The EMBO Journal vol.9 no.3 pp.821 - 831, 1990

Myf-6, a new member of the human gene family of myogenic determination factors: evidence for a gene cluster on chromosome 12

T.Braun, E.Bober, B.Winter, N.Rosenthal[1] and H.H.Arnold

Department of Toxicology, University of Hamburg, Medical School, Grindelallee 117, 2000 Hamburg 13, FRG

[1]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118-2394, USA

Communicated by M.Buckingham

The *Myf-6* gene, a novel member of the human gene family of muscle determination factors has been detected by its highly conserved sequence coding for a putative helix-loop-helix domain. This sequence motif is a common feature of all Myf factors and other regulatory proteins. The new *Myf* gene is located on human chromosome 12, ~ 6.5 Kb upstream of the *Myf-5* locus in a closely linked cluster of myogenic determination genes. Myf-6 cDNAs were isolated from human and mouse skeletal muscle, the only tissue in which expression of the corresponding mRNA was observed. In contrast to human primary muscle cell cultures which express moderate levels of Myf-6 mRNA, most established rodent muscle cell lines completely lack this mRNA. Myogenic 10T1/2 cells, however, induced by the expression of either pEMSV-Myf-4 or pEMSV-Myf-5 activate their endogenous mouse *Myf-6* gene. Constitutive expression of Myf-6 cDNA in C3H 10T1/2 fibroblasts establishes the muscle phenotype at a similar frequency to the previously characterized myogenic factors. Moreover, muscle-specific CAT reporter constructs containing either the human myosin light chain (MLC) enhancer or the promoter of the embryonic myosin light chain gene are activated in NIH 3T3 fibroblasts or in CV1 kidney cells by cotransfection of Myf-6 expression vehicles. This transcriptional activation occurs in the absence of any apparent conversion of the cellular phenotype of the recipient cells. Glutathione-S-transferase fusion proteins with Myf-3, Myf-4 or Myf-5 specifically bind to a MEF-like consensus sequence present in the human MLC enhancer and the MLC1 emb promoter. In contrast, the Myf-6 hybrid protein interacts weakly with the same sequences showing lower affinity and reduced specificity. Since co-expressed pEMSV-Myf-6, nevertheless, is able to activate transcription of the MLC-CAT reporter constructs in non-muscle tissue culture cells, the different DNA binding properties *in vitro* might suggest that transactivation of gene expression by Myf-6 involves distinct binding sites and/or additional protein factors.

Key words: DNA binding/gene complex/helix-loop-helix motif/muscle determination/transactivation

Introduction

The development of skeletal muscle tissues depends on a series of decision making events. During this process cellular transitions from multipotential mesodermal cells to committed but still premitotic muscle precursor cells occur. After withdrawing from the cell cycle, the mononucleated myoblasts differentiate into functional myotubes. These primary myocytes may undergo further maturation whereby expression of various muscle protein isoforms can be altered from one pattern to another depending on incoming signals or stimuli such as innervation or hormones. One of the major controls regulating these developmental events is exerted at the level of transcription. Molecular mechanism underlying this transcriptional control almost certainly involve specific regulatory protein factors which interact with defined *cis*-acting elements in the control regions of genes. Several muscle-specific regulatory factors have recently been identified which play an important role probably in both the determination as well as the differentiation of skeletal muscle cells in vertebrates.

The first cDNA recognized to encode a myogenic determination factor was designed MyoD1 (Davis *et al.*, 1987). It was isolated by applying a subtractive cDNA hybridization procedure to a cDNA library from myoblasts derived from C3H 10T1/2 fibroblasts after a brief treatment with 5-azacytidine (Taylor and Jones, 1979). Another cDNA clone coding for the muscle regulatory factor myogenin was identified by the specific accumulation of the corresponding mRNA in rat L6 myocytes at the onset of fusion and by the abundance of this mRNA in variants of L6 cells which are resistant against the differentiation block exerted by BrdU (Wright *et al.*, 1989). By taking advantage of the structural similarity to MyoD1, we have isolated a cDNA encoding the novel myogenic factor Myf-5 from human skeletal muscle (Braun *et al.*, 1989a) and two additional cDNA clones Myf-3 and Myf-4 representing the human homologues of MyoD1 and myogenin, respectively (Braun *et al.*, 1989b). Similarly, the mouse myogenin cDNA has been identified by sequence homology to MyoD1 (Edmondson and Olson, 1989). More recently, a chicken MyoD1 homologue (Lin *et al.*, 1989) and the MyoD1 homologue of Xenopus (Hopwood *et al.*, 1989) have also been described.

Each of these muscle-specific regulatory factors was shown to be capable of converting 10T1/2 fibroblasts to the muscle phenotype. Interestingly, all these proteins share a putative helix-loop-helix sequence domain which is also present in other regulatory DNA binding proteins and in the *myc* proto-oncogene family (Davis *et al.*, 1987; Caudy *et al.*, 1988). This conserved motif, originally designated as the basic and *myc*-like domain (Davis *et al.*, 1987) was demonstrated to be necessary and sufficient for the generation of the muscle phenotype in 10T1/2 cells by the MyoD1 protein (Tapscott *et al.*, 1988). To date, the existence of one more muscle determination gene *myd* which presumably is distinct from MyoD1 has been suggested. Employing human genomic DNA cloned into a selectable cosmid vector for transfection experiments, this DNA was found to convert the phenotype of 10T1/2 fibroblasts (Pinney *et al.*, 1988).

Although myogenic colonies have been obtained in these experiments, the putative *myd* gene has yet to be characterized in terms of its nucleotide structure.

In our search for the human genes encoding the myogenic factors Myf-4 and Myf-5, we detected an additional distinct *Myf* gene in the human genome. We report here the isolation and characterization of the novel cDNA clone Myf-6 isolated from human skeletal muscle. Apart from the extensively conserved sequence coding for the putative helix−loop−helix structure, this cDNA is completely unrelated in its structure to the previously described *Myf*-cDNAs. Myf-6 is encoded by a unique gene which is located in the immediate vicinity of the *Myf-5* gene on human chromosome 12. The ability of Myf-6 to convert 10T1/2 cells to muscle cells and its capacity to transactivate muscle-specific promoters in non-muscle cells is indistinguishable from the activities observed with Myf-3 (human MyoD1), Myf-4 (human myogenin) and Myf-5. In contrast to these latter proteins, however, bacterially produced Myf-6 protein almost completely fails to bind *in vitro* to a DNA consensus sequence which is present in the muscle-specific enhancer of the gene coding for the human myosin light chain 1/3 and other muscle-specific promoters.

Results

The muscle regulatory protein myf-6 constitutes a new member of the myogenic factor family in humans

During our search for the genes coding for the human myogenic factors Myf-3 to Myf-5, we have detected several recombinant phages carrying genomic sequences which weakly cross-hybridized under low stringency conditions. These clones failed to cross-react to Myf cDNAs when filters were washed at higher stringency. The moderate cross-hybridization was only observed when probes were used which contained the conserved helix−loop−helix sequence. No hybridization signals were obtained with probes generated from each of the 3′ non-coding regions of Myf cDNAs. This observation suggested to us that besides the genes for Myf-3, Myf-4, and Myf-5 additional Myf related sequences might be present in the human genome.

To test whether these genomic sequences were expressed in skeletal muscle, like the other *Myf* genes, a human genomic DNA fragment of ~ 160 bp was generated from the presumed coding region of one of the recombinant phages and used as a probe. With several RNAs isolated from muscle and non-muscle tissues hybridization signals were only obtained on skeletal muscle RNA. This result prompted us to search in a λgt11 cDNA library derived from human fetal skeletal muscle (Seidel *et al.*, 1988) for the corresponding cDNA. The investigation resulted in the isolation of several strongly cross-hybridizing cDNA clones. The longest isolate was designated Myf-6, in continuation of our previous cDNA numbering, and chosen for nucleotide sequence analysis.

As shown in Figure 1A, the cDNA contains 1294 nucleotides and has a single open reading frame for a putative peptide of 240 amino acids. The first AUG codon is part of a translational start site consensus sequence (GAACATGA) as proposed by Kozak (1987) and Cavener (1987). Comparison of the deduced amino acid sequence of Myf-6 with the predicted peptide sequences of Myf-3, Myf-4 and Myf-5 revealed a highly homologous segment encompassing the basic and *myc*-like domain and extending beyond this motif for ~ 15 additional amino acid residues towards the N-terminus. This sequence domain is also present in several proteins regulating development in *Drosophila*, such as twist (Thisse *et al.*, 1988), achaete scute (Alonso and Cabrera, 1988), daughterless (Caudy *et al.*, 1988; Cronmiller *et al.*,

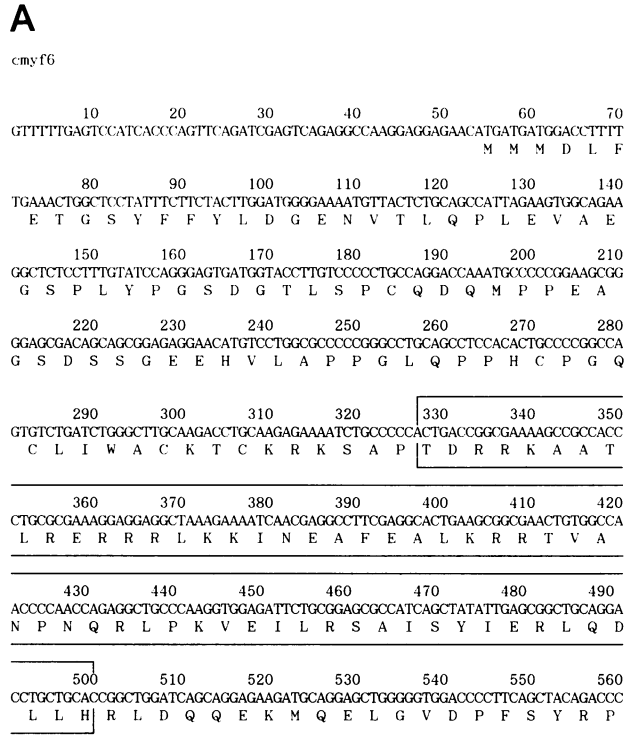


Fig. 1. (**A**) The nucleotide and predicted amino acid sequence of the human Myf-6 cDNA. The highly conserved putative helix−loop−helix structure is indicated by the box. A second conserved region which is also present in Myf-3 and Myf-5 is underlined.

1988), and enhancer of split (Klaembt *et al.*, 1989), and in factors binding to immunoglobulin enhancer, E12 and E47 (Murre *et al.*, 1989a) and the *myc*-protein family (DePinho *et al.*, 1986). The compilation of these sequences is shown in Figure 1B for the segment which was proposed to form a putative amphipathic helix−loop−helix (HLH) structure which is essential for dimerization of proteins and DNA binding (Murre *et al.*, 1989b). From this comparison it becomes evident that in spite of the highly conserved nature of the HLH domain, several amino acid differences exist dispersed over both helices and the putative loop segment. One could speculate that the constant amino acid residues in both helices shared among all proteins are most likely to be responsible for the general structural features, while the subtle differences observed even among the members of one particular family may be responsible for different DNA recognition or hetero- and homodimer formation. A second conserved sequence block in Myf-6 is shared with the other Myf proteins. This motif is located further downstream (amino acid residues 210−224) and is extremely rich in serine and threonine residues which could represent potential sites for protein phosphorylation. Besides the sequence similarities mentioned no additional structural homologies to the other Myf factors exist indicating that Myf-6 is a distinct and novel member of the myogenic factor family.

This assumption was further supported by the expression pattern of Myf-6 mRNA. As shown in Figure 2A, RNAs isolated from fetal and adult skeletal muscle of primarily fast (quadriceps) or slow (soleus) fiber types express Myf-6 mRNA, whereas cardiac muscle and non-muscle tissues do not contain this mRNA. We have observed no marked differences for Myf-6 expression in fetal and adult muscle tissues. The same pattern of mRNA expression with its restriction to skeletal muscle tissues was previously seen for the other human *Myf* genes (Braun *et al.*, 1989a,b). While Myf-6 mRNA is also synthesized in primary human muscle cells in culture (Figure 2B), most established muscle cell lines do not express this mRNA. Using the mouse Myf-6 cDNA probe which we have also isolated (unpublished data), the mouse cell lines C2C12, P2, and BC3H1, do not even show a trace of Myf-6 mRNA, whereas in RNAs isolated from the rat cell lines L6 and L8 extremely weak signals were detected after prolonged exposure of the Northern blot (data not shown).

In 10T1/2 or in the methylcholanthrene transformed 10T1/2 derivatives (MCA C115) which have been converted to muscle cells either with pEMSV−Myf-4 or pEMSV−Myf-5, two Myf-6 mRNA species of ~ 1.7 and 2.3 kb were expressed at similar levels (Figure 2C). While the 1.7 kb mRNA has the same length as Myf-6 mRNA in human skeletal muscle cells, the larger transcript has not been observed otherwise. It is at present unclear whether it is due to alternative splicing or alternative utilization of a second promoter or a second more distal poly(A) addition site. Both mRNA species are synthesized in growing as well as in differentiated cells. It is interesting to note, however, that the myogenic clone MCA−Myf-5/2, in contrast to clone MCA−Myf-5/1 which was derived by the same procedure, does not synthesize any Myf-6 mRNA. This clone has previously been shown to express MyoD1, myogenin, and

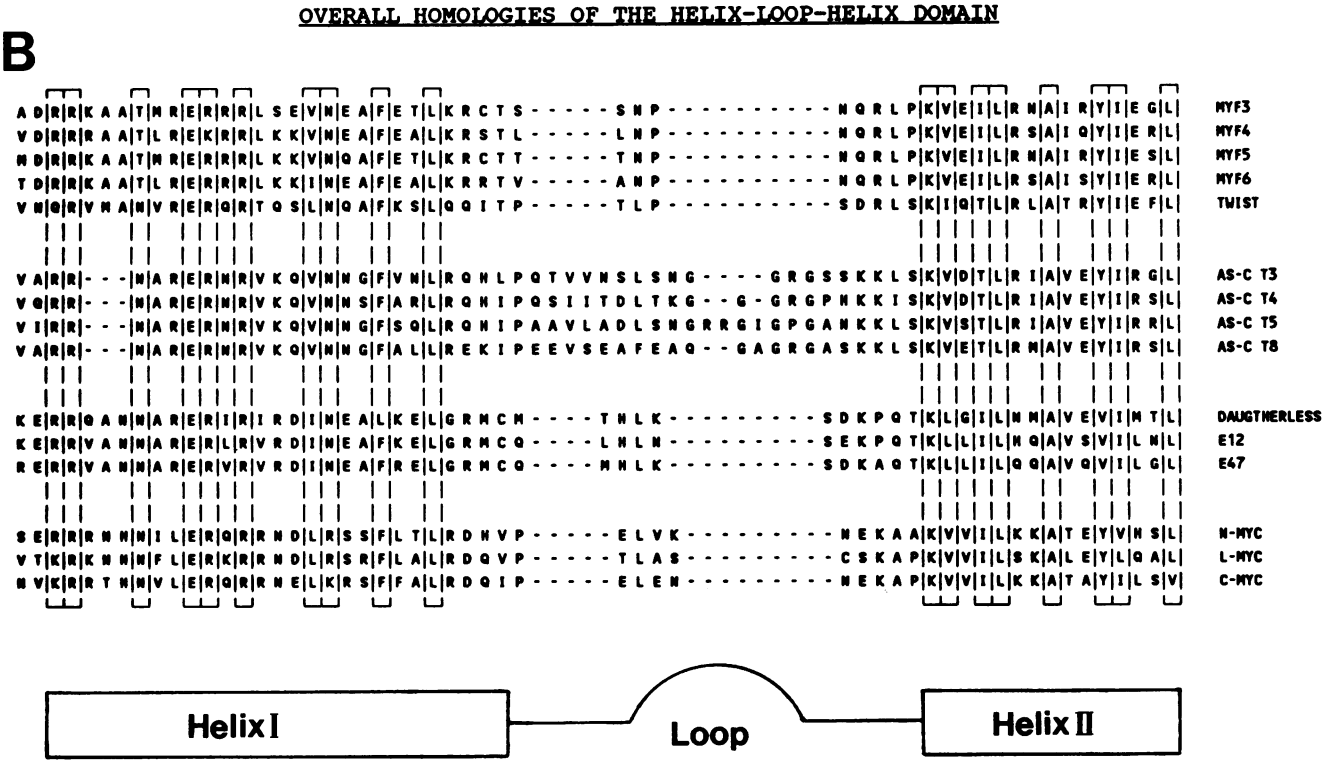


Fig. 1. (**B**) Summary of the putative helix−loop−helix sequences contained in several protein families. Amino acid residues (identical or equivalent) conserved in all proteins are indicated by boxes. The putative Helix I and Helix II forming segments connected by a loop structure of variable length are schematically shown. Sequences were derived from the following cDNAs: Myf-3 to Myf-6, human myogenic factors (Braun *et al.*, 1989a,b), twist, a mesodermal factor of *Drosophila* (Thisse *et al.*, 1988); AS-C T3-T8, neurogenic factors of *Drosophila* (Cabrera *et al.*, 1987; Villares and Cabrera, 1987); daughterless, a ubiquitous regulatory protein of *Drosophila* (Cronmiller *et al.*, 1988), E12 and E47, human IgG enhancer binding proteins (Murre *et al.*, 1989a), the myc proteins, cellular and viral oncogenes (Caudy *et al.*, 1988).

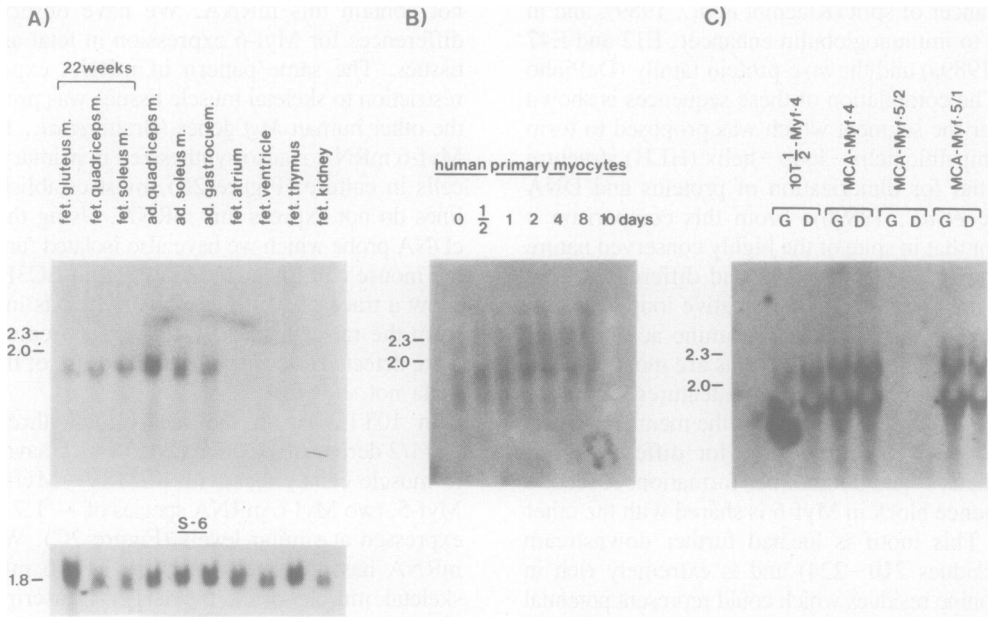


Fig. 2. Expression patterns of Myf-6 mRNA in human tissues (**A**), human primary muscle cells (**B**) and myogenic tissues culture cells (**C**). (**A**) Total RNA (25 μg) from the tissue indicated was isolated from two independently obtained human fetuses (22 weeks of gestation) and from surgically removed adult leg muscle. The blot was hybridized to a Myf-6 specific probe (5′ *Eco*RI/*Sma*I fragment of Myf-6 cDNA). This probe (sp.act. 1×10^8 c.p.m./μg) contained only sequences which are not conserved between the individual Myf factors. The RNA loading of the gels was controlled by hybridization to the S-6 cDNA coding for the constitutively expressed ribosomal protein S-6. (Heinze *et al.*, 1988). The blot was finally washed in 2 × SSC at 55°C and exposed on X-ray film overnight. (**B**) Human primary muscle cultures prepared from leg muscle of a 20 week old human fetus. Cells were cultured in differentiation supporting medium (10% horse serum) for the times indicated. RNA (20 μg) was isolated at each timepoint and hybridized to the Myf-6 specific cDNA probe (1×10^6 c.p.m./ml). The filter was exposed on film for 1 day. (**C**) Total RNA (25 μg) from growing (G) or differentiated (D) tissue culture cells was analysed with the mouse Myf-6 cDNA probe (600 bp *Pst*I fragment; sp.act. 1×10^8 c.p.m./μg). The cells were cloned lines of 10T1/2 or MCA C115 derived myoblasts converted by pEMSV−Myf-4 or pEMSV−Myf-5. MCA−Myf-5/1 and MCA−Myf-5/2 are two clones from independent transfection experiments. The film was exposed for 18 h.

Myf-5 mRNA (Braun *et al.*, 1989b). Taken together the results of *Myf-6* gene activation in some tissue culture cells and the lack of Myf-6 mRNA expression in most established muscle cell lines and one MCA−Myf-5 clone suggest that probably a mechanism of repression exists which can prevent synthesis of Myf-6 mRNA without affecting the other *Myf* genes.

Constitutive Myf-6 expression in 10T1/2 fibroblasts establishes the myogenic program

In order to test the biological activity of the Myf-6 protein, the cDNA was placed in the eucaryotic expression vehicle pEMSV-scribe under the control of the Moloney virus LTR (Davis *et al.*, 1987). When the construct pEMSV−Myf-6 was transfected into C3H 10T1/2 recipient cells together with the plasmid pSV2-neo expressing the selectable marker gene for G418 resistance more than one half of the colonies growing in the presence of the drug G418 had stably acquired the capacity to undergo myogenic differentiation. As shown in Figure 3, these cells when grown in low serum containing medium efficiently fused to multinucleated myotubes. The cells were clearly stained with the MF-20 monoclonal antibody (Bader *et al.*, 1982) indicating that they synthesize sarcomeric myosin heavy chains. By these criteria, the 10T1/2 cells phenotypically converted by Myf-6 appear indistinguishable from those derived by the expression of any of the other myogenic factors. The frequency of muscle cells appearing in Myf-6 converted 10T1/2 cultures is also rather similar to the other Myf-factors which argues for their comparable myogenic capacity.

Muscle-specific gene promoters can be activated in non-muscle cells by transient expression of Myf-6

The nuclear localization of myogenic factors as demonstrated for MyoD1 (Tapscott *et al.*, 1988) and their putative helix−loop−helix structure suggested that Myf proteins might be DNA binding peptides which are involved in transcriptional control. We have previously shown that Myf-5 when cotransfected with muscle-specific promoters into 10T1/2 fibroblasts leads to the activation of the reporter constructs (Braun *et al.*, 1989a). To test the potential transactivating capacity of Myf-6 in a transient assay, we analysed the effect of pEMSV−Myf-6 expression in NIH 3T3 fibroblasts and CV1 monkey kidney cells on the activity of cotransfected CAT constructs. These reporter plasmids contained either 1.2 kb of the promoter and 5′ upstream region of the human embryonic myosin light chain gene (*MLC1 emb*) (unpublished data) or the muscle-specific enhancer of the human *MLC1/3* gene (Seidel and Arnold, 1989). In contrast to 10T1/2 fibroblasts, both recipient cell types used here cannot be converted to the muscle phenotype at high frequency (Davis *et al.*, 1987). As shown in Figure 4, the expression of pEMSV−Myf-6, like the other three Myf clones, resulted in the efficient activation of both reporter constructs in both types of non-muscle recipient cell lines. The CAT-conversion was clearly dependent on the expression of Myf proteins since control experiments conducted with pEMSV-scribe or RSV-neo failed to activate transcription. The reporter CAT activities obtained with any of the four Myf factors were rather similar and were induced even in the absence of any apparent differentiation of the recipient cells. It should be

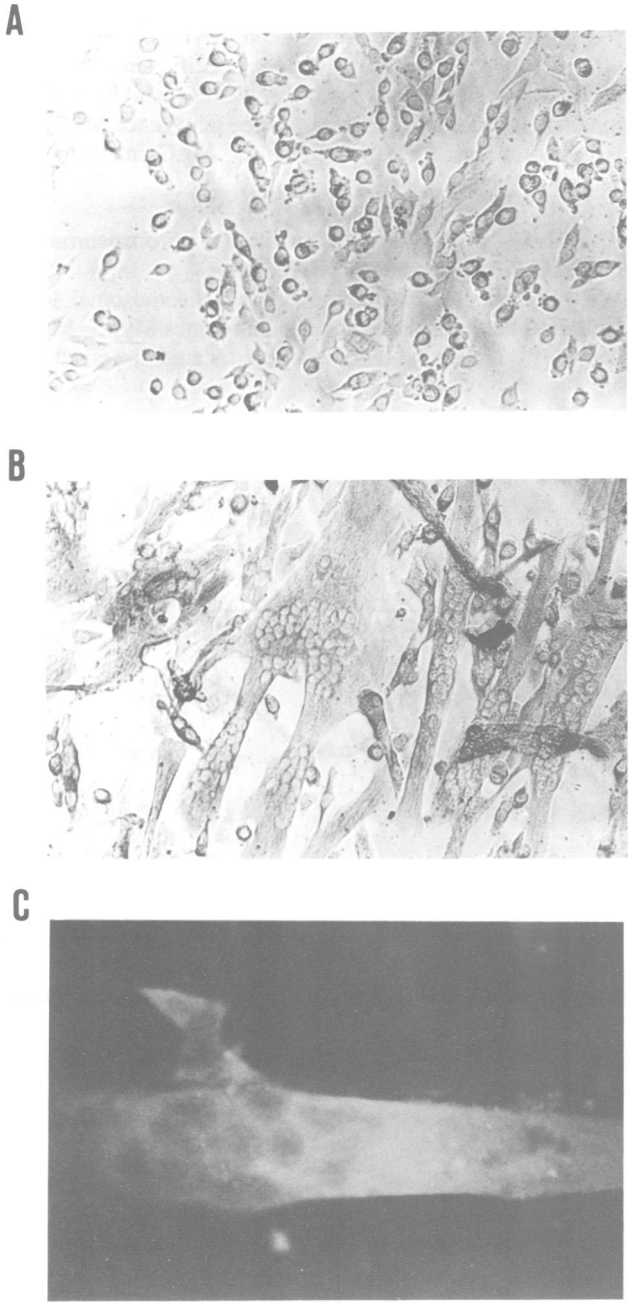


Fig. 3. Myogenic conversion of C3H 10T1/2 cells following expression of pEMSV−Myf-6. (**A**) Phase contrast microscopy of 10T1/2 fibroblast transfected with pEMSV-Myf-6 and pSV2-neo growing in DMEM plus 10% fetal calf serum supplemented with G418 (400 μg/ml). (**B**) Phase contrast image of a single G418 resistant colony of the 10T1/2−Myf-6 culture (same as in **A**) 5 days after serum deprivation (differentiation supporting medium). (**C**) Immunoreaction of a single myotube in pEMSV−Myf-6 converted 10T1/2 colony stained with the antimyosin heavy chain antibody MF20. The immunocomplex was visualized with rhodamine coupled secondary anti-mouse IgG.

noted here that the CAT activity in these experiments was determined 3 days after DNA transfection in cells growing in medium containing 10% fetal calf serum without selection. Constitutive control genes, such as the cytoplasmic β-actin promoter (P1-CAT), the SV 40 promoter (pSV2-CAT) or the herpes virus TK promoter remained virtually unaffected or were only slightly stimulated by the cotransfection of Myf-cDNAs (data not shown). From these results we conclude that DNA control elements which are required for muscle-specific gene transcription can be stimulated in non-muscle cells by transiently expressed Myf proteins. This effect is presumably mediated by DNA−protein interactions which might or might not involve other proteins.

Myf fusion proteins produced in Escherichia coli bind to regulatory sequence elements present in human myosin light chain genes

To investigate the possibility of DNA−protein interactions which could take place during the activation of the MLC-CAT constructs, Myf cDNAs were expressed as glutathione-S-transferase (GST) fusion proteins in *Escherichia coli* for *in vitro* binding studies. The purified hybrid proteins were analysed for their ability to bind to synthetic oligonucleotides in an electrophoretic mobility-shift assay (EMSA). The nucleotide sequences of the test oligonucleotides were selected from the human MLC1emb promoter region (unpublished results) and the human MLC1/3 enhancer (Seidel *et al.*, 1989; and unpublished results) by a computer assisted search for the MEF consensus sequence (Buskin and Hauschka, 1989). Both human muscle-specific *MLC* genes contain the conserved core sequence element CAGCAGGTG which was first postulated to interact with nuclear factors for the muscle-specific enhancer of the mouse *M-CPK* gene and was further shown to bind MyoD1 *in vitro* (Lassar *et al.*, 1989a).

As shown in Figure 5, bacterially produced Myf-3, Myf-4 and Myf-5 GST fusion proteins formed slow-moving DNA−protein complexes with the MLC-enh oligonucleotide in a concentration dependent fashion. When the MLC-Mut oligonucleotide containing nested point mutations which destroy the enhancer core consensus sequence were used for complex formation, no or very inefficient protein binding was obtained. Likewise, when only the control glutathione transferase protein was added to the assay no band shifts were observed (data not shown). The slight differences in complex sizes between the Myf proteins and the multiplicity of bands observed with the Myf-3 fusion protein were probably due to the different size proteins (see Materials and methods) and some protein degradation, respectively. To verify the specificity of the DNA−protein interactions, methylation interference footprinting was conducted on both strands of the binding oligonucleotide. As demonstrated in Figure 6, the Myf-5 protein forms contacts with the GC base pair located in the center of the core consensus sequence CAGCAGG. In addition, two guanosine residues located further outside were also involved in the complex formation. This analysis actually confirmed why the mutated version of the oligonucleotide which does not contain three of the four guanosines involved in binding, failed to form complexes. It should be pointed out here that relatively high concentrations of fusion proteins (50−400 ng) were required for the gel mobility-shift assays and the methylation interference footprinting. This might suggest that the interactions were of rather low affinity. Nevertheless, the results clearly indicate that the protein binding occurred in a sequence-specific manner. By contrast, when Myf-6 GST fusion protein was tested for DNA interaction with the same MLC-enh oligonucleotide, virtually no complexes were formed.

As shown in Figure 7A, in comparison to Myf-5 which specifically interacts with the mouse MCPK binding site MEF (AGGCAGCAGGTGTTGGGGGG), and the MLH-enhancer sequence but not with unrelated sequences such as the Ap1 binding site GATGAGTCAGCCG or the SBE binding site of the MLC2 gene GCCAAAAGTG-GACATGG (Braun *et al.*, 1989c), Myf-6 applied at the same protein concentrations results in extremely faint complexes. At 12-fold higher Myf-6 concentration, however, protein complexes arose not only with the MEF and MLC-enh oligonucleotides but also with the MLC-Mut sequence (Figure 7B). From these observations, we conclude that under the applied testing conditions, Myf-6 only very weakly interacts with the same DNA sequences that are specifically recognized by Myf-3, Myf-4 and Myf-5. The residual affinity, although not completely unspecific since AP1 and SBE binding sites do not form complexes, exhibits different binding characteristics of relaxed specificity since MLC-Mut is bound at least as efficiently as the MLC-enhancer consensus sequence.

The Myf6 gene is located on human chromosome 12 in close proximity to the Myf5 gene

We have previously determined the chromosomal localization of the human myogenic factor genes Myf-3, Myf-4, and Myf-5 and found that they were located dispersed over

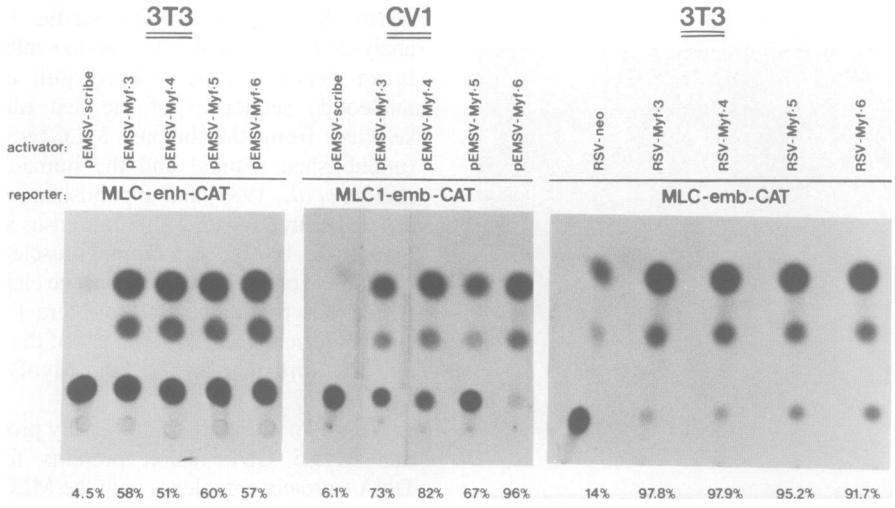


Fig. 4. Muscle-specific promoter activation by transient expression of cotransfected pEMSV−Myf vehicles in NIH 3T3 fibroblasts and CV1 cells. Transcription of the reporter constructs MLC-enh−CAT and MLC1emb−CAT is under the control of the muscle-specific elements of the human MLC1/3 downstream enhancer (Seidel and Arnold, 1989) and the 1.2 kb promoter fragment of the human *MLC 1emb* gene, respectively. Equal concentrations (10 μg) of reporter constructs and Myf expression clones were contransfected as described (Materials and methods). CAT assays were performed with 1% (NIH 3T3) or 5% (CV1) of total cell extracts under standard conditions (Gorman, 1985). The percent conversion of chloramphenicol substrate to its acetylated derivatives is indicated for each assay.

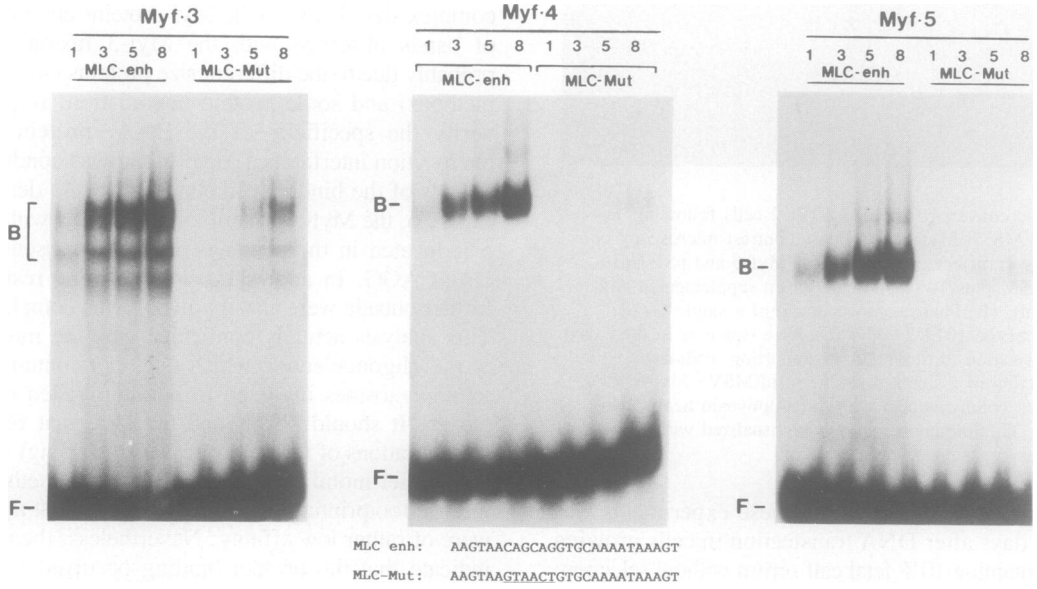


Fig. 5. Gel mobility-shift assays (EMSA) for Myf-3, Myf-4 and Myf-5 glutathione S-transferase fusion proteins binding to the MLC-enh oligonucleotide. Purified Myf-fusion proteins (1−8 μl; 50 ng/μl) were used for each EMSA with either double-stranded MLC-enh or the MLC-Mut oligonucleotides. The nucleotide sequences of the coding strands are indicated. DNA−protein complexes were formed under the conditions described under Materials and methods and separated on 4.5% polyacrylamide gels into bound (B) and free (F) fractions.

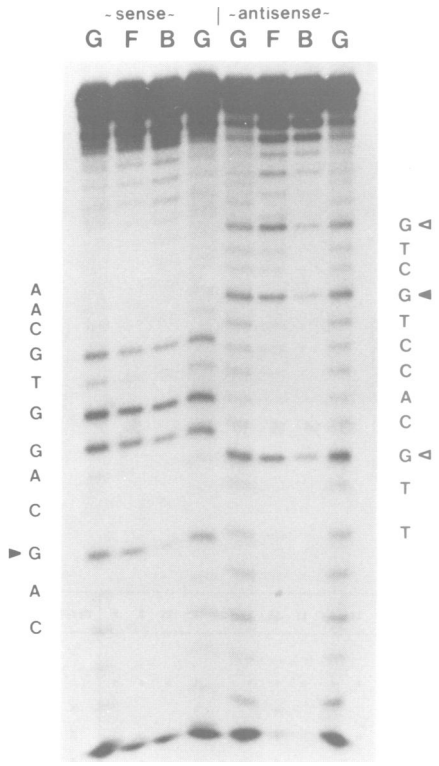


Fig. 6. Methylation interference footprinting of Myf-5–GST fusion protein bound to the MLC-enh oligonucleotide. Sense and antisense strands were analysed as described under Materials and methods. The G residues showing reduced cutting frequency in the bound complex are indicated in the binding sequence by filled arrow heads, weaker G interactions are illustrated by open arrow heads. The same picture was also obtained for Myf-3–GST and Myf-4–GST.

the genome on different human chromosomes (Braun *et al.*, 1989b). In order to map the *Myf-6* gene locus, we utilized a panel of human/rodent somatic hybrid cell lines which had been shown to retain defined subsets of human chromosomes (Balazc *et al.*, 1984). As demonstrated in Figure 8A, on a Southern blot of DNA from numerous hybrid cell lines only those exhibited the diagnostic Myf-6 hybridization signal which contained the human chromosome 12. This pattern of concordance was identical to the one which we had previously determined for the *Myf-5* gene which is also located on chromosome 12 (Figure 8B). An attempt at a more detailed regional assignment by *in situ* hybridization suggested that the two genes might be genetically linked (data not shown). Indeed, when a human genomic library was screened for the *Myf-6* gene, two of the isolated recombinant lambda phages contained the entire or parts of the *Myf-6* gene as well as the *Myf-5* gene. As shown in Figure 8C, the set of overlapping phage allowed us to deduce the precise genomic organization of both *Myf* genes. According to this analysis, the *Myf-6* gene is located upstream of *Myf-5* in the same orientation at a distance of ~ 6.5 kb between the genes.

Discussion

The Myf-6 gene structurally and functionally belongs to the human gene family encoding muscle regulatory proteins

The Myf-6 cDNA isolated from a human fetal skeletal muscle library represents a novel member of a family of genes which codes for muscle regulatory proteins. We have previously reported the isolation of the cDNAs Myf-3 and Myf-4, the human homologues to MyoD1 (Davis *et al.*, 1987) and the rodent myogenins (Edmondson and Olson, 1989; Wright *et al.*, 1989), by their structural similarities

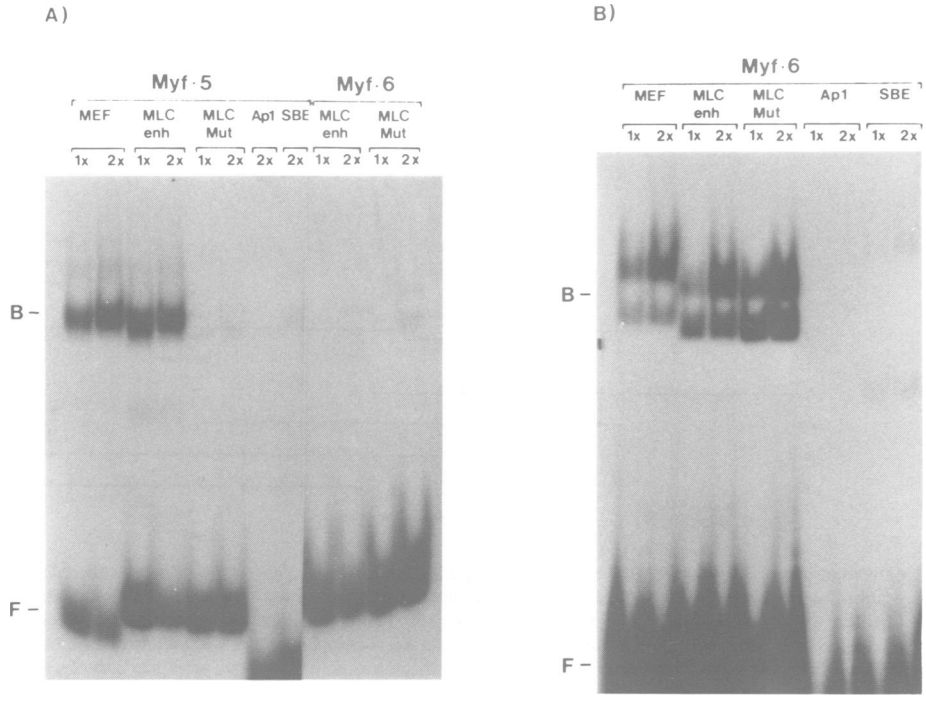
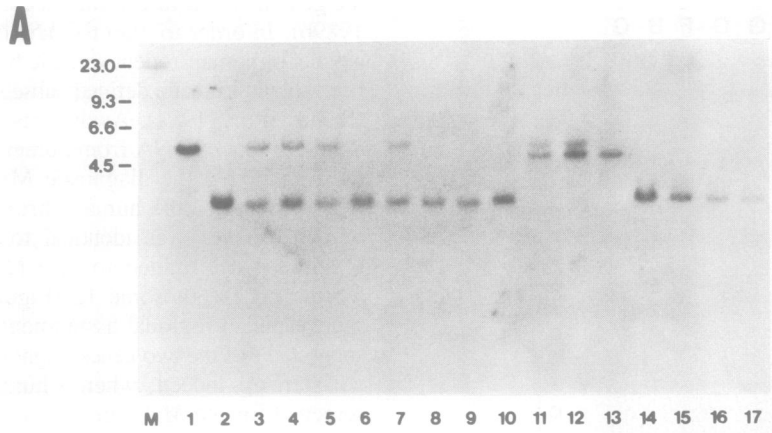


Fig. 7. Comparison of DNA binding of Myf-5 and Myf-6 fusion proteins on various DNA sequences. The gel mobility shift assays in (**A**) were performed with 100 (1×) and 200 (2×) ng of proteins using the indicated oligonucleotides as binding targets. The same assays in (**B**) were performed with 1.2 and 2.4 μg of Myf-6–GST fusion protein. EMSA conditions are described under Materials and methods.



B

Human Chromosomes

	Hybrid Cells	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	Myf-6
1.	GM3104 control																									
2.	B82 MS2 1a-14-1		+	+	+			+	+			+		+		+	+	+	+	+		+	+	+	+	
3.	RAG Anly 1	+		+	+	+	+						+		o				+			+				X
4.	RAG PJ 7-2	+		+		+	+	+	+	+		o	+		+	o			+				+	o		X
5.	RAG PJ 5-15	+		+	+	+	+	+	+		+	+	+		+	+	+	+	+	+	+	+	+	o		X
6.	RAG GM194 7	+		o	+	+			+		+		o							+		+		o		
7.	RAG GM194 5-5			o	+	+	+	+			+		+	+	+		+				+	+	+	o		X
8.	RAG GO 4			+	+			+	+					+	+	+			+				+	o		
9.	RAG SU 3-1-2-3			o		+	o	o														+		+		
10.	A9 SU 1-2		+	o			+							+	+				+			+				
11.	V79 Ly 3-2	+		+	+		+		+		+		+		+			+				+	+	+	+	X
12.	V79 Ly 3-3	+		+	+				+		+		+		+			+				+	+	+	+	X
13.	P3 control																									
14.	A9 GM89 9c-7			+	+			+				+				+	+		+		o	+	+	o		
15.	A9 MS58 2 b		o									+		+	o	+								+		
16.	RAG GM97 8-13-3	o			+					+		+			+			+						o		
17.	RAG control																									

C

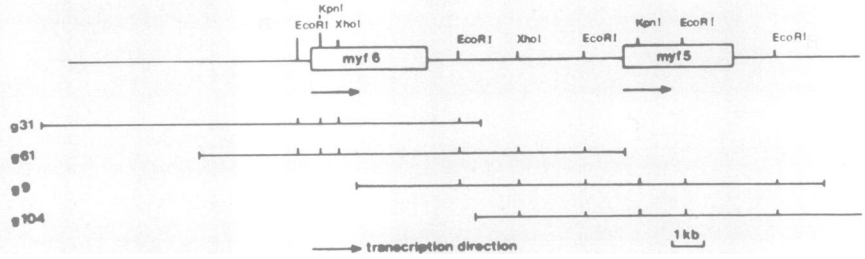


Fig. 8. Chromosomal localization of the *Myf-6* gene and the genomic organization of the *Myf-5–Myf-6* gene cluster. **(A)** Southern blot analysis of DNA from human/mouse hybrid cells (**lanes 2–10** and **14–16**), human/hamster hybrids (**lanes 11** and **12**), human control (**lane 13**) and mouse control cells (**lane 17**). The blot was hybridized to the Myf-6 cDNA (sp.act. 1×10^8 c.p.m./μg) under stringent conditions (washing: 55°C, 0.1 × SSC, 0.1% SDS). **(B)** Schematic representation of hybrid cell lines and their complement of human chromosomes. + stands for the presence of the respective human chromosome, ○ indicates that parts of the chromosome are present, gaps mark the lack of chromosomes and × marks cell lines which exhibit the Myf-6 signal diagnostic for the human gene. Symbols for the chromosomes in the human control cell GM 3104 have been omitted. **(C)** Physical map of the *Myf-5/Myf-6* gene cluster analysed on four overlapping lambda phages (g31, g61, g9, and g104). The location of the *Myf* genes was derived from hybridization analyses on the four isolated phages using Myf-6 and Myf-5 specific probes, respectively. The orientation of each gene was analysed with 5′ and 3′-specific probes as described under Materials and methods. Only restriction sites used for the analysis are shown.

to MyoD1 cDNA (Braun *et al.*, 1989b). We have furthermore identified the Myf-5 cDNA, another distinct myogenic determination factor in humans (Braun *et al.*, 1989a). We describe here the isolation and characterization of a fourth muscle regulatory protein from the same species and demonstrate that it also contains the highly conserved potential amphipathic helix−loop−helix structure which appears important for the biological activity of MyoD1 (Tapscott *et al.*, 1988) and which possibly involves nuclear localization, protein dimerization (Murre *et al.*, 1989b) and specific DNA binding (Lassar *et al.*, 1989a). In the course of the preparation of this manuscript, we were informed of the isolation of the MRF-4 cDNA from adult rat muscle (S.Konieczny, personal communication). This clone is highly homologous in sequence to Myf-6 and therefore might constitute the rat homologue of Myf-6. Like the previously identified myogenic cDNAs, constitutively expressed Myf-6 is capable of efficiently converting the phenotype of 10T1/2 fibroblasts to myoblasts. The 10T1/2 derived myoblasts are readily recognized in serum deprived growth medium by their ability to form multinucleated myotubes and to activate the expression of muscle-specific proteins. This myogenic conversion appears initially solely dependent on the expression of one of the myogenic factors. Clearly, no neural or hormonal input or specific cellular interactions are required for the formation of normal differentiating muscle cells which emphasizes the autonomous regulatory capacity of the *Myf* genes in cytodifferentiation.

It is also clear, however, that the induction of the myogenic phenotype by any of the Myf factors leads to auto- or cross-activation of the other members of the *Myf* gene family (Braun *et al.*, 1989b; Thayer *et al.*, 1989). This autoregulatory loop of *Myf* gene expression may be an important factor for the maintenance of the complete myogenic programme. In this context it is interesting to note that fully differentiated skeletal muscle tissues as well as primary muscle cell cultures simultaneously express all four Myf mRNAs, whereas established muscle tissue culture cells only express subsets. For example, Myf-6 mRNA was not found in any muscle cell line except at trace amounts in rat L-6 cells. The lack of Myf-6 expression, or any other *Myf* gene expression, might be due to a specific suppression mechanism which comes into action during the immortalization of stable muscle cell lines. It is evident from the fact that the *Myf-6* gene can be activated in myocytes derived from 10T1/2 or MCA cells that the gene is not irreversibly shut down in established tissue culture cells and the repression is subject to regulation. Candidate proteins for the reversible repression of *MyoD1* gene transcription have been identified as mutated H-*ras* and v-*fos* are both capable of blocking myogenic differentiation by virtue of preventing MyoD1 synthesis. (Lassar *et al.*, 1989b; Sternberg *et al.*, 1989). In the light of the different patterns of Myf expression seen in established muscle cell lines, it is interesting to ask how complete the myogenic programs actually are in these cells.

The Myf-6 gene is part of a gene cluster on human chromosome 12

The generation of the muscle phenotype in 10T1/2 cells was demonstrated in transfection experiments with DNA from C2C12 muscle cells or 10T1/2 myoblasts converted by 5-azacytidine but not with DNA from normal 10T1/2 fibroblasts (Lassar *et al.*, 1986). The frequency of the conversion event and its stable and heritable nature suggested that a single gene or genetic locus would be responsible for the generation of the myogenic phenotype (Konieczny and Emerson, 1984; Lassar *et al.*, 1986). Meanwhile, at least four distinct gene products have been identified which individually or in concert are capable of turning 10T1/2 fibroblasts into muscle cells. For the human factors Myf-3, Myf-4 and Myf-5 we have shown that the corresponding genes are scattered around the human genome (Braun *et al.*, 1989b) and they therefore represent independent biological activities which are functionally equivalent to each other and to the putative *myd* gene (Pinney *et al.*, 1988). In contrast, the new *Myf-6* gene is closly linked to the *Myf-5* gene on chromosome 12. Since *Myf-5* and *Myf-6* are structurally *not* more closely related than the other *Myf* genes, this gene cluster probably is not the result of a recent gene duplication. Although this is the first instance for gene clustering of *Myf* genes, numerous examples exist where genes involved in developmental regulation are organized in gene complexes. For instance, the *achaete-scute* genes involved in neuronal development of *Drosophila* are linked in one genetic locus comprised of several structurally related genes (Alonso and Cabrera, 1988). Particularly interesting examples exist in the antennapedia and bithorax gene complexes which control the anterior−posterior specialization in the developing fly (Duboule and Dole, 1989; Graham *et al.*, 1989). These genes are clustered on the chromosome in the 3′-5′ direction reflecting the order of the anterior limits of their expression along the body axis (Ingham, 1985; Graham *et al.*, 1989).

A similar situation seems to be realized in *Hox* gene clusters, which represent mouse homologues of the antennapedia and bithorax complexes. In contrast other vertebrate homologous of *Drosophila* pattern forming genes such as the paired box containing genes Pax (Deutsch *et al.*, 1988) or engrailed like genes (Davis and Joyner, 1988) do not show this organization. At present, it is unknown whether additional *Myf* related genes exist and if so whether they are linked to the known *Myf-5−Myf-6* gene cluster or to any of the other two human *Myf* genes located on different chromosomes.

Myf-6, most likely acts as transcriptional activator, like the other Myf proteins

Based on the cellular localization of Myf proteins in the nucleus (Tapscott *et al.*, 1988; our unpublished results) and their structural features it appears reasonable to assume that these proteins interact with DNA to regulate gene transcription in muscle cells. Short of a cell-free transcription assay, we demonstrate here that muscle-specific reporter genes can indeed be activated in non-muscle cells by cotransfection of expression vectors carrying any of the four Myf cDNAs. Although this experiment certainly does not prove the direct effect of Myf proteins on regulatory elements of muscle-specific genes, it strongly argues for their critical involvement in the transcriptional activation. This inference is further supported by the *in vitro* findings that bacterially produced Myf proteins, specifically bind to MEF-like consensus elements present in the human myosin light chain genes. For mouse MyoD1 it has been demonstrated that dimerization with the immunoglobulin enhancer binding protein E12 greatly increases the specific DNA binding affinity (Murre *et al.*, 1989b). We have preliminary data

which indicate that this could also be the case for all Myf proteins (unpublished results). An interesting observation, however, has been made with the Myf-6 protein which essentially does not bind to the MEF-consensus sequence *in vitro* but nevertheless activates reporter constructs *in vivo*. Several explanations are possible. Activation by Myf-6 could involve *cis*-elements which are different from the MEF-binding motif. These distinct binding sites are likely to exist since transcription of the α-cardiac actin promoter which lacks any obvious MEF consensus sequence can nevertheless be activated by cotransfected CMD1 (Lin *et al.*, 1989) or any of the Myf expression vectors (unpublished data). Alternatively, Myf-6, like the other Myf factors could require other proteins such as E12 to effectively bind to the MEF consensus sequence in the cell, but in contrast to Myf-3, Myf-4 and Myf-5, Myf-6 completely fails to bind *in vitro* even at high concentrations possibly due to its inability to form homodimers. Since these secondary proteins (E12, E47) are ubiquitous factors and therefore should be present in the transfected cells but are of course absent in the *in vitro* binding, this could explain the *in vivo* activation in the absence of *in vitro* binding. In any case, the different behaviour of Myf-6 in the EMSA as compared to Myf-3, Myf-4, and Myf-5 is a first indication for possible differences in the biological targets for the Myf proteins or a hint for alternative mechanisms by which these proteins might act.

Materials and methods

Isolation of Myf-6 cDNA and determination of its nucleotide sequence

The Myf-6 cDNA clone was isolated by screening 2×10^6 p.f.u. of a human fetal muscle library (Seidel *et al.*, 1989) using a 160 bp *Pst*I fragment derived from the 5'-coding region of the human *Myf-6* gene (unpublished result). Labelling of DNA and hybridization conditions have been described elsewhere (Braun *et al.*, 1989a). DNA was isolated and subcloned for dideoxy-sequencing in either pBS plasmids (Stratagene) or M13mp18 or mp19 vectors (Sanger *et al.*, 1977). The nucleotide sequence from the longest cDNA clone was determined on both strands using restriction fragments and unidirectional *Exo*III deletions (Henikoff, 1984). The corresponding mouse Myf-6 cDNA was isolated from 1×10^6 p.f.u. of a newborn mouse muscle cDNA library (the library was kindly provided by M.Buckingham) using the same 160 pb probe. The mouse Myf-6 clone was unequivocally identified by its nucleotide sequence analysis which revealed a remarkable conservation to the human Myf-6 sequence (unpublished). All DNA work was performed with standard procedures as outlined in Maniatis *et al.* (1982).

Overlapping lambda phages containing the human *myf-6/myf-5* gene locus (see Figure 3C) were isolated by screening a human genomic EMBL-3 library (kindly provided by A.M.Frischauf) with full-length myf-5 cDNA under reduced stringency. The isolated clones were analysed by restriction mapping, partial sequencing, and hybridization with appropriate probes to localize both genes on a contiguous segment of DNA. Clone λg61 contained the major part of both genes on a single 10.5 kb *Kpn*I fragment.

Plasmid constructions

For the expression in eucaryotic cells the Myf-6 cDNA was subcloned into the *Eco*RI site of pEMSV-scribe vector (kindly supplied by A.Lassar). The expression constructs for Myf-3, Myf-4 and Myf-5 have been described previously (Braun *et al.*, 1989a,b).

To express glutathione S-transferase-Myf fusion proteins in *E.coli*, the pGEX expression vectors (Smith and Johnson, 1988) were utilized. In detail, the complete Myf-3 *Eco*RI cDNA fragment was cloned into the *Eco*RI site of pGEX1, the Myf-4 cDNA starting with the codon for amino acid 39 was subcloned into the *Eco*RI site of pGEX 2, the Myf-5 *Bam*HI/*SpHI* fragment starting in the untranslated 5' leader of Myf-5 and extending to the codon for amino acid 168 was subcloned into the *Bam*HI/*Sma*I sites of pGEX2. The Myf-6 cDNA was cloned as *Eco*RI fragment starting with the codon for amino acid 2 (shorter cDNA isolate) into the pGEX1 vector. All constructs were examined to contain the inserted DNA in the correct reading frame of GST by double-strand sequencing and restriction mapping for proper orientation.

Preparation of GST-Myf fusion proteins in E.coli

The various GST-Myf plasmids were expressed in *E.coli* and purified as described by Smith and Johnson (1988) with the modifications introduced by Lassar *et al.* (1989a). Each preparation used for DNA binding experiments was analysed on SDS−PAGE for purity and integrity. Fusion proteins were identified by Western blot analysis with monoclonal antibodies against MyoD1 (kindly supplied by S.Kohtz) and myogenin (kindly provided by W.E.Wright) and with polyclonal antiserum raised against a β-Gal-Myf-5 fusion protein (unpublished). The polyclonal Myf-5 antiserum cross-reacts with both Myf-3 and Myf-6 but not with Myf-4 in an immunoprecipitation assay (data not shown).

Electrophoretic mobility-shift (EMSA) and methylation interference footprinting

The binding and reaction conditions used for EMSA and methylation interference footprinting have been described in Braun *et al.* (1989c). For a typical band shift experiment ~ 50−400 ng of purified fusion-protein was used in a reaction volume of 25 μl. Preparative EMSA for the methylation interference footprinting were scaled up five times. The coding strand sequences were prepared as synthetic oligonucleotides containing either the wild type binding site (MLC-enh) AAGTAACAGCAGGTGCAA-AATAAAGT or the mutant binding site (MLC-Mut) AAGTAAGTAAC-TGTGCAAAATAAAGT. These sequnces were derived from the human *MLC1/3* gene enhancer (unpublished). The purification and labelling procedure of double stranded oligonucleotides was described previously (Braun *et al.*, 1989c).

Blot analysis of RNA from human tissue culture cells and DNA from rodent/human somatic cell hybrids

RNA was isolated from frozen tissues of two 22 week old human fetuses, adult leg muscle (surgery) and tissue culture cells by the methods of Auffray and Rougeon (1980) and Chomczynski and Sacchi (1987). Gel electrophoresis, RNA transfer and hybridization conditions have been described by Braun *et al.* (1989b). To generate human Myf-6 specific probe, a 5'cDNA fragment was isolated by digestion with *Eco*RI and *Sma*I resulting in a 250 bp fragment located upstream of the conserved helix−loop−helix motif. A mouse specific Myf-6 probe was generated from the 3'-half of the mouse cDNA by digestion with *Pst*I resulting in a 600 bp fragment located downstream of the conserved region. The source of the human−rodent hybrid cell lines used to establish the chromosomal localization of Myf-6 was described previously (Braun *et al.*, 1989b). The Southern blot hybridization was performed with the full length human Myf-6 cDNA probe applying hybridization and washing conditions (final washing step: 0.1 × SSC, 0.1% SDS at 55°C) which prevent cross-hybridization to other human *Myf* genes.

Cell culture, DNA transfection, and immunostaining

Mouse C2C12 (Yaffe and Saxel, 1977a), BC3H1 cells (Schubert *et al.*, 1974), rat L6 and L8 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, P2 aza-myoblasts were kindly supplied by A.Lassar, Seattle. NIH 3T3 and CV1 cells were obtained from W.Ostertag and W.Deppert, Hamburg, respectively. All cells were grown as recommended by the American Type Culture Collection (ATCC) or in Dulbecco's modified Eagles Medium (DMEM) containing 10% fetal calf serum. Differentiation of the myogenic cell lines was generally induced by serum withdrawal. Stable myogenic clones from C3H 10T1/2 or MCA C115 cells were obtained by transfection of ~ $1-5 \times 10^5$ cells with 1 μg of pSV2-neo plasmid and 30 μg of the various pEMSV-Myf expression vectors using the Ca-phosphate precipitation procedure (Graham and Van der Eb, 1973). Selection of G418 resistant colonies and induction of myogenic differentiation were carried out as described previously. (Braun *et al.*, 1989a,b). For immuno-cytochemistry cell cultures were rinsed with cold phosphate buffered saline (PBS; 13mM NaCl, 0.15 mM KH_2PO_4, 0.27 mM KCl., 0.8 mM Na_2PO_4 pH 7.4) and fixed in a 20:2:1 solution of ethanol (70%), formaldehyde and acetic acid for few min at 4°C. Fixed cultures were washed several times with PBS and incubated with anti-myosin monoclonal antibody MF-20 (Bader *et al.*, 1982) for 1 h at 4°C. Following two washes in PBS, cultures were reincubated with rhodamine coupled anti-mouse IgG second antibody for an additional hour. Immuno-complexes were visualized with a fluorescence microscope.

Transient transfection of pEMSV-Myf plasmids with CAT reporter constructs in NIH3T3 and CV1 cells

Approximately 5×10^5 CV1 or NIH 3T3 cells were transfected in 10 cm dishes with double-banded supercoiled plasmid DNA. For cotransfection experiments 10 μg of the CAT reporter-constructs were mixed with 10 μg of the various Myf expression constructs and precipitated with calcium

phosphate as described elsewhere (Arnold *et al.*, 1988). Cells were harvested after 3 days in serum rich medium and CAT activity was determined as described by Gorman (1985) with 1–5% of total cellular extracts standardized either to constant concentrations or constant β-Gal activity from a cotransfected *LacZ* expression vector. Both procedures gave essentially the same results. Transfection efficiency for different recipient cells was calibrated to the β-actin CAT construct P1-CAT (Lohse and Arnold, 1988). The construction of MLC1-Enh-CAT was described in Seidel and Arnold (1989). The experiments leading to the construction of MLCemb-CAT will be described elsewhere (Bober *et al.*, in preparation).

Acknowledgements

We wish to thank A.Lassar for generously providing the pGex expression vectors, W.Wright for his rat L6 subclone and the MF20 hybridomas, S.Kohtz for the RNA blot of human primary muscle cells, K.H.Grzeschik for the panel blot with DNA from the somatic cell hybrids and J.Buskin for the MEF oligonucleotide. We also are grateful to S.Konieczny for communicating the rat MRF4 sequence prior to publication. We greatly appreciate the comments on the manuscript by M.Buckingham. The skilful technical assistance of H.Eberhardt and the secreterial help of A.Broecker-Nagel are gratefully acknowledged. This work was supported by Deutsche Forschungsgemeinschaft and Deutsche Muskelschwundhilfe e.V.

References

Alonso,M.C. and Cabrera,C.V. (1988) *EMBO J.*, **7**, 2585–2591.
Arnold,H.H., Tannich,E. and Paterson,B.M. (1988) *Nucleic Acids Res.*, **16**, 2411–2429.
Auffray,C. and Rougeon,F. (1980) *Eur. J. Biochem.*, **107**, 303–313.
Bader,D., Masaki,T. and Fishmann,D.A. (1982) *J. Cell. Biol.*, **95**, 763–770.
Balazc,J., Purello,M., Aldaheff,B., Grzeschik,K.W. and Szabo,P. (1984) *Hum. Genetics.* **68**, 57–61.
Braun,T., Buschhausen-Denker,G., Bober,E., Tannich,E. and Arnold,H.H. (1989a) *EMBO J.*, **8**, 701–709.
Braun,T., Bober,E., Buschhausen-Denker,G., Kohtz,S., Grzeschik,K.H. and Arnold,H.H. (1989b) *EMBO J.*, **8**, 3617–3625.
Braun,T., Tannich,E., Buschhausen-Denker,G. and Arnold,H.H. (1989c) *Mol. Cell. Biol.*, **9**, 2513–2525.
Buskin,J.N. and Hauschka,S.D. (1989) *Mol. Cell. Biol.* **9**, 2627–2640.
Cabrera,C.V., Martinez-Arias,A. and Bate,M. (1987) *Cell*, **50**, 425–433.
Caudy,M., Vaessin,H., Brand,M., Tuma,R., Jan,L.Y., and Jan,Y.N. (1988) *Cell*, **55**, 1061–1067.
Cavener,D.R. (1987) *Nucleic Acids Res.*, **15**, 1353–1361.
Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
Cronmiller,C.R., Schedl,P. and Cline,T.W. (1988) *Genes Dev.*, **2**, 1666–1676.
Davis,C. and Joyner,A. (1988) *Genes Dev.*, **2**, 1736–1744.
Davis,R.L., Weintraub,H. and Lassar,A.B. (1987) *Cell*, **51**, 987–1000.
DePinho,R.A., Legony,E., Feldman,L.B., Kohl,N.E., Yancopoulos,G.D. and Alt,F.W. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1827–1831.
Deutsch,U., Dressler,G.R. and Gruss,P. (1988) *Cell*, **53**, 617–625.
Duboule,D. and Dolle,P. (1989) *EMBO J.*, **8**, 1497–1506.
Edmondson,D.G. and Olson,E.N. (1989) *Genes Dev.*, **3**, 628–640.
Gorman,C.M. (1985) In Glover,D.M. (ed.), *DNA Cloning II.* IRL Press, Oxford, UK, pp. 143–190.
Graham,A., Papalopulu,N. and Krumlauf,R. (1988) *Cell*, **57**, 367–378.
Graham,F.L., and van der Eb,A.J. (1973) *Virology*, **52**, 456–467.
Heinze,H., Arnold,H.H., Fischer,D. and Kruppa,J. (1988) *J. Biol. Chem.*, **263**, 4139–4144.
Henikoff,S. (1984) *Gene*, **28**, 351–359.
Hopwood,N.D., Pluck,A. and Gurdon,J.B. (1989) *EMBO J.*, **8**, 3409–3417.
Ingham,P. (1985) *Trends Genet.*, **1**, 74–80.
Klaembt,C., Knust,E., Tietze,K. and Campos-Ortega,J.A. (1989) *EMBO J.*, **8**, 203–210.
Konieczwy,S.F. and Emerson,C.P. (1984) *Cell*, **38**, 791–800.
Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
Lassar,A.B., Paterson,B.M. and Weintraub,H. (1986) *Cell*, **47**, 649–656.
Lassar,A.B., Buskin,J.N., Lockshon,D., Davis,R.L., Apone,S., Hauschka,S.D. and Weintraub,H. (1989a) *Cell*, **58**, 823–831.
Lassar,A.B., Thayer,M.J., Overell,R.W. and Weintraub,H. (1989b) *Cell*, **58**, 659–667.
Lin,Z., Duchesne,C.A., Eldridge,J. and Paterson,B.M. (1989) *Genes Dev.*, **3**, 986–996.
Lohse,P. and Arnold,H.H. (1988) *Nucleic Acids Res.*, **16**, 2787–2803.
Maniatis,T., Fritsch,E.F. and Sambrook,J.C. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Murre,C., Schonleber-McCaw,P. and Baltimore,D. (1989a) *Cell*, **56**, 777–783.
Murre,C., Schonleber-McCaw,P., Vaessin,H., Caudy,M., Jan,L.Y., Jan,Y.N., Cabrera,C.V., Buskin,J.N., Hauschka,S.D., Lassar,A.B., Weintraub,H. and Baltimore,D. (1989b) *Cell*, **58**, 537–544.
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.
Schubert,D., Harris,A.J., Devine,C.E. and Heinemann,S. (1974) *J. Cell. Biol.*, **61**, 398–413.
Seidel,U., Bober,E., Winter,B., Lenz,S., Lohse,P., Goedde, H.W., Grzeschik,K.H. and Arnold,H.H. (1988) *Gene*, **66**, 135–146.
Seidel,U. and Arnold,H.H. (1989) *J. Biol. Chem.*, **264**, 16109–16118.
Smith,D.B. and Johnson,K.S. (1988) *Gene*, **67**, 31–40.
Sternberg,E.A., Spizz,G., Perry,M.E. and Olson,E.N. (1989) *Mol. Cell. Biol.*, **9**, 594–601.
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.
Thayer,M.J., Tapscott,S.J., Davis,R.L., Wright,W.E., Lassar,A.B. and Weintraub,H. (1989) *Cell*, **58**, 241–248.
Thisse,B., Stoetzel,C., Gorostiza-Thisse,C. and Perrin-Schmitt, F. (1988) *EMBO J.*, **7**, 2175–2185.
Villares,R. and Cabrera,C.V. (1987) *Cell*, **50**, 415–424.
Wright,W.E., Sassoon,D. and Lin,V.K. (1989) *Cell*, **56**, 777–783.
Yaffe,D. and Saxel,O. (1977a) *Nature*, **270**, 725–727.
Yaffe,D. and Saxel,O. (1977b) *Differentiation*, **7**, 159–166.

Received November 20, 1989