A bovine homolog to the human myogenic determination factor *myf-5*: sequence conservation and 3' processing of transcripts

Theodore G. Clark, Julie Morris, Maho Akamatsu¹, Royal A. McGraw¹ and Robert Ivarie*

Department of Genetics and ¹Department of Physiology and Pharmacology, University of Georgia, Athens, GA 30602, USA

Received March 23, 1990; Accepted April 26, 1990

EMBL accession no. X52526

ABSTRACT

A bovine cDNA library from fetal skeletal muscle myoblasts was screened with a 274 bp probe to a conserved region of the mouse *MyoD1* cDNA. One positive recombinant, designated *bmyf*, was found to contain a 1931 bp insert with an open reading frame encoding a predicted protein highly related to the human myogenic factor *myf-5*. Human and bovine factors are 96% homologous in their predicted amino acid sequences. At the nucleotide level, *bmyf* and *myf-5* are 92% identical in the coding region and 74 and 80% homologous in their 5'- and 3'-untranslated regions, respectively. The *bmyf* cDNA, nevertheless, extends 475 nucleotides beyond a polyadenylation signal common to both cDNAs. *Bmyf* transcripts are expressed exclusively in skeletal muscle where three transcripts of 1.5, 2 and 3 kb were detected. While the 1.5 kb transcript lacks sequences 3' to the polyadenylation signal at nt 1415 in the *bmyf* cDNA, both the 2 and 3 kb RNAs contain these sequences suggesting that *bmyf* transcripts are alternatively polyadenylated. *Bmyf* cDNA can activate the expression of the myogenic program in C3H10T1/2 fibroblasts as assayed by stable and transient transfection experiments.

INTRODUCTION

The developmental program leading to the formation of vertebrate skeletal muscle is widely thought to be under the control of a hierarchy of regulatory genes which direct successive stages in development (see, for example, 1-3). Several genes which appear to play a role in myogenic lineage determination and/or myocyte differentiation have recently been identified and are designated *MyoD1* (4), *myd* (2), *myogenin* (5, 6), *myf-5* (7), *mrf4* (8) and *ski* (9). *MyoD1*, *myf-5*, *myogenin* and *mrf4* were isolated from cDNA libraries of murine, human and rat skeletal muscle cells, respectively. When placed under the control of retroviral promoters, they stably convert fibroblasts to a skeletal muscle phenotype following DNA transfection. In transient co-transfection assays with C3H10T1/2 cells, they activate the

expression of muscle-specific gene constructs, and in the case of *MyoD1*, induce expression of endogenous muscle-specific genes in differentiated cells of both mesodermal and non-mesodermal origin (7,10,11). A *MyoD1* fusion protein also binds to the muscle-specific creatine kinase promoter (12). Transcripts from these genes are present only in skeletal muscle tissues, and both *MyoD1* and *myogenin* appear in presumptive muscle at early stages of embryogenesis (5,13,14).

Based on DNA sequence comparisons, *MyoD1*, *myf-5*, *myogenin* and *mrf4* are the products of different genes, but share relatively long stretches of sequence homology and constitute members of a related gene family (4,5,7,8). Indeed, *myf-5*, *mrf4* and mouse *myogenin* were isolated from respective cDNA libraries using probes against the mouse *MyoD1* cDNA (6-8). Apparent homologs to mouse *MyoD1* have also been isolated from chicken and *Xenopus* skeletal muscle cDNA libraries using the same approach (10,14). At least two regions within the myogenic factor cDNAs are strongly conserved and are essential for conversion of C3H10T1/2 cells to the skeletal muscle phenotype (15). These include nucleotide sequences encoding a highly basic group of amino acids (residues 102-135 in *MyoD1*), and an adjacent set (residues 143-162) having similarity to a region within the *myc* family of proteins (4,15). Related segments are present in a number of *Drosophila* regulatory gene products and in two proteins which bind to the immunoglobulin κ chain gene enhancer (4; 7,16). Such segments have the potential of forming amphipathic helices separated by an intervening loop, and apparently constitute a novel DNA binding and protein dimerization motif (16).

We set out to isolate myogenic regulatory factors from bovine skeletal muscle using a probe which spanned functional domains of the *MyoD1* cDNA. Of several positive cDNA clones isolated, one was identified as the bovine equivalent of human *myf-5*. Here we describe its sequence, expression and functional properties.

MATERIALS AND METHODS

Cells and growth conditions

C3H10T1/2 (clone 8) from the American Type Culture Collection was grown in DMEM supplemented with 10% heat-inactivated

* To whom correspondence should be addressed

Figure 1. Nucleotide and amino acid sequence of bovine *myf* cDNA.

Various tissues were excised from 20 cm bovine fetuses and frozen in liquid nitrogen. Skeletal muscle came from hind and forelimbs. For RNA preparation, frozen tissue was ground to a fine powder and lysed in 5 M guanidine thiocyanate, 10 mM EDTA, 50 mM Tris·HCl (pH 7.5) and 8% (v/v) β -mercaptoethanol. RNA was precipitated directly from the lysate with 4 M LiCl as described (29). After solubilization and extraction with phenol, RNA was ethanol precipitated and stored either as LiCl pellets or as aqueous solutions at -70°C (29). PolyA⁺ RNA was isolated as above, electrophoresed on 1.1% agarose-formaldehyde gels, and transferred onto Biotrans nylon membranes in 20 $\times$ SSC. Filters were UV-irradiated, baked at 70°C for 60–90 min, and hybridized to ^{32}P -labelled cDNAs as follows. Filters were prehybridized for 4 hr at 65°C in 0.5 M sodium phosphate (pH 7.2), 7% SDS, 1% BSA, and 1 mM EDTA, then hybridized with base-denatured probe ($>10^6$ cpm/ml) for 18–24 hr under the same conditions (30). Filters were washed twice at 65°C in 40 mM sodium phosphate buffer (pH 7.2), 5% BSA, 5% SDS and 1 mM EDTA for 30 min each, followed by two washes at 65°C in 40 mM sodium phosphate (pH 7.2), 1 mM EDTA and 1% SDS (30). Fragments from cDNAs were gel purified before labelling as probes (31).

DNA transfections were carried out using cDNAs inserted into the *Eco*RI site between the Moloney sarcoma virus LTR and

SV40 polyadenylation signals of the pEMSV-scribe expression plasmid (4,32). Full-length *bmyf* cDNA was inserted in both sense and anti-sense orientations. For transient transfections (5), C3H10T1/2 cells were plated at 1×10^5 cells per 10 cm dish and grown for 2 days in DMEM containing 10% heat-inactivated fetal calf serum (FCS). Medium was changed 3 hr prior to transfection and plasmid DNA (7.5 μ g/dish) was added as a calcium phosphate precipitate as described (33). Cells were incubated 24 hr, washed with fresh medium and grown for 2 days before transfer to differentiation medium (DMEM with 2% heat-inactivated horse serum). For stable transfections, C3H10T1/2 cells were plated at 2×10^5 cells per 10 cm dish and cotransfected as above with the G418-resistance plasmid pRSVneo (0.2 μ g/dish) and appropriate *bmyf* or *MyoD1* expression vectors (10 μ g/dish) in the presence of 10 μ M chloroquine (4). After 24 hr the medium was replaced with DMEM containing 15% heat-inactivated FCS and cultures were maintained for an additional 24 hr before replating cells at 1/10th dilution in the presence of 0.4 mg/ml G418. Once macroscopic colonies had formed, cells were maintained for 3 days in differentiation medium containing 0.2 mg/ml G418. For immunostaining, cells were washed in Tris-buffered saline (150 mM NaCl, 20 mM Tris $\cdot$ HCl, pH7.5), fixed in ethanol:formaldehyde:glacial acetic acid (70:3.7:5) (4), then reacted sequentially with monoclonal antibody to the heavy chain of skeletal muscle myosin (MF-20; 34) followed by goat anti-mouse IgG conjugated to alkaline phosphatase. Antibody reactions were carried out at room temperature for 1 hr and immunostaining was performed with NBT and BCIP color reagents (Promega). Photomicrographs were taken with a Zeiss inverted microscope using Kodak Panatomic-X film.

Figure 2. Predicted protein sequence of *bmyf* and human *myf-5*. The *bmyf* sequence is shown above that of *myf-5*. Rectangular boxes indicate amino acid differences. Regions of sequence conservation with *MyoD1* are underlined. The *myc* homology region is double underlined.

RESULTS

Isolation of bovine *myf* cDNA

A cDNA library from fetal myoblasts was probed under conditions of reduced stringency with a 274 bp fragment (nt 339–612) from a conserved region of the mouse *MyoD1* cDNA. Inserts from several positive recombinants were characterized by restriction endonuclease mapping and Southern hybridization with bovine genomic DNA (data not shown). Based on DNA sequence, one cDNA was found to be nearly identical to the recently identified human myogenic factor *myf-5* (7), and is designated bovine *myf* (or, *bmyf*). The nucleotide sequence of the *bmyf* insert is shown in Figure 1. It contains 1,889 bases, excluding a 42 nt polyA-tail, and a long open reading frame of 765 nt beginning with a potential ATG start site at nt 69. The deduced amino acid sequence can encode a protein of 255 amino acids identical in size to the protein from the longest open reading frame of the human *myf-5* cDNA (7). As shown in Figure 2, *bmyf* and human *myf-5* share extensive homology particularly in their coding regions, with 96% identity at the amino acid level. Fifteen amino acid substitutions are found, nearly half being conservative. At the nt sequence level, the cDNAs are 92% identical in their coding regions. Both cDNAs contain short 5' untranslated regions (68 nt for *bmyf*, and 42 nt for human *myf-5*) upstream of the deduced ATG start codons in each. These regions are also related showing 74% sequence conservation. Although it is not known whether the first nt is the actual 5' transcriptional start site of the *bmyf* mRNA, an independently cloned and smaller bovine cDNA (data not shown) contains the identical 5' sequence including the first nt suggesting that *bmyf* is a full-length cDNA.

The principal difference between *bmyf* and human *myf-5* occurs in the 3'-untranslated region where *bmyf* is over 400 nt longer. From the stop codon to an AATAAA polyadenylation signal at nt 1415 in *bmyf* and nt 1407 in human *myf-5*, the two cDNAs are highly related and have 80% sequence similarity. *Myf-5* terminates within a short distance of this polyadenylation signal, while *bmyf* extends an additional 475 nt past a second polyadenylation signal at nt 1872. Presumably, the human *myf-5* cDNA came from a transcript terminating at the upstream site, while *bmyf* came from a transcript terminating at the downstream site. As shown next, RNA from fetal skeletal muscle contains a 1.5 kb *bmyf* transcript apparently arising from adenylation at AATAAA₁₄₁₅.

Northern analysis

When polyA⁺ RNA from fetal skeletal muscle was analyzed on Northern blots using *bmyf* cDNA as a probe, three RNA species were detected migrating at 1.5, 2 and 3 kb (Figure 3a). The 1.5 kb band was the most abundant, being present at 2–3 times the level of the two larger transcripts. The existence of three distinct bands on Northern blots suggested that transcripts from the *bmyf* gene are differentially processed. For example, the 1.5 kb RNA band might be equivalent to transcripts which generated the human *myf-5* cDNA and lack the region distal to the polyadenylation signal AATAAA₁₄₁₅ in *bmyf*. To test this, Northern blots of polyA⁺ RNA from bovine fetal skeletal muscle were probed with a 354 bp *AvaII/EcoRI* fragment from the distal end of *bmyf*, 3' to the polyadenylation signal at nt 1415. As shown in Figure 3b, the distal 354 bp fragment hybridized strongly to the 2 and 3 kb bands, but very weakly to the 1.5 kb RNA, suggesting that this species lacks the same 3' sequence missing in the human *myf-5* cDNA. The 3' probe shares 61% similarity to a 105 nt region (nt 1186–1290) upstream of the

Figure 3. *Bmyf* transcripts in fetal skeletal muscle RNA and alternative processing. (A) 2 μ g polyA⁺ RNA from bovine fetal skeletal muscle was electrophoresed on a 1.1% agarose/formaldehyde gel and subjected to Northern hybridization analysis with ³²P-labelled *bmyf* cDNA. Three RNA species are clearly evident migrating at approximately 1.5, 2 and 3 kb. (B) Adjacent lanes of a 1.1% agarose gel were loaded with 2 μ g of polyA⁺ RNA from the same preparation as in (A) above, then electrophoresed and subjected to Northern hybridization analysis with either a 3' *bmyf* fragment (nt 1589–1942) (lane 1), or with the 1.9 kb *bmyf* cDNA (lane 2).

Figure 4. Tissue specificity of *bmyf* expression. PolyA⁺ RNA from different fetal bovine tissues was subjected to Northern hybridization analysis with the full-length *bmyf* cDNA as a probe. Conditions were the same as those used in Figure 3. The lane on the far right (marked *bmyf*) contained 2 μ g polyA⁺ RNA from fetal kidney and approximately 8 pg of anti-sense *bmyf* RNA. A pGEM[®] -3zf(-) plasmid containing the full-length *bmyf* cDNA was used as the template for the synthesis of anti-sense RNA in reactions with T7 RNA polymerase.

polyadenylation signal AATAAA₁₄₁₅ and most likely accounts for the weak hybridization signal seen with this fragment.

To determine whether *bmyf* is expressed in a tissue-specific manner, polyA⁺ RNAs from several fetal bovine tissues were subjected to Northern analysis and probed with the 1.9kb *bmyf* cDNA (Figure 4). Only RNA from skeletal muscle gave rise to detectable bands. By comparison with the signal intensity from known amounts of *bmyf* RNA synthesized from a pGEM[®] vector containing the *bmyf* insert, the levels of *bmyf* transcripts in skeletal muscle are low and estimated to be 0.01%-0.001% of total message. Little or no signal was detected in RNA from mouse azamyoblast cell lines at the levels of stringency used above (data not shown).

***Bmyf* is myogenic In C3H10T1/2 cells**

Human *myf-5* cDNA converts mouse C3H10T1/2 cells into myoblasts that subsequently form myotubes and activate muscle-specific genes (7). We have used a pEMSV-scribe retroviral expression vector containing the *bmyf* cDNA to show that the bovine *myf* cDNA can convert C3H10T1/2 cells into myoblasts following both transient and stable transfection. In transient assays, C3H10T1/2 cells were transfected with the sense-oriented expression vector (pEMSV*bmyf*+), grown for 2 days and then transferred to low mitogen-containing medium before fixing and reacting with a monoclonal antibody to skeletal muscle myosin

heavy chain (MF20; 34). Cells transfected with the vector carrying *bmyf* in its anti-sense orientation (pEMSV*bmyf*-), or mouse *MyoD1* in its sense orientation (pEMc11s) served as negative and positive controls, respectively. Cells which reacted positively with the muscle myosin antibody probe following transfection with the sense-oriented *bmyf* and *MyoD1* expression plasmids are shown in Figure 5A,B. In random fields, *MyoD1* transfected cultures contained ~8-fold more positive cells than *bmyf* transfected cultures. No antibody-positive cells were observed in cultures transfected with the bovine *myf* expression plasmid in the anti-sense orientation. Cells converted to the muscle myosin phenotype with the bovine *myf* expression vector were present at ~10⁻³ frequency which is several orders of magnitude higher than the spontaneous activation level in C3H10T1/2 cells (35).

In stable transfection assays, C3H10T1/2 cells were subjected to cotransfection with pRSVneo and the sense *bmyf* or *MyoD1* expression vectors described above. Following selection in G418, colonies with large multinucleated syncytia which stained positive with muscle-specific myosin antibodies were clearly visible in both *bmyf* and *MyoD1* transfected cultures (Figure 5C,D). Although virtually the same percentage of colonies were positive for muscle myosin expression in the *bmyf* and *MyoD1* plates (~15%), the total number of G418-resistant colonies was ~4-fold less in the *MyoD1* transfected cultures (data not shown).

Figure 5. Immunostaining of skeletal muscle myosin heavy chain in C3H10T1/2 cells transfected with myogenic factor cDNAs. Upper panels show individual cells from transient assays. Lower panels show multinucleated myotubes in G418 resistant colonies resulting from stable transfections. Cultures were transfected with sense-oriented *MyoD1* (panels A and C) or *bmyf* (B and D) expression vectors. Darkly stained cells are expressing skeletal muscle myosin heavy chain. (Magnification A and B = 270 ×; C and D = 72 ×).

As discussed previously (4,10), overexpression of *MyoD1* may lead to rapid, direct conversion of single cells to the differentiated phenotype and prevent the formation of proliferating stem cells which could give rise to individual myogenic colonies.

DISCUSSION

We have cloned a bovine cDNA homolog (*bmyf*) to the human myogenic factor, *myf-5*. Based on sequence comparisons, the degree of conservation between the human and bovine cDNAs is striking. Both contain open reading frames encoding predicted proteins of 255 amino acids of which 96% are identical. While there are 15 amino acid differences between *bmyf* and *myf-5*, human *MyoD1* and myogenin differ from their rodent counterparts by 31 and 33 amino acids, respectively (36), indicating that *myf-5* factors are more highly conserved than other members of the *MyoD1*-related gene family. The region of strongest conservation between *bmyf* and human *myf-5* is a 116 amino acid segment extending from residues 46-161, which has undergone only a single conservative amino acid substitution (glu → asp). This region includes the basic/*myc* homology region thought to play a role in DNA binding (16), contains the minimal region capable of myogenic conversion of C3H10T1/2 cells to myoblasts (15) and is highly conserved among other members of the *MyoD1*-related family, including *MyoD1*, *myogenin*, and *mrf4* (4,5,8). Aside from the basic/*myc* homology region, two ser/thr-rich segments (aa 157-167 and 202-216 in *bmyf*) are also shared among *bmyf*, human *myf-5* and *MyoD1* (Figure 2). In these domains, only a single amino acid substitution (tyr → cys) has taken place between human *myf-5* and *MyoD1*, while 3 amino acid substitutions (exclusively ser → asn) have occurred between human *myf-5* and *bmyf*. Although the functional significance of these changes is unclear, ser/thr-rich segments are absent in *myogenin*, and are only partially conserved in *mrf4*.

In addition to similarities in their coding regions, close homologies between *bmyf* and human *myf-5* are also found in the 5' and 3'-untranslated regions, which contain 74% and 80% sequence conservation, respectively. This contrasts with mouse and human cDNAs encoding the structural protein α -actin, which show higher sequence conservation across species lines at the coding level (100% amino acid identity between human and mouse actins) but which are substantially more diverged in their non-coding sequences (37,38,39). As is the case with *bmyf* and *myf-5*, the 3' untranslated regions of rat and human *myogenin*, and mouse and human *MyoD1* are also substantially conserved (7). These observations suggest that these regions have an important functional role in the overall expression of the myogenic factor genes, perhaps in the processing of nuclear transcripts or turnover of mature messages. In this regard, an obvious feature of the 3'-untranslated regions of *bmyf* and human *myf-5* cDNAs are the presence of AT-rich sequences (AUU[U]_n and AUUUA) which have been proposed to play a role in selective degradation of transiently expressed mRNAs (40). Sequences beginning with a single adenosine residue followed by a tract of three or more U's occur frequently in the 3'-untranslated regions of a variety of low abundance cytokine, lymphokine and proto-oncogene transcripts, and the insertion of a region bearing such sequences into the 3'-untranslated region of normally stable β -globin mRNA substantially increases turnover of the transcript *in vivo* (40).

The principal difference between *bmyf* and human *myf-5* occurs at the distal end of the 3'-untranslated region. A segment

extending from nt 1441–1889 in *bmyf* is absent from the human cDNA homolog. The smallest of three *bmyf* transcripts detectable by Northern analysis (1.5, 2 and 3 kb) also appears to lack this distal segment. Since there are two polyadenylation signals in the 1.9 kb *bmyf* cDNA (beginning at residues 1415 and 1872), the most likely explanation for these observations is that *bmyf* transcripts can be adenylated at either of two sites. In one case, a ~2.0 kb message would be generated, which presumably gave rise to the *bmyf* cDNA described here. In the second case, a shorter message (~1.5 kb) would be generated which would lack the region distal to the AATAAA₁₄₁₅ polyadenylation signal. The human *myf-5* cDNA described previously by Braun *et al.* (7) contains a single polyadenylation signal within a short distance of its 3'-terminus and virtually in the same location as the upstream polyadenylation signal in *bmyf*. Thus, human *myf-5* cDNA may have arisen from an equivalent of the ~1.5 kb RNA species seen in bovine skeletal muscle. The possibility of alternative polyadenylation has also been raised in connection with recent studies of CMD1, an avian homolog to the mouse *MyoD1* cDNA. In this case, a polyadenylation signal is present roughly 200 residues upstream of the 3'-terminus of the CMD1 cDNA, and transcripts corresponding in size to what would be expected from polyadenylation occurring at alternative sites are present in RNA from chicken skeletal muscle (10). Whether differential processing at the 3' end of myogenic factor cDNAs plays a functional role in their activity remains to be determined.

ACKNOWLEDGEMENTS

The authors extend their thanks to A. B. Lassar and H. Weintraub for the plasmid vector pVZc11b and the expression vectors pEMc11s and pEMSVscribe, and to D. A. Fischman for MF-20 antibody supernatant. We also thank Todd Arnold for his expertise and help with the transient transfection assays, and Mr. Norm Shapiro of Shapiro Meat Packing Company, Augusta, GA for bovine fetuses. This work was supported in part by a grant from the National Cancer Institute (CA34066) to RDI.

REFERENCES

- Blau, H. M. (1988). *Cell* **53**, 673–674.
- Pinney, D.F., Pearson-White, S.M., Konieczny, S.F., Latham, K.E. and C.P. Emerson (1988). *Cell* **53**, 781–793.
- Gurdon, J. B., Mohun, T. J., Sharpe, C. R. and M. V. Taylor (1989). *Trends Gen.* **5**, 51–56.
- Davis, R. L., Weintraub, H. and A. B. Lassar (1987). *Cell* **51**, 987–1000.
- Wright, W. E., 13, D. A. and V. K. Lin (1989). *Cell* **56**, 607–617.
- Edmonson, D. G. and E. N. Olson (1989). *Genes Dev.* **3**, 628–640.
- Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. and H. H. Arnold (1989). *EMBO J.* **8**, 701–709.
- Rhodes, S.J. and S.F. Konieczny (1989). *Genes Dev.* **3**, 2050–2061.
- Colmenares, C. and E. Stavnezer (1989). *Cell* **59**, 203–213.
- Lin, Z.-Y., Dechesne, C. A., Eldridge, J. and B. M. Paterson (1989). *Genes Dev.* **3**, 986–996.
- Weintraub, H., Tapscott, S. J., Davis, R. L., Thayer, M. J., Adam, M. A., Lassar, A. B. and A. D. Miller (1989). *Proc. Nat. Acad. Sci. (USA)* **86**, 5434–5438.
- Lassar, A. B., Buskin, J. N., Lockshon, D., Davis, R. L., Apone, S., Hauschka, S. D., and H. Weintraub (1989). *Cell* **58**, 823–831.
- Sassoon, D., Lyons, G., Wright, W. E., Lin, V., Lassar, A., Weintraub, H., and M. Buckingham (1989). *Nature* **341**, 303–307.
- Hopwood, N. D., Pluck, A. and J. B. Gurdon (1989). *EMBO J.* **8**, 3409–3417.
- Tapscott, S. J., Davis, R. L., Thayer, M. J., Cheng, P.-F., Weintraub, H. and A. B. Lassar (1988). *Science* **242**, 405–411.

16. Murre, C., Schonleber McCaw, P. and D. Baltimore (1989). *Cell* **56**, 777–783.
17. Konieczny, S. F. and C. P. Emerson (1984). *Cell* **38**, 791–800.
18. Gospodarowicz, D., Weseman, J., Moran, J. S. and J. Lindstrom (1976). *J. Cell Biol.* **70**, 395–405.
19. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and W. J. Rutter (1979). *Biochemistry* **18**, 5294–5299.
20. Freeman, G.J., Clayberger, C., DeKruyff, R., Rosenblum, D.S. and H. Cantor (1983). *Proc. Nat. Acad. Sci. (USA)* **80**, 4094–4098.
21. Krug, M. S. and S. L. Berger (1987). *Meth. Enzymol.* **152**, 316–325.
22. Gubler, U. (1987). *Meth. Enzymol.* **152**, 330–335.
23. Saiki, R. K., Scarf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. and N. Arnheim (1985). *Science* **230**, 1350–1354.
24. Mullis, K.B. and Faloona, (1987). *Meth. Enzymol.* **155**, 335–350.
25. Feinberg, A. P. and Vogelstein, B. (1983). *Anal. Biochem.* **132**, 6–13.
26. Carlock, L., (1986). *Focus* **8**, 6–8.
27. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). *Proc. Nat. Acad. Sci. (USA)* **74**, 5463–5467.
28. Devereux, J., Haeblerli, P. and Smithies, O. (1984). *Nucleic Acids Res.* **12**, 387–395.
29. Cathala, G., Savouret, J.-F., Mendez, B., West, B. L., Karin, M., Martial, A. and Baxter, J. (1983). *DNA* **2**, 329–335.
30. Mahmoudi, M. and Lin, V. K. (1989). *Biotechniques* **7**, 331–333.
31. Dretzen, G., Bellard, M., Sassone-Corsi, P. and Chambon, P. (1981). *Anal. Biochem.* **112**, 295–298.
32. Harland, R. and Weintraub, H. (1985). *J. Cell Biol.* **101**, 1094–1099.
33. Gorman, C. (1985). *DNA cloning volume II. A practical approach.* ed., D. M. Glover. pp. 143–190. IRL Press (Oxford, U.K.).
34. Bader, D., Masaki, T. and Fischman, D.A. (1982). *J. Cell Biol.* **95**, 763–770.
35. Lasser, A.B., Paterson, B.M. and Weintraub, H. (1986). *Cell* **47**, 649–656.
36. Braun, T., Bober, E., Buschhausen–Denker, G., Kotz, S., Grzeschik, K-H. and Arnold, H. H. (1989). *EMBO J.*, **8**, 3617–3625.
37. Hanauer, A., Levin, M., Heilig, R., Daegelen, D., Kahn, A. and Mandel, J. L. (1983). *Nucleic. Acids Res.* **11**, 3503–3516.
38. Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H. and Kedes, L. (1983). *Mol. Cell. Biol.* **3**, 787–795.
39. Leader, D. P., Gall, I., Campbell, P. and Frischauf, A.-M. (1986). *DNA* **5**, 235–238.
40. Shaw, G. and Kamen, R. (1986). *Cell* **46**, 659–667.