The EMBO Journal vol.8 no.12 pp.3617-3625, 1989

# Differential expression of myogenic determination genes in muscle cells: possible autoactivation by the *Myf* gene products

**Thomas Braun, Eva Bober, Gregor Buschhausen-Denker, Stave Kotz[1], Karl-H.Grzeschik[2] and Hans Henning Arnold**

Department of Toxicology, Medical School, University of Hamburg, Grindelallee 117, 2000 Hamburg 13, FRG, [1]Mount Sinai School of Medicine, New York, NY, USA and [2]Department of Human Genetics, University of Marburg, Marburg, FRG

Communicated by M.Buckingham

**The development of muscle cells involves the action of myogenic determination factors. In this report, we show that human skeletal muscle tissue contains, besides the previously described Myf-5, two additional factors Myf-3 and Myf-4 which represent the human homologues of the rodent proteins MyoD1 and myogenin. The genes encoding Myf-3, Myf-4 and Myf-5 are located on human chromosomes 11, 1, and 12 respectively. Constitutive expression of a single factor is sufficient to convert mouse C3H 10T1/2 fibroblasts to phenotypically normal muscle cells. The myogenic conversion of 10T1/2 fibroblasts results in the activation of the endogenous *Myo*D1 and *Myf*-4 (myogenin) genes. This observation suggests that the expression of Myf proteins leads to positive autoregulation of the members of the *Myf* gene family. Individual myogenic colonies derived from MCA Cl15 cells (10T1/2 fibroblast transformed by methylcholanthrene) express various levels of endogenous MyoD1 mRNA ranging from nearly zero to high levels. The *Myf*-5 gene was generally not activated in 10T1/2 derived myogenic cell lines but was expressed in some MCA myoblasts. In primary human muscle cells Myf-3 and Myf-4 mRNA but very little Myf-5 mRNA is expressed. In mouse C2 and P2 muscle cell lines MyoD1 is abundantly synthesized together with myogenin. In contrast, the rat muscle lines L8 and L6 and the mouse BC3H1 cells express primarily myogenin and low levels of Myf-5 but no MyoD1. Myf-4 (myogenin) mRNA is present in all muscle cell lines at the onset of differentiation. From these differential patterns of *Myf* expression we conclude that, although in principle each Myf factor is capable of activating its own and related *Myf* genes, only a subset of myogenic factors is actually expressed in most muscle cells and this is sufficient to generate and maintain the differentiated phenotype.**

*Key words:* autoregulation/cell determination/human myogenic factors/regulatory network

## Introduction

Mouse C3H 10T1/2 embryonic fibroblasts can be converted into stable muscle cells by a brief treatment with 5-azacytidine (Taylor and Jones, 1979). This observation and the subsequent demonstration that the muscle phenotype can also be obtained by transfection of hypomethylated DNA (Konieczny *et al.*, 1985; Lassar *et al.*, 1986) into the same recipient cells, led to the hypothesis that one or few myogenic factor genes exist which upon activation cause the development of muscle cells. Indeed, a cDNA encoding the mouse myogenic factor MyoD1 was isolated from 10T1/2 myoblasts which had been derived by 5-azacytidine treatment (Davis *et al.*, 1987). This cDNA clone when expressed in non-muscle recipient cells was capable of generating myoblasts which differentiate into myobtubes under appropriate growth conditions (Weintraub *et al.*, 1989). Similarly, transfection of cloned human genomic DNA which was presumably unmethylated resulted in myogenic colonies due to the activation of the putative *myd* gene (Pinney *et al.*, 1988). More recently another cDNA, myogenin, was isolated by subtraction hybridization on the basis of its high level expression in bromodeoxyuridine resistant clones of the rat L6 muscle cell line and also in normal L6 myoblasts at the onset of cell fusion. (Wright *et al.*, 1989). We have recently described the isolation of the cDNA from human skeletal muscle encoding the novel myogenic determination factor Myf-5 (Braun *et al.*, 1989). This factor was detected by its moderate cross-hybridization to the mouse MyoD1 cDNA and was demonstrated to be capable of initiating the myogenic program in non-muscle cells of mesodermal origin. During that screening procedure two additional MyoD1 related cDNAs, Myf-3 and Myf-4, have been detected.

To investigate the nature of these human cDNAs and their encoded proteins in more detail and to explore their mutual relationship, we have determined the nucleotide sequences of Myf-3 and Myf-4 and their pattern of mRNA expression in human tissues and several myogenic cell lines. We report here that Myf-3 and Myf-4 represent the human homologues to MyoD1 and myogenin respectively. The *Myf* genes were found to be localized on three different human chromosomes. We present evidence for the potential autoregulatory capacity of myf proteins which are capable of activating transcription of their corresponding endogenous genes as well as the related members of the *myf* gene family. We furthermore demonstrate that established muscle cell lines generally do not co-express all three myf proteins. While myogenin (Myf-4) mRNA was found in all differentiating myocytes, MyoD1 mRNA is expressed in some but not in every cell line. Myf-5 specific mRNA was detected in rat L8 and L6 myoblasts which do not express MyoD1. A slightly larger mRNA which cross-hybridizes to the Myf-5 probe was found to be present in mouse BC3H1 cells and at low level also in C2 cells. In summary, our results suggest that the myogenic factors form an autoregulatory network but expression of a subset of myogenic factors may be sufficient to maintain the muscle phenotype.

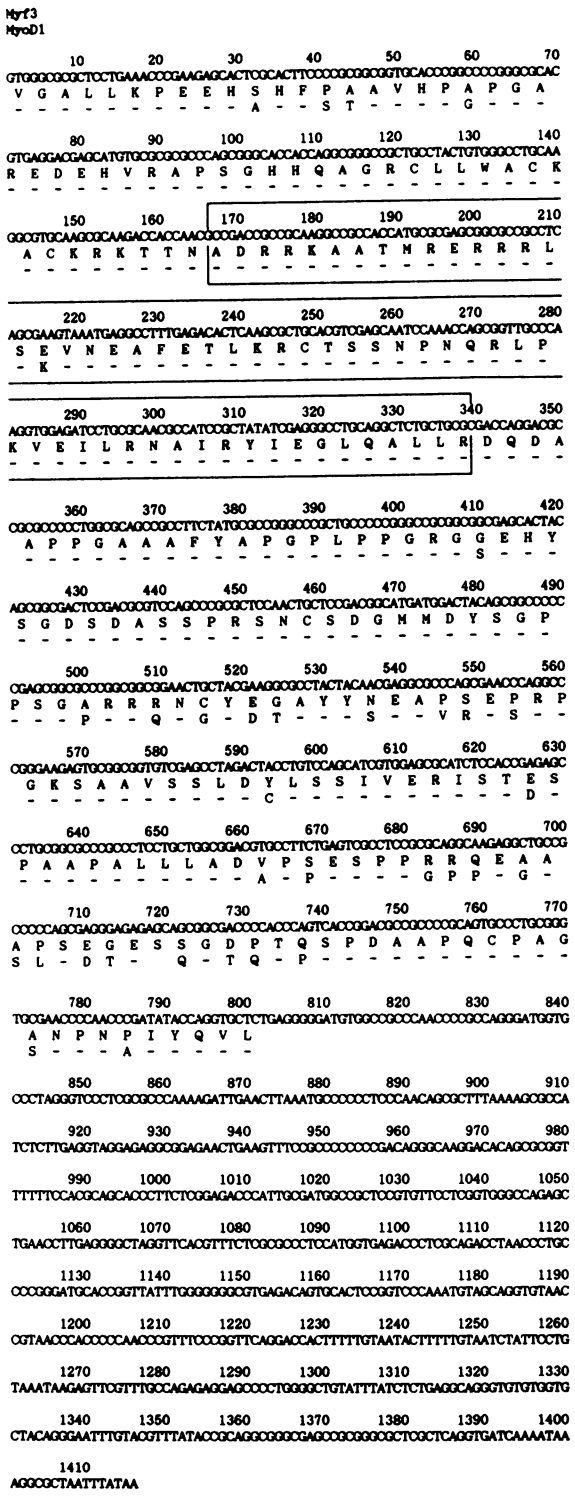


**Fig. 1.** Partial nucleotide sequence and the deduced amino acid sequence of the human cDNA clone Myf-3. The protein sequence of the mouse MyoD1 is shown for comparison (lower line). Identical amino acid residues are indicated by dashes. The highly conserved amphipathic helix-loop-helix motif is boxed in.

## Results

### *The Myf-3 cDNA represents the human homologue to mouse MyoD1*

When a cDNA library from human skeletal muscle was screened with the mouse MyoD1 probe, three types of recombinants were obtained. Besides the Myf-5 cDNA (Braun *et al.*, 1989), two additional distinct clones, Myf-3 and Myf-4, have been isolated by their moderate cross-hybridization to the MyoD1 probe. Although we were unsuccessful in isolating a full-length Myf-3 clone, we determined the partial sequence of ~1400 nucleotides, which exhibits a remarkable similarity to mouse *Myo*D1. This homology extends over the entire coding sequence and is not restricted to the helix-loop-helix domain or the myc-similarity region that has been identified in several developmental genes of *Drosophila*, e.g. the *achaete scute* complex (Villares and Cabrera, 1987), the *twist* gene (Thisse *et al.*, 1988) and the *daughterless* locus (Caudy *et al.*, 1988). More recently, a related sequence motif has also been found in the E12 and E47 proteins which bind to the immunoglobulin light chain gene enhancer (Murre *et al.*, 1989). As shown in Figure 1, ~90% of the deduced amino acid sequences of the available human and mouse proteins are identical. Even in the 3′ non-coding regions of both cDNAs, considerable conservation of the nucleotide sequences exists (not shown). When a full-length cDNA clone of *Myf*-3 (this clone was provided by C.Emerson and S.Pearson-White) was transfected into mouse C3H 10T1/2 fibroblasts and expressed under the control of the retroviral promoter (MSV-LTR), myogenic colonies were obtained which were indistinguishable from those observed with mouse *Myo*D1 (data not shown). The structural similarity together with the myogenic activity, and the chromosomal localization, which is presented below, suggest that Myf-3 constitutes the human homologue of the mouse myogenic factor MyoD1.

### *The Myf-4 cDNA encodes a protein homologous to rat myogenin*

The nucleotide sequence of the Myf-4 clone was also determined. As shown in Figure 2, Myf-4 contains ~1800 nucleotides, the sequence of which is clearly distinct from *Myf*-3 and *Myf*-5 except for the short helix-loop-helix homology region of ~180 nucleotides which is conserved between the three human cDNA isolates and is also shared by *Myo*D1 (Davis *et al.*, 1987). The most striking structural similarity, however, was observed when the Myf-4 sequence was compared with the recently published sequence of rat myogenin (Wright *et al.*, 1989). The deduced amino acid sequences for both proteins appear almost identical in the N-terminal half of the molecules with a slightly higher divergence in the C-terminal half. The overall nucleotide sequence conservation between rat and man is >85% including the 5′ and 3′ non-coding sequences. The structural similarity suggests that Myf-4 represents the human homologue of rat myogenin. This proposition was further supported by the myogenic capacity of Myf-4. As illustrated in Figure 3, when Myf-4 cDNA was constitutively expressed in C3H 10T1/2 recipient cells, multinucleated myotubes were observed at high frequency following the shift to low serum containing medium. These cells synthesize sarcomeric myosin as shown by immunostaining with antimyosin antibodies. Moreover, the converted cells express functional acetylcholine receptors identified by their binding of radiolabeled α-bungarotoxin (data not shown).

### *The expression of the three human myogenic factors is restricted to skeletal muscle tissues*

Having identified three distinct myogenic factors with moderate structural relatedness to one another and other

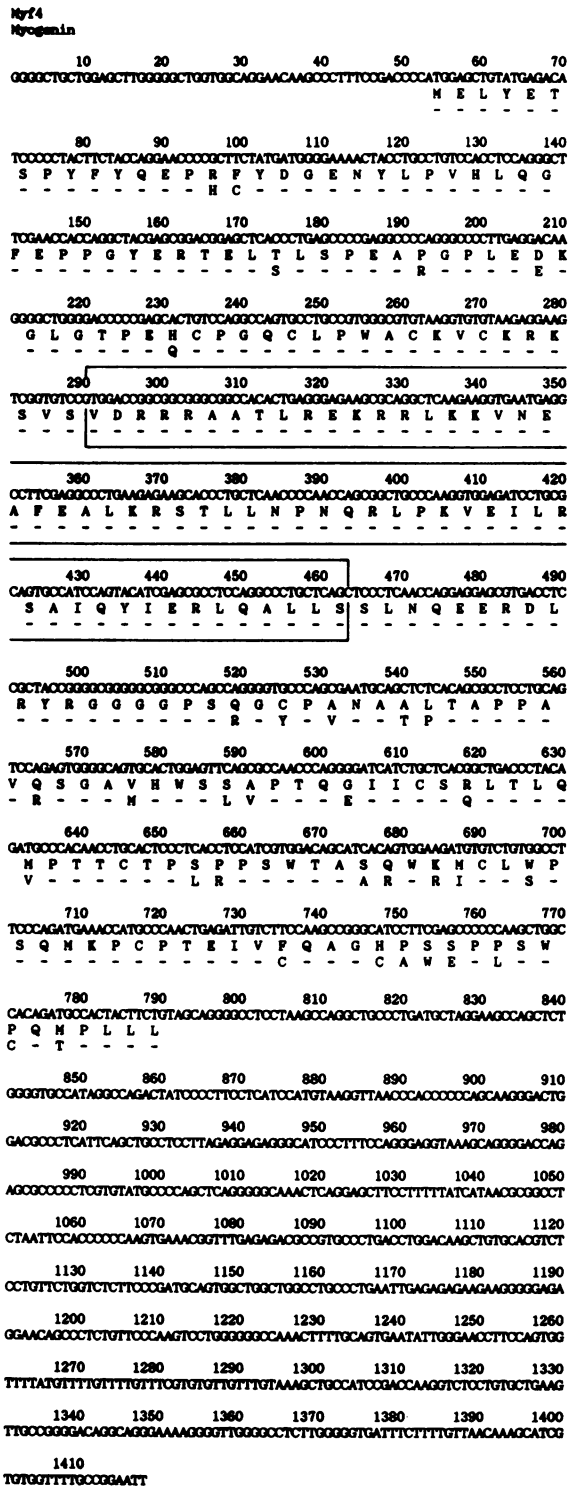


**Fig. 2.** Nucleotide sequence and the predicted amino acid sequence of the human Myf-4 cDNA supplemented by a 5′ genomic fragment. Nucleotides 1−170 were derived from the isolated Myf-4 gene (unpublished) and added to the cDNA sequence by an overlap of >100 nucleotides. The rat myogenin protein sequence is shown in the lower line for comparison. The C-terminal sequence of the rat myogenin exceeding the human sequence has been omitted. The box frames the Myf-homology motif.

developmentally important proteins, we analyzed the pattern of mRNA expression in various human tissues. As shown in Figure 4, the cDNA probes, specific for each Myf isolate, detected mRNAs of similar size in RNA preparations from

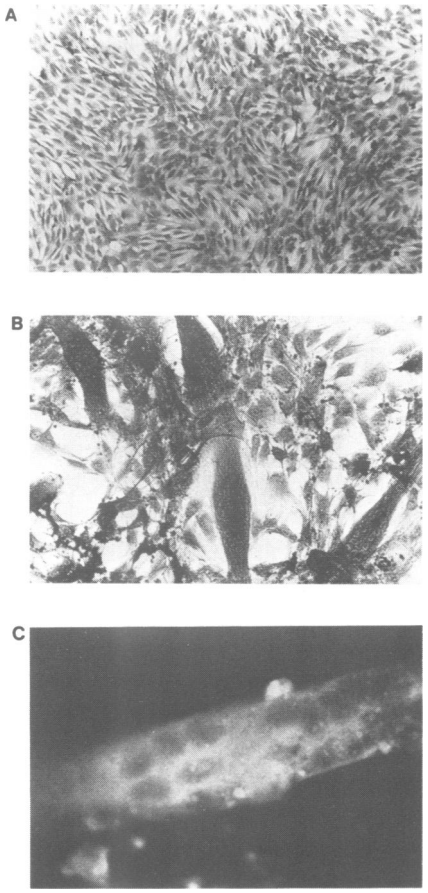


**Fig. 3.** Conversion of C3H 10T1/2 fibroblasts to myotubes following the expression of the human cDNA clone pEMSV-Myf4. (**A**) Phase contrast microscopy of normal C3H 10T1/2 mouse cells in differentiation medium. (**B**) Myotubes derived by the transfection of pEMSV-Myf4 into 10T1/2 fibroblasts 3 days after shift to differentiation medium. (**C**) C3H 10T1/2 derived myotube immunostained with anti-skeletal myosin and fluorescence coupled anti-IGG antibody.

several human skeletal muscle samples of fetal or adult tissues. No expression was found in heart muscle, non-muscle and smooth muscle tissues or in established culture cells of smooth muscle origin (HISM) (data not shown). The level of mRNA expression was extremely low for Myf-3, particularly in early fetal skeletal muscle (22 week old fetus), relative to the expression of the mRNAs coding for Myf-4 and Myf-5 which were moderately abundant at the same stage of development. The mRNA coding for Myf-5 seemed preferentially expressed in fetal skeletal muscle and dropped to lower levels in the adult muscle. In addition to the major transcript of ~1.8 kb, we have observed larger mRNA species hybridizing to the Myf-5 probe which might be due to unprocessed precursor molecules or might indicate that either different transcripts are generated from the Myf-5 gene or alternative splicing pathways exist.

### *The myogenic factor genes are located on different human chromosomes*

The chromosomal localization of the human *Myf* genes was determined by analyzing DNA blots from mouse−human and hamster−human somatic cell hybrids which had been previously shown to retain different subsets of human chromosomes (Balazs *et al.*, 1984). As demonstrated in

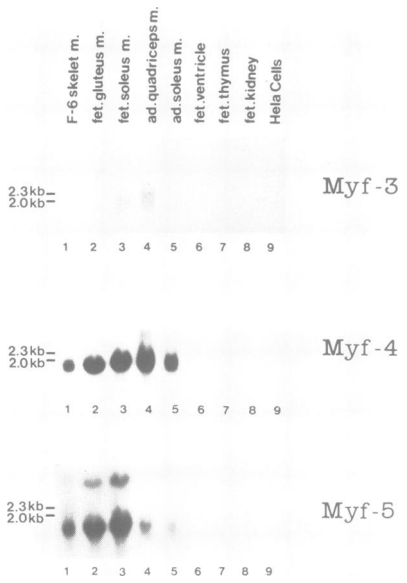


**Fig. 4.** Expression pattern for Myf-3, Myf-4 and Myf-5 mRNAs in various human muscle and non-muscle tissues. Approximately 40 $\mu$g of total RNA from the indicated tissues were applied per lane. The fetal tissues were prepared from two individual 22 week old human fetuses. F-6 skeletal muscle was taken from the upper leg. Adult muscle tissue was obtained from amputated leg muscle. Hybridizations were performed with $2 \times 10^6$ c.p.m./ml radioactively labelled probe (sp. act. $1 \times 10^8$ c.p.m./$\mu$g) overnight, the films were exposed for 24 h. The hybridization probes (see Materials and methods) do not cross-hybridize with each other under the applied conditions (final wash: 0.1 × SSC, 0.1% SDS at 65°C). The same RNA blot was reused in the indicated order after removing the hybridization signals of the prior probe.

Figure 5A, using specific cDNA probes for each of the three Myf proteins and choosing restriction fragments which allow the unambiguous distinction of the human versus the homologous rodent genes, we localized the *Myf* genes on three different human chromosomes. In human DNA digested with *Eco*RI restriction endonuclease, the *Myf*-3 gene is contained on a 7.5 kb restriction fragment, the *Myf*-4 gene is located on a 6.5 kb fragment and the *Myf*-5 gene resides on two *Eco*RI fragments of 3 and 3.5 kb. As summarized in Figure 5B, the diagnostic signal for the *Myf*-3 gene segregates concordantly with the human chromosome 11, the signal for the human *Myf*-4 gene co-segregates with chromosome 1 and the signals for the *Myf*-5 gene appear only in cells containing the entire or parts of human chromosome 12. This result indicates that the *Myf* gene family for the myogenic proteins is not clustered in the human genome but the gene loci are rather dispersed over several human chromosomes.

### *Muscle cell lines in culture express distinct subsets of myogenic factors*

To test the mRNA expression for Myf-3 (mouse MyoD1), Myf-4 (myogenin) and Myf-5 in human primary muscle cultures and in established rodent muscle cell lines, RNA blots from mouse C2 (Yaffe and Saxel, 1977a) BC3H1 (Schubert *et al.*, 1974), and P2 cells (Lassar *et al.*, 1986) as well as from rat L8 and L6 cells (Yaffe and Saxel, 1977b) were analyzed with cDNA hybridization probes specific for each myogenic factor. As demonstrated in Figure 6, various cell lines expressed different mRNA patterns for the three myogenic factors. While C2 and P2 cells accumulate very high levels of MyoD1 (Myf3) mRNA in myoblasts and even more so in myotubes, no MyoD1 mRNA was detectable in BC3H1, L6 and L8 myoblasts or myotubes. In contrast, Myf-4 (myogenin) mRNA was present in all muscle cell lines and generally appeared to be expressed after the cells had been cultured in differentiation supporting media. The RNA from P2 myoblast growing stage, shown in Figure 6, was derived from a particular experiment in which the cells had grown to complete confluence and therefore already expressed the Myf-4 mRNA. Low but detectable amounts of Myf-5 mRNA (1.8 kb) were found in rat L6 and L8 cells, at similar concentrations in growing myoblasts and differentiated myotubes. The expression of a Myf-5 related mRNA of ~2.2 kb in BC3H1 cells and at very low levels also in the original C2 line (Yaffe and Saxel, 1977a) was concluded from the hybridization signals obtained with either a genomic probe from the 5′ end of the mouse *Myf*-5 gene (unpublished data) or the complete human *Myf*-5 cDNA probe. The mRNA detected with these probes is slightly larger than the authentic Myf-5 mRNA detected in other rodent cells, e.g. L6, etc. The P2 myoblasts (provided by A.Lassar) failed to synthesize Myf-5 mRNA as did the original MCA Cl15 and 10T1/2 fibroblasts used as controls. As also shown in Figure 6B, primary human muscle cells in culture synthesize very low levels of Myf-5 mRNA, during 10 days after shift to differentiation medium. These cells contain Myf-3 and also Myf-4 mRNAs which both seem to slightly increase in concentration with the onset of differentiation. In summary, Myf-4 (myogenin) mRNA is always expressed in differentiated myocytes irrespective of the origin of cells or the procedure by which the cell lines had been established. MyoD1 mRNA is present at relatively high levels in some muscle cell lines which are known to differentiate efficiently to myotubes and also in primary muscle cells in culture. With the exception of myoblasts derived from the malignantly transformed MCA Cl15 cells and possibly a Myf-4 derived 10T1/2 cell (see Table I), we have not observed coexpression of MyoD1 and Myf-5 mRNAs in the same cell line. The BC3H1 cell which exhibits a striated muscle phenotype, but is incompetent of fusing to multinucleated syncythia, expresses myogenin mRNA and moderate levels of a Myf-5 related mRNA but no MyoD1.

### *Muscle conversion of fibroblasts by MyoD1, Myf-4 or Myf-5 leads to activation of the endogenous myogenic genes*

When C3H 10T1/2 fibroblasts or MCA Cl15 cells were transfected with plasmids expressing MyoD1, Myf-4 and Myf-5 from a viral promoter, the recipient cells not only morphologically converted to muscle cells but they also activated the expression of their endogenous myogenic genes. Using hybridization probes which specifically recognize mouse MyoD1, Myf-4 or myogenin, and the Myf-5 mRNAs, we analyzed RNAs from 10T1/2 and MCA Cl15 derived myocytes which had been converted by the constitutive expression of one of the myogenic factor cDNAs. As shown in Figure 7, 10T1/2 derived myogenic cells, converted either by the LTR driven expression of Myf-4 or Myf-5, activate transcription of the endogenous mouse *Myo*D1 gene. This is concluded from the appearance of the authentic MyoD1 mRNA of the correct size (1.8 kb) in all converted 10T1/2 cells. MyoD1 mRNA is absent in the original 10T1/2 and MCA Cl15 recipient cells shown as control (Figure 7, lanes 11 and 12). A similar result has been obtained for MyoD1 converted 10T1/2 myoblasts which also activate the

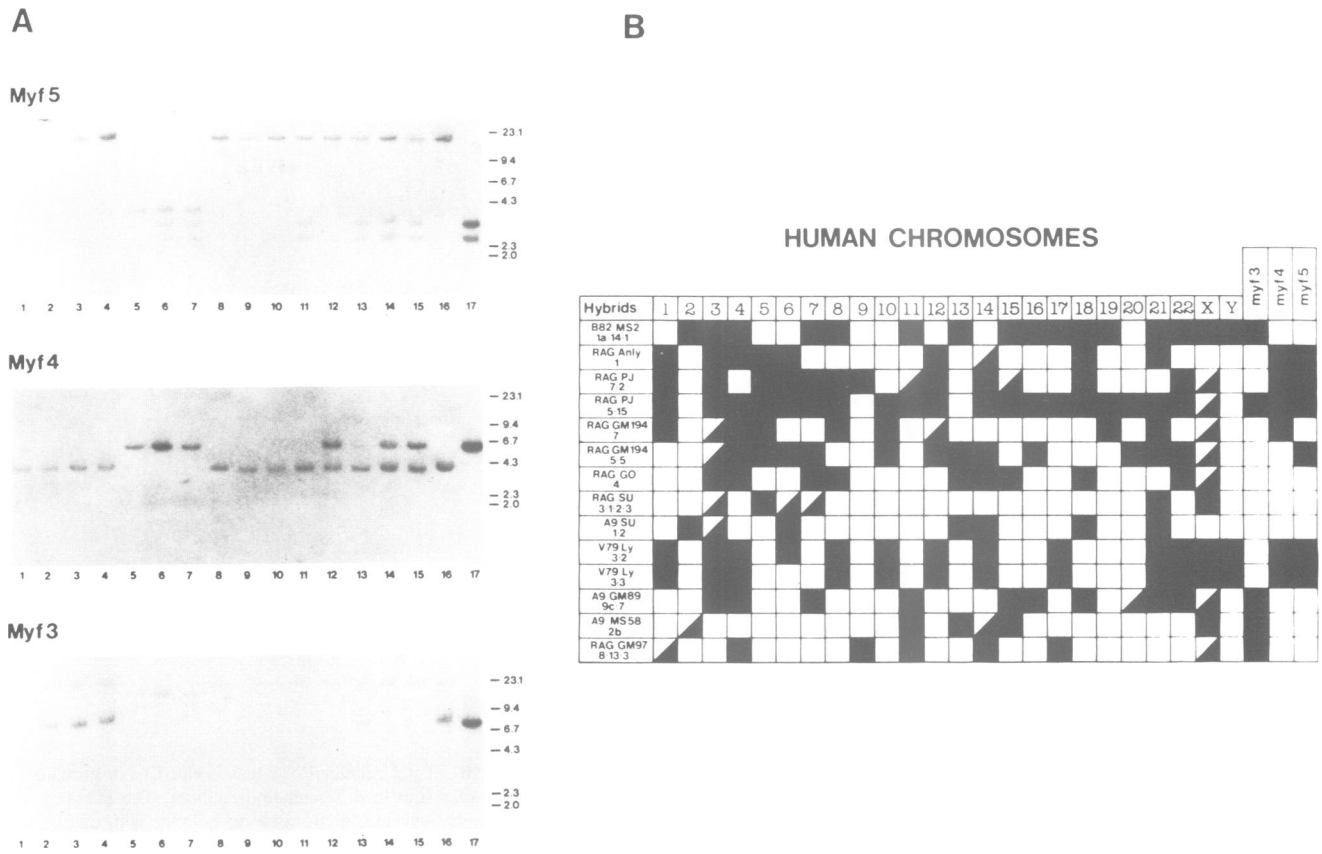


**Fig. 5.** Chromosomal assignment of the human *Myf* genes. (**A**) Southern blot analysis of *Eco*RI digested genomic DNA from mouse (**lane 1**), hamster (**lane 5**), mouse/human somatic hybrid cell lines (**lanes 2–4** and **8–16**), hamster/human hybrid cells (**lanes 6–7**) and human DNA (**lane 17**) hybridized with the indicated specific cDNA probes. For a detailed description of the somatic cell hybrids see Balazs *et al.* (1984). (**B**) Schematic representation of the localization of *Myf*-3, *Myf*-4 and *Myf*-5 on human chromosomes 11, 1 and 12 respectively. The listed cell lines (from top to bottom) correspond to the gel lanes 16–6 and 4–2 respectively. Black boxes indicate the presence of the complete human chromosome, (horizontal numbers), half-filled boxes show the presence of parts of the chromosome in the particular cell hybrid. Hybridization signals for the human *myf* alleles are also shown in black boxes.

endogenous *Myo*D1 gene (M.Thayer and A.Lassar, personal communication). In contrast, two myogenic clones, derived from MCA Cl15 cells by the LTR driven expression of MyoD1, synthesize large amounts of exogenous MyoD1 mRNA (~2.2 kb) generated from the expression vector but do not activate the endogenous *Myo*D1 gene. Since the exogenous MyoD1 mRNA is larger than the endogenous mRNA [due to the additional SV40 poly(A) addition sequence in the expression vehicle], both transcripts can clearly be discriminated on Northern blots. The probe used does not cross-hybridize to the human Myf-4 or Myf-5 sequences (Figure 7, lane 13). Confirmation for the activation of the endogenous myogenic genes came from independent experiments shown in Figure 8. As has been noticed before (see Figure 7), the mouse *Myo*D1 gene was activated in the 10T1/2-Myf 4 converted clone and its expression appeared up-regulated in differentiated myotubes. This myogenic clone also synthesized mouse Myf-4 (myogenin) mRNA at increasing concentrations with differentiation. The MCA Cl15 derived myogenic colonies analyzed in this experiment also activated their endogenous *Myo*D1 gene but to very different degrees. Whereas the MCA-Myf 4 clone expressed very low levels of MyoD1 mRNA, two independently isolated MCA-Myf 5 clones transcribed high levels of MyoD1 RNA which is in contrast to a previously studied MCA-Myf 5 isolate (Braun *et al.*, 1989). This observation indicated that there is considerable clonal variation of MCA Cl15 derived myoblasts with respect to the activation of the *Myo*D1 gene. By comparison the activation of the endogenous Myf-4 gene was approximately equal in all MCA Cl15 derived myoblast clones. The level of mouse Myf-4 mRNA could be specifically determined even in cells expressing exogenous human Myf-4, since the transfected expression construct lacks the 5′ sequences (*Eco*RI/*Sst*I) which were used as hybridization probes (see Materials and methods). Likewise, the expression of the mouse Myf-5 mRNA could be specifically measured in the clones 10T1/2-Myf5/1 and MCA-Myf5/1 which were derived by expressing the human cDNA Myf5/18 lacking the 5′ sequences up to the first *Pst*I restriction site (see Materials and methods). Using this Myf-5 sequence as hybridization probe, activation of the endogenous mouse Myf-5 gene was observed in MCA-Myf5/1 cells but not in the analogous clone derived from 10T1/2 cells. The activation of endogenous myogenic factor genes by the expression of exogenous determination factors is summarized in Table I.

## Discussion

### *Apparent redundancy of myogenic determinaton genes in humans*

The work presented here shows that at least three distinct myogenic factors are expressed in human striated muscle

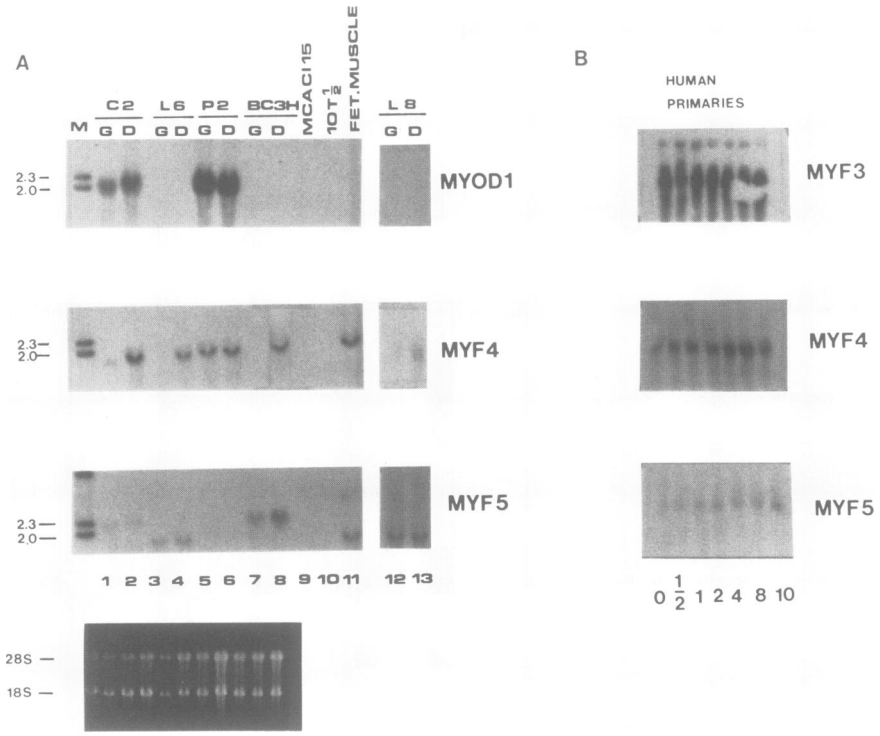


**Fig. 6.** Northern analysis of Myf mRNAs in the cell lines (**A**) and human primary muscle cells (**B**). Total cellular RNA was isolated from growing (G) and differentiated (D) cell lines and from human primary muscle cells at the indicated days after shift to differentiation medium. The filters containing RNA (50 μg) from cell lines (L8 cells were analyzed on separate blot) were consecutively hybridized (from top to bottom) with the radioactive probe for mouse MyoD1, the 5′ specific Myf-4 probe (lacking the homology sequence) and the mouse Myf-5 probe (see Materials and methods). The bottom panel shows the ethidium bromide staining of a parallel gel to confirm that equivalent amounts of RNA were loaded on each lane. Note, the larger RNA signal (~2.2 kb) detected with the mouse Myf-5 probe in C2 and BC3H1 cells. (**B**) Three parallel filters containing RNA (25 μg) of human primary muscle cells grown in full-medium (0) and 0.5–10 days after transfer to differentiation medium were hybridized to Myf-3, Myf-4 and human Myf-5 probes (top to bottom). The mRNAs detected with the human probes are approximately of the same length. The filters probed with Myf-3 and Myf-4 were exposed on film for 24 h, the filter probed with Myf-5 was exposed for 72 h.

tissues at fetal and later stages of development but not in heart and smooth muscle. We have re-examined the expression of Myf mRNAs in recently obtained samples from human uterus tissue and human intestinal smooth muscle cells (HISM) and failed to confirm our previous observation of Myf transcripts in smooth muscle (Braun *et al.*, 1989). We attribute this discrepancy to the nature of the original specimen which might not have been exclusively smooth muscle. The protein factors Myf-3, Myf-4 and Myf-5 are structurally related to each other owing to a highly conserved amphipathic helix-loop-helix domain which is also shared by other regulatory proteins from *Drosophila* (Villares and Cabrera, 1987; Caudy *et al.*, 1988; Thisse *et al.*, 1988), human B cells (Murre *et al.*, 1989) and the myc gene family (Battey *et al.*, 1983; Alt *et al.*, 1986; DePinho *et al.*, 1987). Detailed sequence comparison and the functional tests reveal that Myf-3 constitutes the human homologue of the mouse MyoD1 (Davis *et al.*, 1987). Myf-4 represents the human counterpart of myogenin (Edmondson and Olson, 1989; Wright *et al.*, 1989). All three human cDNAs when individually expressed from a constitutive promoter are indiscriminately capable of converting 10T1/2 and MCA C115 fibroblasts into muscle cells. The muscle phenotype obtained appears indistinguishable regardless which of the three factor cDNAs has been transfected into the cells. The genes encoding the Myf proteins are located on three different human chromosomes and therefore do not constitute a single genetic locus with one or a few linked myogenic genes as previously postulated based on DNA transfection experiments (Konieczny and Emerson, 1984; Lassar *et al.*, 1986). They rather represent a set of genes which seem structurally and functionally equivalent according to the '*in vitro*' myogenic conversion assay. The human *Myo*D1 gene had previously been located on human chromosome 11 using the heterologous mouse cDNA probe (Tapscott *et al.*, 1988). We confirm this chromosomal assignment for *Myf*-3 with the homologous human probe. Fine mapping has shown that *Myo*D1 is located on the small arm of chromosome 11, distal to 11p13 and probably close to 11p15 (M.Gessler and H.H.Arnold, unpublished results). The existence of several *Myf* genes in the human genome is reminiscent of the *achaete scute* complex in *Drosophila* which is also comprised of several homologous genes, which are, however, located in a closely linked gene cluster (Villares and Cabrera, 1987; Alonso and Cabrera, 1988).

### *The muscle phenotype does not require the simultaneous expression of all three determination factors*

The pattern of Myf expression, analyzed on the level of specific mRNAs in several established muscle cell lines reveals that generally only a subset of one or two Myf mRNAs accumulates. As demonstrated for the mouse C2 and P2 cells and also found for the MM14 line (S.Hauschka, personal communication), several mouse muscle cells apparently express the *Myo*D1 gene together with the

**Table I.**

| Myogenic cell clones | Activation of endogenous myogenic factor genes | | |
|---|---|---|---|
| | MyoD1[e] | Myf4[f] | Myf5 |
| 10T1/2-Myf4[a] | + | + | (+) |
| 10T1/2-Myf5/1[b] | + | + | –[g] |
| 10T1/2-Myf5/2[c] | + | + | – |
| MCA1-MyoD1[d] | – | + | – |
| MCA3-MyoD1[d] | – | + | – |
| MCA-Myf4[a] | (+) | + | – |
| MCA-Myf5/1[b] | + | + | +[g] |
| MCA-Myf5/2[c] | + | + | n.d. |

[a]Conversion with pEMSV-Myf4/*Sst*.
[b]Conversion with pEMSV-Myf5/18.
[c]Conversion with pEMSV-Myf5.
[d]Conversion with pEMSV-MyoD1.
[e]3′ Noncoding region of mouse MyoD1 cDNA (*Hpa*II–*Eco*RI) was used as probe.
[f]5′ Sequence (*Nco*I–*Sst*I) of human *Myf*-4 gene was used as probe. This 115 bp fragment is diagnostic for the endogenous mouse gene in all cells studied since it is not present in the pEMSV-Myf-5/*Sst* expression vehicle.
[g]5′ Sequence (*Eco*RI–*Pst*I) of Myf-5 cDNA was used as probe. This fragment is diagnostic for the endogenous *Myf*-5 gene in pEMSV-Myf5/18 derived cells.
+, Abundant or moderate mRNA levels; (+) barely detectable mRNA levels and n.d. was not determined.

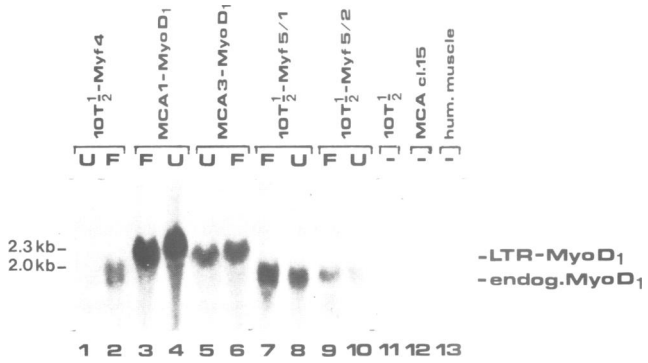


**Fig. 7.** Expression of endogenous and LTR driven (transfected) mouse MyoD1 mRNA in 10T1/2 and MCA Cl15 derived myoblasts converted by different myogenic factor cDNAs. Total cellular RNA (40 μg) isolated from unfused (U) or fused (F) cell lines was analyzed on Northern blots as indicated. The blot was probed with the mouse MyoD1 cDNA fragment which lacks the basic and Myc homology region (see Materials and methods). Note that the RNA expressed from the pEMSV-MyoD1 expression vector is clearly larger than the endogenous MyoD1 mRNA due to additional SV40 poly(A) addition sequences. MCA1-MyoD1 and MCA3-MyoD1 represent independently derived clones of the same transfection experiment. 10T1/2-Myf5/1 was obtained by expression of pEMSV-Myf5/18. 10T1/2-Myf5/2 was generated by expression of pEMSV-Myf5.

myogenin gene. In contrast, the human primary muscle cells which might represent a mixture of different myocytes synthesize Myf-3 and Myf-4 plus some Myf-5 mRNA. The BC3H1 line, derived from a mouse brain tumor and exhibiting markers characteristic of striated muscle, synthesizes probably a Myf-5 related protein, since it contains a cross-hybridizing mRNA species and the cells can be specifically stained with Myf-5 antiserum (not shown). It does not contain MyoD1. Similarly, the rat muscle lines L8 and L6, obtained from the ATCC, contain myogenin and low levels of Myf-5 mRNA but do not show detectable amounts of MyoD1. A subclone of L6, however, which was selected for efficient myotube formation, expresses appreciable amounts of Myf-5 mRNA in addition to myogenin (W.Wright, personal communication). Since BC3H1 cells are incompetent at fusing, L8 and L6 cells are very slow and inefficient in forming myotubes and might express only a partial myogenic program (Whalen *et al.*, 1978; Minty *et al.*, 1986), whereas mouse C2, P2 and MM 14 cells fuse very efficiently in low serum, one might speculate that the expression of MyoD1 and myogenin which correlates with the well differentiating phenotype might support a more complete myogenic program than the expression of myogenin alone or in conjunction with Myf-5.

Alternatively, the observation of the differential expression of the myogenic determination genes might suggest that distinct muscle subtypes, e.g. slow or fast fibers, or different muscle cell types, e.g. satellite cells or primary myoblasts, could be determined by the expression of one or the other *Myf* gene. In fact, we have not yet detected a normal muscle cell line which co-expresses MyoD1 and Myf-5. (The nature of the Myf-5 related mRNA in C2 cells must await further clarification.) This observation would be compatible with the idea that MyoD1 and Myf-5 might mark distinct myogenic cell lineages which are reflected by the permanent lines. The fact that myogenin has been found to be synthesized in all muscle cell lines might indicate that it is either absolutely required for the generation or maintenance of the muscle phenotype or that it is activated in development prior to the expression of the other *Myf* genes and therefore present in most established cell lines. In this context it is interesting to note that during the formation of the myotome in mouse development, myogenin expression actually precedes the expression of MyoD1 (Sassoon *et al.*, 1989). In summary, our results indicate that probably one myogenic factor (myogenin) is sufficient to maintain at least some aspects of the muscle phenotype in cultured cells but generally the complete muscle cell contains more than one factor.

### *Expression of myogenic factor cDNAs leads to activation of endogenous muscle determinaton genes*

Transfection and expression of Myf-4 and Myf-5 cDNAs results in the activation of the endogenous *Myo*D1 and *myogenin* genes in 10T1/2 and MCA Cl15 fibroblasts. Similarly expression of Myf-3 or mouse MyoD1 cDNA causes activation of the mouse *Myo*D1 and *myogenin* genes in 10T1/2 cell fibroblasts (unpublished results and personal communication by M.Thayer). Using MCA Cl15 cells as recipients, we have previously observed myogenic colonies which did not activate their endogenous *Myo*D1 gene to an appreciable degree (Braun *et al.*, 1989). We have now independently obtained MCA-Myf5 colonies which express high levels of mouse MyoD1 mRNA in addition to mouse Myf-5 mRNA and myogenin. It has previously been shown that the expression of the activated H-*ras* oncogene will prevent terminal differentiation of muscle cells (Olson *et al.*, 1987). Recently it has been documented that methyl-cholanthrene transformed 10T1/2 fibroblasts, e.g. the MCA Cl15 cells, carry a mutated H-*ras* allele which when expressed is probably associated with the transformed phenotype of these cells (Chen and Hershman, 1989). The fact that MCA Cl15 cells can be converted to muscle cells

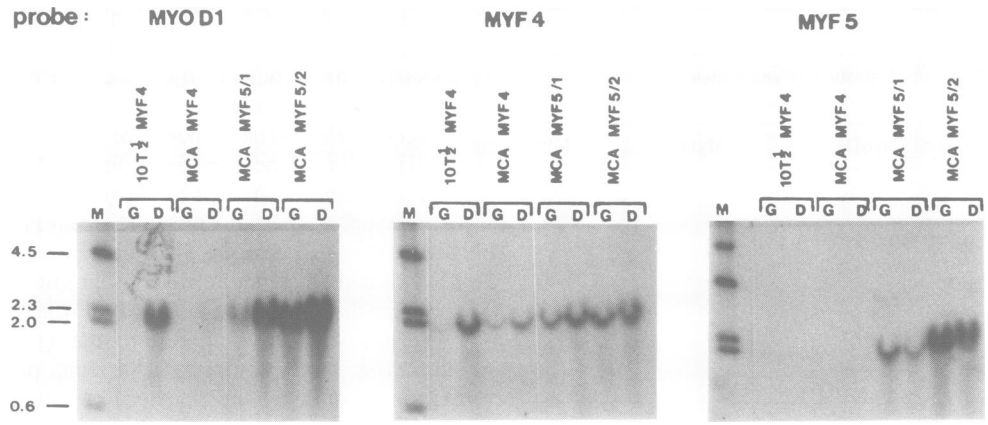


**Fig. 8.** Activation of endogenous Myf mRNA expression in various myogenic clones. Total RNAs (40 μg) of the indicated myogenic cell clones from growing (G) or differentiated cells (D) were analyzed by the Northern blot technique and consecutively hybridized with the mouse specific MyoD1 probe (left), the 5′ end Myf-4 (*Nco*I−*Sst*I) probe (middle) and the 5′ end Myf-5 (*Eco*RI−*Pst*I) probe (right) which are each able to detect the mouse myogenic factor mRNAs and distinguish between the transfected and the endogenous RNA species. Except for MCA-Myf-5/2 which was obtained by expression of the total Myf-5 cDNA (pEMSV-Myf-5 generates slightly larger transcript), all myogenic clones shown in this experiment were derived from N-terminally truncated Myf cDNAs (pEMSV-Myf-4/*Sst*; pEMSV-Myf-5/18).

by constitutive expression of any of the known myogenic factors and the observation that this conversion is associated with varying levels of endogenous gene expression suggests that a delicate balance of activated ras protein and exogenous myogenic factors may be determining the state of the endogenous MyoD1 gene. While these results taken together suggest that all Myf proteins are potentially able to activate their homologous mouse counterparts and the other members of the *myf* gene family as well, they also show that the activation is further controlled by other cellular factors of the recipient cells. Whether the expression of myogenic determination genes is the result of a direct interaction of the gene products and regulatory gene sequences is currently obscure, although preliminary results suggest that the upstream regions of Myf genes can bind to Myf-5 protein *in vitro* (unpublished observation).

### *The potential relationship between myogenic regulatory genes*

It is now clear that at least three distinct myogenic regulatory genes exist which belong to a family of potential DNA binding proteins (Tapscott *et al.*, 1988). In addition, the presence of the human myd gene has been proposed as a result of DNA transfections leading to the muscle phenotype (Pinney *et al.*, 1988). Southern blot analysis of myd transfectants (the DNA blot was given to us by C.Emerson) with all three Myf probes indicates that myd is a genetic locus which is distinct from the *Myf* genes and probably constitutes a different class of genes, since no obvious cross-hybridization of the Myf homology region was observed (unpublished results).

The hypothetical hierarchy among the Myf proteins remains enigmatic. While it seems evident that the expression of the factors is differentially regulated in cell culture and *in vivo* (Sassoon *et al.*, 1989), no obvious pattern emerges which would allow us to recognize a cascade of events. The autoregulatory activation of *Myf* genes by constitutively expressed factor cDNAs only argues for a possible role in maintaining the muscle phenotype but gives no information on the initial order of activation with respect to the individual *Myf* genes. The expression pattern of *Myf*-4 (myogenin) is particularly interesting since it seems to be the only gene which is expressed in all skeletal muscle type cells. In addition its expression seems to be inhibited in proliferating myoblasts suggesting the existence of a negative control acting prior to the onset of fusion.

## Materials and methods

### *Isolation of cDNAs and determination of nucleotide sequences*

The cDNA clones Myf-3 and Myf-4 were isolated from a human fetal skeletal muscle λGT11 library as previously described (Braun *et al.*, 1989). The cDNA inserts were released from the recombinants with *Eco*RI restriction endonuclease and subcloned into mp18 and mp19M13 vectors for dideoxy sequencing (Sanger *et al.*, 1977). Appropriate restriction subfragments or unidirectional *Exo*III deletions (Henikoff, 1984) were also cloned and sequenced. The final nucleotide sequences were established on both DNA strands for the majority of subfragments. The Myf-4 sequence up to the *Sst*I restriction site located at nucleotide 170 was generated from the genomic fragment which we have isolated from a human genomic lambda library (unpublished results).

### *Expression of Myf cDNAs*

To construct the expression vehicles pEMSV-Myf4/Sst, pEMSV-Myf5 and pEMSV-Myf5/18, cDNA inserts of the original clones Myf-4 and Myf-5 (Braun *et al.*, 1989) and of a shorter cDNA version of Myf-5 (Myf5/18) were cloned into the *Eco*RI site of the pEMSV α-scribe vector (kindly supplied by A.Lassar) to allow transcription in eukaryotic cells. Specifically, the cDNA of Myf-4, starting at the *Sst*I site was joined with its 5′ end to an *Nco*I linker to supply an AUG translational start signal in frame. Likewise, Myf5/18 starting with nucleotide 185 was supplemented with an *Nco*I linker supplying the start signal. Each modified cDNA insert was subsequently inserted via *Eco*RI linkers into pEMSV α-scribe. The pEMSV-Myf5 construct has been described previously (Braun *et al.*, 1989). Alternatively, the full length Myf-4 expression vector pEMSV-Myf4 was constructed by fusing a 5′ genomic fragment, covering the 5′ leader sequence up to the *Sst*I site in exon 1, to the same site in the Myf-4 cDNA. The resulting construct was then cloned into pEMSVscribe via *Eco*RI linkers. The full length cDNA insert of a human MyoD1 clone, which was provided to us by C.Emerson, was also inserted into the *Eco*RI site of pEMSVscribe in order to express Myf-3.

### *Cell culture and DNA transfection procedure*

Mouse C2 (Yaffe and Saxel, 1977a), BC3H1 cells (Schubert *et al.*, 1974), rat L8 and L6 cells (Yaffe and Saxel, 1977b) were obtained from the American Type Culture Collection (ATCC). MCA C115 and C3H 10T1/2 cells were provided by H.Marquardt, Hamburg and P2 aza-myoblasts were kindly supplied by A.Lassar, Seattle, USA. Cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS) or

according to the procedures recommended by the ATCC. To initiate differentiation C2, P2 and 10T1/2 or MCA derived myoblasts were transferred to DMEM with 2% horse serum, BC3H1 cells were shifted to DMEM with 0.5% FCS, and L6 cells were exposed to DMEM with 10% horse serum and 5 μg/ml insulin (differentiation media). To generate stable myogenic clones from 10T1/2 or MCA Cl15 cells, $1-5 \times 10^5$ cells/10 cm diameter dish were co-transfected with 1 μg of supercoiled pSV2-neo plasmid and either 30 μg of pEMSV-Myf5 (Braun *et al.*, 1989) or 30 μg of pEMSV-Myf4 or the truncated versions of the expression clones (see above) using the standard calcium phosphate precipitation method of Graham and von der Eb (1973) and as described previously (Braun *et al.*, 1989). G418 resistant colonies were selected in medium containing 400 μg/ml G418 (Geneticin, Gibco, USA) and myogenic conversion was scored by the appearance of multinucleated myofibres 3–6 days after shift to the low serum containing differentiation medium (DMEM plus 2% horse serum). Myogenic colonies were picked as multi-clones and recloned twice to obtain pure muscle cell lines.

### *Immunostaining*

For the immunostaining with anti-skeletal myosin rabbit serum (Bio Yeda, Rehovot, Israel), cell cultures were fixed in 70% ethanol, 3.7% formaldehyde and 5% acetic acid for 5 min and subsequently washed extensively in phosphate buffered saline (PBS). Cultures were then incubated overnight in 1:20 diluted antimyosin serum at 4°C. Unbound antibody was removed by washing with PBS and cells were incubated with fluorescence coupled goat anti-rabbit serum for 30 min at 37°C. After extensive washing to remove excess of second antibody, cells were analyzed under the fluorescence microscope.

### *Northern blot analysis of RNA from human tissues and cell cultures*

RNA was isolated from frozen human tissues by the LiCl–urea method (Auffray and Rougeon, 1980) and from cultured cells by the guanidinium method (Chomczynski and Sacchi, 1987). Usually $5-10 \times 10^6$ cells were used, yielding between 200 and 300 μg of total RNA 25–50 μg of total RNA were denatured by glyoxylation, separated on agarose gels and transferred to PALL membrane (Biodyne) as described by Thomas (1980). Gel purified DNA fragments were radioactively labeled to a sp. act. of $1-2 \times 10^8$ c.p.m./μg with [$^{32}$P]dCTP (3000 Ci/mmol) by random priming using the kit system from Amersham. Filter hybridizations were carried out in 50% formamide, 5 × Denhardt's solution, 5 × SSC (1 × SSC contains 0.15 M sodium chloride and 0.015 M sodium citrate), 50 mM Na phosphate, 0.1% sodium dodecylsulfate (SDS) and 200 μg/ml denatured herring sperm DNA at 42°C for 18 h followed by washing steps in 2 × SSC, 0.1% SDS at 55°C and 0.1 × SSC, 0.1% SDS at 65°C for 30 min each. Filters were exposed to X-ray film for 24–72 h at −80°C using intensifying screen.

The following probes were used for hybridizations. (i) The mouse MyoD1 fragment was cleaved from the cDNA (Davis *et al.*, 1987) with *Hpa*II and *Eco*RI restriction endonucleases and gel purified. It represents a 800 nucleotide fragment encoding the 3′ end sequence. (ii) the human probes Myf-3, Myf-4 and Myf-5, unless otherwise specified, constitute the total cDNAs as previously published (Braun *et al.*, 1989). They coincide with the following sequences described in this report: Myf-3, nucleotides 595–1415; Myf-4, nucleotides 170 (*Sst*I site)–1420; and Myf-5, nucleotides 185–1430. (iii) A 5′ specific Myf-4 probe, was generated by *Nco*I/*Sst*I digestion of the Myf-4 gene resulting in a fragment of 115 bp. (iv) A 5′ specific Myf-5 probe was obtained by *Eco*RI–*Pst*I digestion of the Myf-5 cDNA yielding a fragment of 180 nucleotides. (v) A mouse Myf-5 probe was generated by *Ball*–*Pst*I digestion of the mouse genomic sequence (unpublished) yielding a 280 nucleotide fragment which contains the 5′ leader sequence plus 100 nucleotides of the N-terminal coding sequence.

## Acknowledgements

We are grateful to C.Emerson and S.Pearson-White for providing DNA blots of myd transfected clones and the full length cDNA encoding the human MyoD1 which they isolated independently. We thank A.Lassar for the MyoD1 plasmid and the pEMSVscribe expression vector, A.Lassar, W.Wright and E.Olson for helpful discussions and information prior to publication. We are particularly grateful to M.Buckingham for a detailed discussion of the manuscript and many helpful suggestions for improvement. We acknowledge the technical assistance of H.Eberhardt and D.Nollmeyer, and thank A.Broecker for relentlessly typing the manuscript and K.D.Soehren for the artwork. The work was supported by Deutsche Forschungsgemeinschaft and Deutsche Muskelschwundhilfe.

## References

Alonso,M.C. and Cabrera,C.V. (1988) *EMBO J.*, **7**, 2585–2591.
Alt,F.W., DePinho,R.A., Zimmerman,K., Legony,E., Hutton,K., Ferrier,P., Tesfaye,A., Yancopoulos,G.D. and Nisen,P. (1986) *Cold Spring Harbor Symp. Quant. Biol.*, **51**, 931–941.
Auffray,C. and Rougeon,F. (1980) *Eur. J. Biochem.*, **107**, 303–313.
Balazs,J., Purello,M., Aldaheff,B., Grzeschik,K.W. and Szabo,P. (1984) *Hum. Genet.*, **68**, 57–61.
Battey,J., Moulding,C., Taub,R., Murphy,W., Stewart,T., Potter,M., Lenoir,G. and Leder,P. (1983) *Cell*, **34**, 779–787.
Braun,T., Buschhausen-Denker,G., Bober,E., Tannich,E. and Arnold,H.H. (1989) *EMBO J.*, **8**, 701–709.
Caudy,M., Vassin,H., Brand,M., Tuma,R., Jan,L.Y. and Yan,Y.N. (1988) *Cell*, **55**, 1061–1067.
Chen,A.C. and Hershman,H.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 1608–1611.
Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
Davis,R.L., Weintraub,H. and Lassar,A.B. (1987) *Cell*, **51**, 987–1000.
DePinho,R.A., Hutton,K.S., Tesfaye,A., Yancopoulos,G.D. and Alt,F.W. (1987) *Genes Dev.*, **1**, 1311–1326.
Edmondson,D.G. and Olson,E.N. (1989) *Genes Dev.*, **3**, 628–640.
Graham,F.L. and von der Eb,A.J. (1973) *Virology*, **52**, 456–467.
Henikoff,S. (1984) *Gene*, **28**, 351–359.
Konieczny,S.F. and Emerson,C.P.J. (1984) *Cell*, **38**, 791–800.
Konieczny,S.F., Baldwin,A.S. and Emerson,C.P.J. (1985) In Emerson,C., Fishman,D.A., Nadal-Ginard,B. and Siddiqui,M.A.Q. (eds) *Molecular Biology of Muscle Development*. UCLA Symposium on Molecular and Cellular Biology, New Series, Vol. 29, Alan R.Liss, New York, pp.21–34.
Lassar,A.B., Paterson,B.M. and Weintraub,H. (1986) *Cell*, **52**, 179–184.
Minty,A., Blau,H. and Kedes,L. (1986) *Mol. Cell. Biol.*, **6**, 2137–2148.
Murre,C., McCaw,P.S. and Baltimore,D. (1989) *Cell*, **56**, 777–783.
Olson,E.N., Spizz,G. and Tainsky,M.A. (1987) *Mol. Cell. Biol.*, **7**, 2104–2111.
Pinney,D.F., Pearson-White,S.M., Konieczny,S.F., Latham,K.E. and Emerson,C.P.Jr (1988) *Cell*, **53**, 781–793.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Sassoon,D., Lyons,G., Wright,W.E., Lin,V., Lassar,A., Weintraub,H. and Buckingham,M. (1989) *Nature*, **341**, 303–307.
Schubert,D., Harris,A.J., Devine,C.E. and Heinemann,S. (1974) *J. Cell. Biol.*, **61**, 398–413.
Tapscott,S.J., Davis,R.L., Thayer,M.J., Cheng,P.F., Weintraub,H. and Lassar,A.B. (1988) *Science*, **242**, 405–411.
Taylor,S.M. and Jones,P.A. (1979) *Cell*, **17**, 771–779.
Thisse,B., Stoetzel,C., Gorostiza-Thisse,C. and Perrin-Schmitt,F. (1988) *EMBO J.*, **7**, 2175–2185.
Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
Villares,R. and Cabrera,C.V. (1987) *Cell*, **50**, 415–424.
Weintraub,H., Tapscott,S., Davis,R., Thayer,M., Adam,M., Lassar,A. and Miller,A.D. (1989) *Proc. Natl. Acad. Sci. USA*, in press.
Whalen,R.G., Butler-Brown,G.S. and Gros,F. (1978) *J. Mol. Biol.*, **126**, 415–431.
Wright,W.E., Sassoon,D.A. and Lin,V.K. (1989) *Cell*, **56**, 607–617.
Yaffe,D. and Saxel,O. (1977a) *Nature*, **270**, 725–727.
Yaffe,D. and Saxel,O. (1977b) *Differentiation*, **7**, 159–166.

*Received on August 1, 1989*