pAG1_{\Delta 258}$ lacks all E boxes, the signal it generated was not reproducibly different from that of pAG1 (Fig. 3A). There were detectable, but markedly reduced signals in $pAG1_{\Delta 187}$–$pAG1_{\Delta 151}$; these were 10% or less of those from pAG1. These differences did not simply reflect the level to which DNA had replicated in the embryo (Fig. 4). Fig. 2 shows that $pAG1_{\Delta 258}$ contains all four CArG boxes, but $pAG1_{\Delta 187}$ lacks CArG box4; $pAG1_{\Delta 167}$ and $pAG1_{\Delta 151}$ in addition lack CArG box3. Greater deletions did not give detectable transcription and it is interesting that $pAG1_{\Delta 142}$, which is the cut-off point for transcription, is only slightly shorter than $pAG1_{\Delta 151}$ and retains CArG box3.

These results therefore indicate that, while a cardiac promoter clone lacking a MyoD1 binding E box has little or no transcriptional activity in the neurula, a few hours later it is highly active in the muscle-forming region in the tailbud tadpole.

4. DISCUSSION

The restriction of promoter activity to particular types of cells usually depends on combinations of sites which bind ubiquitous transcription factors and those that bind tissue-specific factors. In the case of striated muscle-specific α-actin genes the important ubiquitous factors are the CArG boxes, or serum-response elements, which bind members of the serum response factor (SRF) family [10]. For example in the case of the *X. laevis* cardiac actin gene, at least one CArG box is required for the function of cloned genes injected into the embryo [6] and *Xenopus* embryos contain several species of SRF [9,11]. A *X. borealis* skeletal actin gene seems to require at least two CArG boxes for activity in the embryos [5]. However, while SRF binding is necessary, it is not sufficient for actin expression in the muscles of the embryo, and one class of transcription factor shown to be capable of providing the necessary tissue-specific component are the muscle-specific helix-

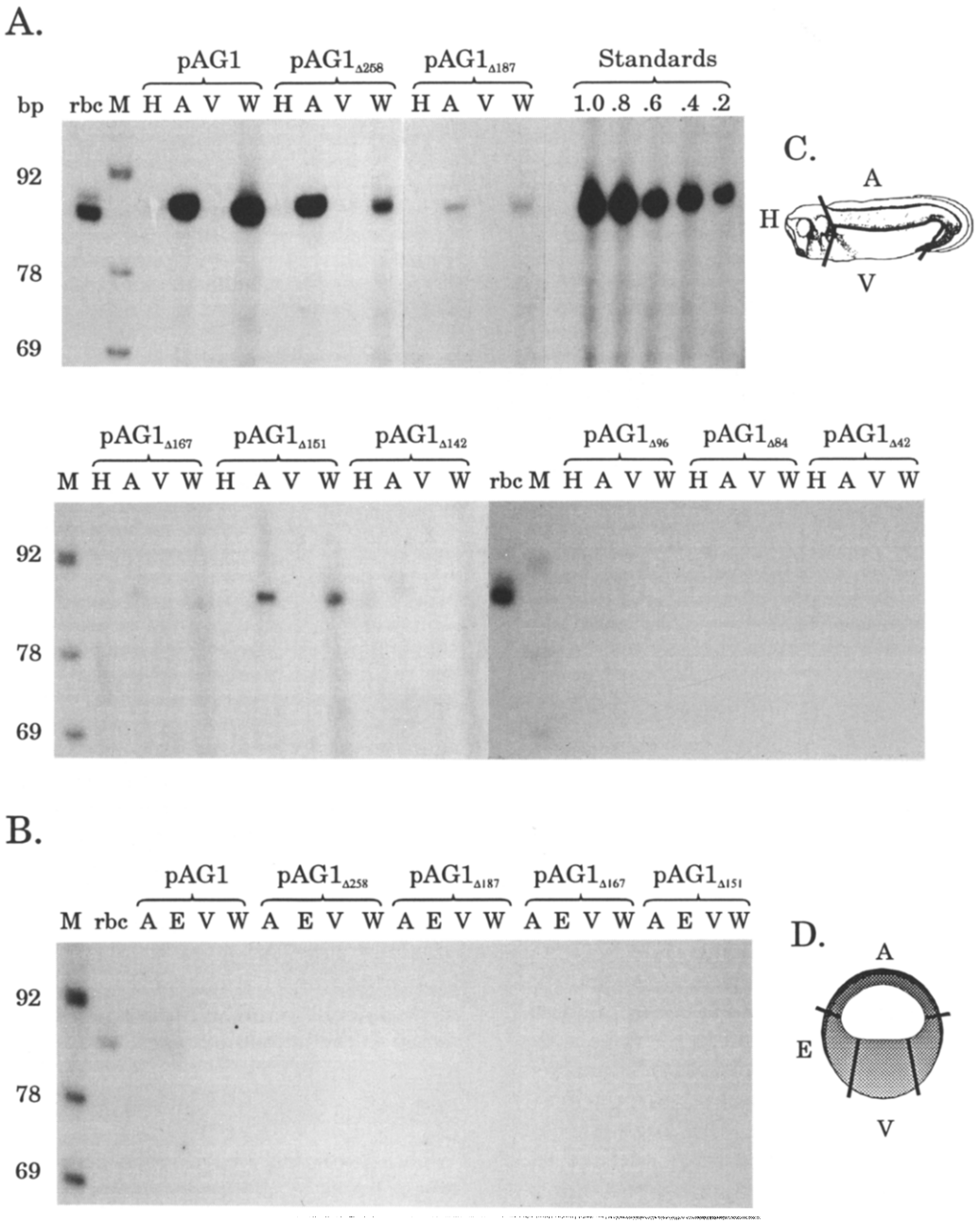

Fig. 3. Primer extension analyses of stage 26 tailbud embryos injected with actin/globin fusion genes pAG1–pAG$_{\Delta 42}$. (A) Embryos injected at the 2-cell stage and dissected at stage 26, as shown in C. Molecular weights of the markers in lane M are on the left. W indicates whole embryos and H, A and V are as in panel C; rbc is a primer extension product from whole red blood cell RNA used as a standard and fortuitously the same size as that produced by the actin promoter. For the dissected parts the amount contained in one embryo was loaded; allowing for experimental variability, the parts should therefore add up to the signal given by the whole embryos. On the right is a series of dilution standards of the pAG1 A sample. (B) Stage 19 analyses. The embryos were dissected into animal, A, vegetal, V, and equatorial, E, fragments at stage 9, as shown in (D), and incubated on to stage 19. Other details as in A.

loop-helix-containing proteins, of which MyoD1 was the prototype. This was identified by virtue of its ability to switch on the terminal myotube phenotype in C3H 10T$^1/_2$ fibroblasts [12]. Other members of the family include myogenin, Myf-5, and MRF4, also called herculin and myf-6 [13–17].

Xenopus embryos are known to express two MyoD genes, myf-5 and MRF4 [18–23]. Ubiquitous over-expression of MyoD will turn on the cardiac actin gene in tissues where it is not normally expressed, but the fate of these cells is not diverted to muscle [24,25]. However, Taylor et al. [7] have argued that the E boxes are essential for expression of the cloned *X. laevis* cardiac α-actin gene. Their assays were performed on tissue fragments from neurulae and included a number of internal deletions of the promoter. Our results agree with theirs in

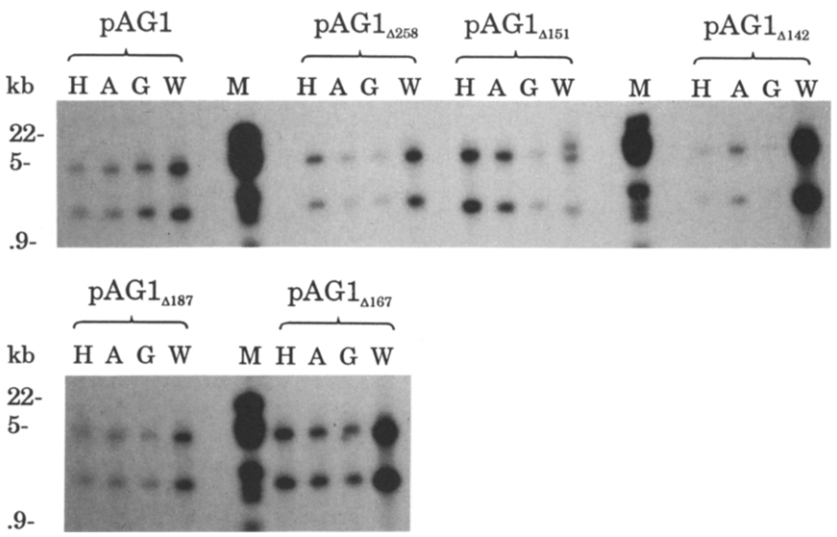

Fig. 4. Southern blot analyses of injected DNA in samples also analysed for actin RNA. Aliquots of the total nucleic acid extracts analysed by primer extension in Fig. 3A, were digested with *Bam*HI, which produces two DNA fragments from the replicated linear concatenates of injected DNA found in embryos injected with linear actin gene constructs. The DNA was analysed by agarose gel electrophoresis, Southern blotted and probed with nick-translated pAG1.

that removing all three E boxes in the promoter gave indetectable transcription in the muscle-forming region of the neurula. There was also very little transcription from a promoter containing a single E box, as opposed to the three in the wild-type sequence. However, at the tailbud tadpole stage, only a few hours later, both deletion clones were highly active in the appropriate region. Thus in the tailbud, embryo sequences other than the E box sequences must be capable of directing appropriate expression of the gene. In the intact gene there are therefore likely to be at least two regions separately controlling the gene; indeed, functional redundancy in the promoter of a *Xenopus* α3 skeletal actin gene has been noted previously [5]. It is not clear what region has taken over control of the gene in the absence of the E box sites. The data in Fig. 3 suggest that there is an important sequence between −151 and −142. At this point there is the same binding sequence for EMF-1, also called C/EBP, as was reported in the α1-antitrypsin gene [26]. However, the significance of this is unclear since C/EBP is not expressed in skeletal muscle, at least in mammals [27].

The fact that a gene should not necessarily depend on the MyoD family for transcription in skeletal muscle is not unprecedented, for example, the β-myosin gene enhancer does not require E boxes [28]. Further, gene knock-out experiments indicate that MyoD and myf-5 are individually redundant for the formation of most muscles and some muscles can form in the absence of myogenin [29–32]. However, it is not yet clear whether skeletal muscle can ever form in the collective absence of the whole family. Lastly some, at least, of the family can be expressed in a cell without activating many muscle-specific genes. This was shown in the electric organ of *Torpedo* where the MyoD, myogenin and myf-5 levels are similar to those in skeletal muscle, but the proper differentiated muscle phenotype is not expressed [33]. What we show here is that for the cardiac actin gene there is an early developmental stage when its E box is essential, but a few hours later it is not. Whether this is true for other genes in these developing somites remains to be seen.

REFERENCES

[1] Mohun, T.J., Brennan, S., Dathan, N., Fairman, S. and Gurdon, J.B. (1984) Nature 311, 716–721.
[2] Mohun, T., Garrett, N. and Gurdon, J. (1986) EMBO J. 5, 3185–3193.
[3] Wilson, C., Cross, G.S. and Woodland, H.R. (1986) Cell 47, 589–599.
[4] Boardman, M., Cross, G.S., Jones, E. and Woodland, H.R. (1992) Eur. J. Biochem. 208, 241–249.
[5] Lakin, N.D., Boardman, M. and Woodland, H.R. (1993) Eur. J. Biochem. 214, 425–435.
[6] Mohun, T.J., Garrett, N. and Taylor, M.V. (1989) EMBO J. 8, 1153–1161.
[7] Taylor, M.V., Gurdon, J.B., Hopwood, N.D., Towers, N. and Mohun, T.J. (1991) Genes Dev. 5, 1149–1160.
[8] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[9] Patient, R.K., Harris R., Walmsley, M.E. and Williams, J.G. (1983) J. Biol. Chem. 258, 8521–8523.
[10] Minty, A. and Kedes, L. (1986) Mol. Cell. Biol. 6, 2125–2136.
[11] Mohun, T.J., Chambers, A.E., Towers, N. and Taylor, M.V. (1991) EMBO J. 10, 933–940.
[12] Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) Cell 51, 987–1000.
[13] Wright, W.E., Sassoon, D.A. and Lin, V.K. (1989) Cell 56, 607–617.
[14] Edmonson, D.G. and Olson, E.N. (1989) Genes Dev. 3, 628–640.
[15] Braun, T., Rudnicki, M.A., Arnold, H.H. and Jaenisch, R. (1992) Cell 71, 369–382.

[16] Rhodes, S.J. and Konieczny, S.F. (1989) Genes Dev. 3, 2050–2061.
[17] Miner, J.H. and Wold, B. (1990) Proc. Natl. Acad. Sci. USA 87, 1089–1093.
[18] Hopwood, N.D., Pluck, A. and Gurdon, J.B. (1989) EMBO J. 8, 3409–3417.
[19] Harvey, R.P. (1990) Development 108, 669–680.
[20] Harvey, R.P. (1991) Proc. Natl. Acad. Sci. USA 88, 9198–9202.
[21] Harvey, R.P. (1992) Mech. Dev. 37, 141–149.
[22] Scales, J.B., Olson, E.N. and Perry, M. (1990) Mol. Cell. Biol. 10, 1516–1524.
[23] Jennings, C.G.B. (1992) Dev. Biol. 150, 121–132.
[24] Hopwood, N.D. and Gurdon, J.B. (1989) Nature 347, 197–200.
[25] Rashbass, J., Taylor, M.V. and Gurdon, J.B. (1992) EMBO J. 11, 2981–2990.
[26] Costa, R.H., Grayson, D.R., Xanthopoulos, K.G. and Darnell, J.E. (1988) Proc. Natl. Acad. Sci. USA 85, 3840–3844.
[27] Birkenmeier, E.H., Gwyn, B., Howard, S., Jerry, J., Gordon, J.I., Landschulz, W.H. and McKnight, S.L. (1989) Genes Dev. 3, 1146–1156.
[28] Thompson, W.R., Nadal-Ginard, B. and Mahdavi, V.J. (1991) Biol. Chem. 266, 22678–22688.
[29] Rudnicki, M.A., Braun, T., Hinuma, S. and Jaenisch, R. (1992) Cell 71, 383–390.
[30] Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. and Arnold, H.H. (1989) EMBO J. 8, 701–709.
[31] Hasty, P., Bradley, A., Morris, J.H., Edmondson, D.G., Venuti, J.M., Olson, E.N. and Klein, W.H. (1993) Nature 364, 501–506.
[32] Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S.W., Nonaka, I. and Nabeshima, Y. (1993) Nature 364, 532–535.
[33] Neville, C.M. and Schmidt, J. (1992) FEBS Lett. 305, 23–26.