2234–2241 *Nucleic Acids Research, 1994, Vol. 22, No. 12* © 1994 Oxford University Press

MyoD1 promoter autoregulation is mediated by two proximal E-boxes

Jean-Marc Zingg, Gustavo Pedraza-Alva and Jean-Pierre Jost*
Friedrich Miescher Institute, PO Box 2543, 4002 Basel, Switzerland

Received April 8, 1994; Revised and Accepted May 26, 1994

ABSTRACT

We show that in mouse myoblasts the MyoD1 promoter is highly stimulated by MyoD1 expression, suggesting that it is controlled by a positive feedback loop. Using deletion and mutation analyses, we identified the targets for MyoD1 promoter autoregulation as the two proximal E-boxes located close to the MyoD1 core promoter. Gel mobility shift competition assays with MyoD1 antibodies as competitor suggest that the MyoD1 protein is binding directly to these E-boxes. Autoregulation did not occur in fibroblasts cotransfected with the expression vector of MyoD1. It is assumed that autoregulation is controlled by the stoichiometry between the MyoD1 protein and negatively regulatory proteins like Id, which is known to be highly expressed in fibroblasts. When the MyoD1 promoter was methylated, autoregulation only occurred when the density of methylated sites was low. The density of DNA methylation, therefore, can determine the accessibility of the MyoD1 promoter to transcription factors and interfere with the auto- and crossregulatory loop. The MyoD1 promoter *in vivo* was found to be only partially methylated in all tissues tested except in skeletal muscle where it was demethylated. We propose that high level expression of the MyoD1 gene is a result of release from constraints such as negative regulatory factors and/or DNA methylation interfering with MyoD1 autoregulation.

INTRODUCTION

Muscle development has been described as a two step process of determination of pluripotential precursor cells to the myogenic lineage and of differentiation of myoblasts to multinucleated myotubes and myofibers (for reviews see 1–4). A family of myogenic regulatory factors, MyoD1, Myogenin, Myf-5 and MRF-4 is involved both in myogenic determination as well as in myogenic differentiation (5–8). Each of these myogenic proteins is able to convert cells derived from all three embryonic layers to the myogenic lineage, and some members directly transactivate muscle-specific genes (9–16). Overexpression of exogenous myogenic genes in non-muscle cells also leads to activation of the endogenous myogenic genes (10,16). The potential of some myogenic genes to autoactivate and autoregulate each other's expression and, therefore, to form a network of interactive regulation ultimately may stabilise the myogenic determination state and prevent transdetermination of myoblasts.

Overexpression of the MyoD1 protein activates its own expression as well as the expression of the other members of the myogenic family and converts various cell types to the myogenic lineage (10,17). Since the frequency of myogenic conversion is dependent on the cell history of the transfected cell, MyoD1 may require additional factors or specific modifications such as phosphorylation for its action (17–21). Indeed, MyoD1 is known to be phosporylated and in the case of the homologous Myogenin protein, phosphorylation of a specific site in the DNA binding domain prevents DNA binding and myogenic activity (22,23). Furthermore, MyoD1 forms transcriptionally active dimers with proteins such as E12 or E47, inactive dimers with proteins such as Id or c-jun and an active myogenic complex with pRb (24–27). The stoichiometry of activating and inactivating dimerisation partners in a given cell may ultimately determine the myogenic potential of the MyoD1 protein. A cell line such as the fibroblast cell line C3H/10T1/2 which converts relatively easily to the myogenic lineage may represent a stage of differentiation closer to myoblasts than the HELA cell line which is difficult to convert (18,19). Treatment of the transfected HELA cell line with the DNA demethylating agent 5-azacytidine enhances the myogenic conversion suggesting that demethylation increases the accessibility of regulatory sites to the MyoD1 protein and/or activates further regulatory proteins involved in myogenic conversion (18,19).

In this report we show that the MyoD1 promoter contains regulatory elements which autoregulate its own expression. Using deletion and mutation analyses of a MyoD1 promoter–luciferase construct we localised the elements required for MyoD1 promoter autoregulation. We furthermore analysed, both in myoblasts and in fibroblasts, the ability of the autoregulatory elements and of the SV40 enhancer to activate the MyoD1 promoter from a distance. The potential of the MyoD1 protein to autoregulate its own expression in fibroblasts was low compared to myoblasts suggesting that fibroblasts lack a myoblast-specific coactivator or contain a fibroblast-specific inhibitor. The effect of DNA methylation on MyoD1 promoter autoregulation was studied by transfection of an *in vitro*

*To whom correspondence should be addressed

methylated MyoD1 promoter−luciferase construct. To elucidate the role of DNA methylation for the organ-specific expression of the MyoD1 gene, we determined the DNA methylation pattern of the MyoD1 gene in various tissues of the mouse. The relevance of MyoD1 autoregulation and of DNA methylation in the MyoD1 promoter to establishment and maintenance of the specific identity of a cell is discussed.

MATERIALS AND METHODS

Plasmid constructions

MyoD1 promoter fragments were cloned either into pGL2-basic (b) or into pGl2-enhancer (e) plasmids containing the firefly luciferase reporter gene (LUC) (Promega). The nomenclature used for plasmid construction is pMOb/e,where pMOb/− indicates cloning into pGl2-basic and pMO−/e indicates cloning into pGl2-enhancer. Plasmid pMO54/56 is the HindIII/BsmI (with HindIII-linker) MyoD1 promoter fragment cloned into the HindIII site of pGl2-b/e, respectively. Plasmid pMO60/61 is the PvuII/BglII(filled with Klenow) fragment deleted from pMO54/56. Plasmid pMO66/67 is the XbaI (filled with Klenow)/BsmI (with HindIII-linker) MyoD1 promoter fragment cloned into the SmaI/HindIII site of pGl2-b/e, respectively. Plasmid pMO68/69 is the AccI (filled with Klenow)/BsmI (with HindIII-linker) MyoD1 promoter fragment cloned into the SmaI/HindIII site of pGl2-b/e, respectively. Plasmid pMO58/59 corresponds to plasmid pMO54/56 deleted for the SmaI/SmaI MyoD1 promoter fragment. Plasmid pMO80 is plasmid pMO60 containing the oligonulceotide O1 cloned into the BamHI site whereas pMO82 contains the same oligonucleotide O1 cloned into the BamHI site of pMO68. Plasmids pMO70, pMO72 pMO74, pMO76 and pMO78 contain the oligonucleotides O2, O3, O4, O5 and O6 cloned into the BglII/PvuII site of pMO54, respectively. Plasmid pMO19, plasmid pMO19anti containing the MyoD1 gene in the inverted orientation and plasmid pMO29 are described in Zingg *et al.* (28). Plasmid pMO37 is the AccI/BsmI (both sites filled with Klenow) MyoD1 promoter fragment and plasmid pMO25 is the PvuII/BsmI (filled with Klenow) MyoD1 promoter fragment cloned into the HindIII/BamHI (both sites filled with Klenow) sites of plasmid pPALU (29).

Oligonucleotides

Oligonucleotides were synthesised with an Applied Biosystems oligonucleotide synthesiser. Only the upper strands are shown, E-boxes are underlined and mutated bases are shown in lower case letters.

O1: 5′*GATC*GATAGCCAAGTGCTACCGCGTGTGGCTGCCAGTCTCTCT-GCCCTCCTTCCTAGCTAGGCAGCTGCCCCA3′
O2: 5′*GATC*GATAGCgtAGTGCTACCGCGTGTGGCTGCCAGTCTCTCTG-CCCTCCTTCCTAGC'TAGGCAG3′
O3: 5′*GATC*CAAGCTTGTGGCTGCCAGTCTCTCTGCCCTCCTTCCTAG-CTAGGCAG3′
O4: 5′*GATC*CAAGCTTTGCCCTCCTTCCTAGCTAGGCAG3′
O5: 5′*GATC*CAAGCTTGTGGCTGCCAGTCTCTCTGCCCTCCTTCCTAG-CTAGGgtG3′
O6: 5′*GATC*GATAGCgtAGTGCTACCGCGTGTGGCTGCCAGTCTCTCTG-CCCTCCTTCCTAGCTAGGCAG3′
O7: 5′GGCTGCTAGGCAGCTGCCCCAG3′

Cell culture, DNA transfections and luciferase assays

G8 myoblasts and C3H/10T1/2 fibroblasts were cultured as described previously (28). Plasmid DNA (5 μg) was transfected using liposomes (DOTAP) according to the manufacturer's instructions (Boehringer-Mannheim). As internal control, 1 μg of the plasmid pCH110 (Pharmacia) expressing the β-galactosidase gene under the control of the SV40 early promoter was used. Extracts were prepared as described previously (28,30) and luciferase activity was measured in a luminometer (Autolumat LB 953 EG, Berthold) according to de Wet *et al.* (30). β-galactosidase activity was measured with a microtiter plate reader as described by Lucibello *et al.* (31). Each bar diagram represents the average of 3 measurements normalised to the β-galactosidase activity obtained from the internal control and the standard deviations of the mean are indicated by error bars.

Nuclear extract preparation and gel mobility shift assay

Nuclear extract were prepared as described by Zingg *et al.* (28). Gel mobility shift assays were performed as described by Pedraza *et al.* (32) using the oligonucleotide O7 containing the E-box P1. Immunobandshift experiments were performed by preincubation of the antibody with muscle nuclear extracts for one hour at 4°C before the gel mobility assay was performed as mentioned above. MyoD1 anti-peptide antibody 169 was generated using a peptide from the amino-terminal region of the MyoD1 protein (amino acids 1−25), and the polyclonal antibody 160 was generated using a peptide covering the basic region of the MyoD1 protein (amino acids 102−126). The specificity of the antibodies was assayed by immunoblots.

DNA methylations

Plasmid DNA was methylated with SssI, HpaII or HhaI according to the manufacturer's instructions (NewEngland Biolabs). Completion of the methylation reaction was checked by digestion with the isoschizomeric restriction enzymes and separation of the products on an agarose gel.

Isolation of genomic DNA and Southern blotting

Genomic DNA from tissues was isolated according to Saluz and Jost (33). To assay the *in vivo* methylation pattern 15 μg of genomic DNA was digested overnight in 300 μl final volume with 150 units (6 base cutter) or 200 units (4 base cutter) of the appropriate restriction enzyme. The DNA was purified by phenol extraction, ethanol precipitation, separated over an agarose gel and capillary blotted to a Hybond N (NEN) membrane. Prehybridisation and hybridisation were done using the conditions outlined by Church and Gilbert (34). Either the HindIII/SmaI MyoD1 promoter fragment (probe X) or the SmaI/SmaI MyoD1 gene fragment (probe Y) labelled by random priming was used as the probe (see Fig.7 for location of restriction sites). Complete digestion was confirmed when the same blot was hybridised with the second probe and found to be fully demethylated.

RESULTS

A fragment containing the two proximal E-boxes of the MyoD1 promoter mediates MyoD1 autoregulation in myoblasts but not in fibroblasts

When the construct pMO54 containing the HindIII/BsmI fragment of the MyoD1 promoter linked to the luciferase reporter gene (LUC) of plasmid pGL2-basic was cotransfected with a MyoD1 expression vector (pMO19) into G8 myoblasts, MyoD1 promoter activity was increased about 30 fold suggesting that the target(s) for MyoD1 promoter autoregulation are located within the MyoD1 promoter sequence (Fig.1A). Transfection of a vector expressing the antisense to the MyoD1 gene (pMO19anti) did

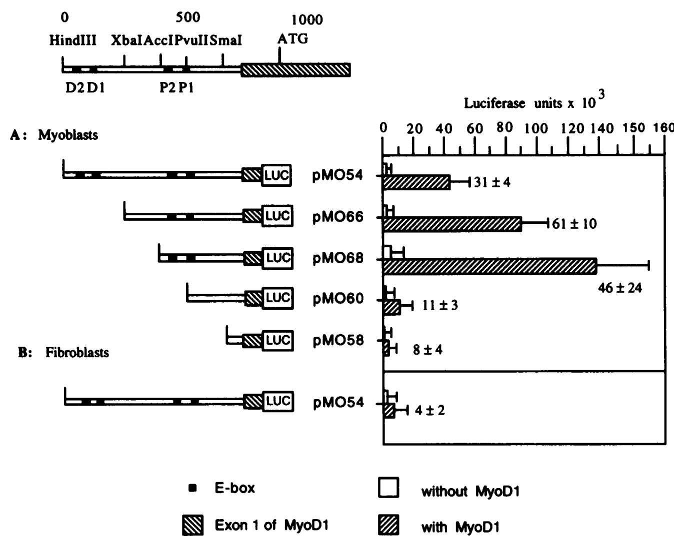

Figure 1. Localisation of the elements for the autoregulation of the MyoD1 promoter. The MyoD1 deletion constructs were transfected with or without the MyoD1 expression vector pMO19 into G8 myoblasts (A) or into C3H/10T1/2 fibroblasts (B) as outlined in Materials and methods. The numbers on the side of the bar diagrams are the fold increase in luciferase activity (with the standard deviation) obtained from transfection experiments with and without cotransfected MyoD1 expression vector pMO19.

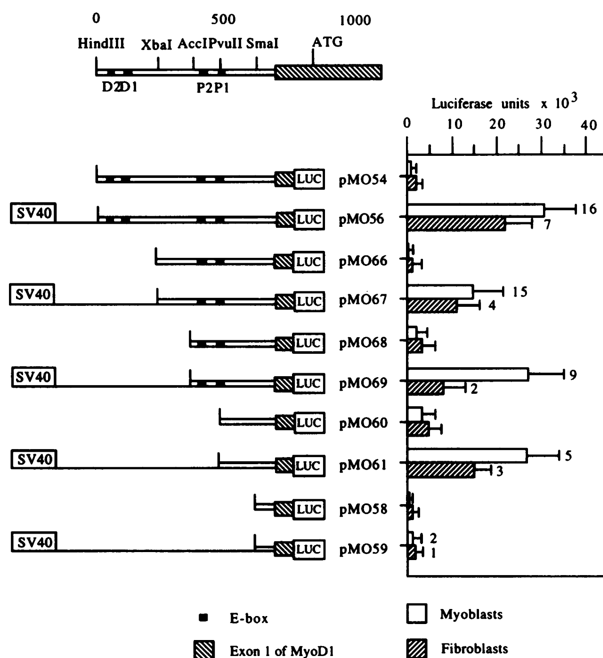

Figure 2. Effect of the SV40 enhancer on the activity of the MyoD1 promoter deletion mutants transfected either into G8 myoblasts (open boxes) or into C3H/10T1/2 fibroblast (hatched boxes). The numbers on the side of the bar diagrams are the fold increase in luciferase activity (with the standard deviation) obtained from transfection experiments with and without SV40 enhancer.

not have any effect on MyoD1 promoter activity (data not shown). MyoD1 autoregulation may be either directly mediated by one or several of the E-boxes previously identified (28), or indirectly mediated by regulatory factors induced by MyoD1 expression. Progressive deletion analysis was used to evaluate the mechanism of the observed induction by MyoD1. As shown in Figure 1A, removal of the two distal E-boxes D1 and D2 (pMO66 and pMO68) did not impair MyoD1 autoregulation, whereas removal of the two proximal E-boxes P1 and P2 (pMO60 and pMO58) significantly reduced but did not completely abolish induction of MyoD1 promoter activity. The residual level of MyoD1 promoter activation with plasmids pMO60 and pMO58 lacking any E-box can be explained either by MyoD1 dependent induction of the M-CAAT protein (35), or by an unspecific effect of E-boxes present in the plasmid pGl2-basic (there are 16 consensus E-boxes in pGl2-basic).

Interestingly, when the same constructs were transfected into C3H/10T1/2 fibroblasts, only a weak induction to a maximum level of four fold was observed by cotransfecting the MyoD1 expression vector (Fig. 1B). Since stable transfection of the same MyoD1 expression vector converted fibroblasts to myoblasts (28) and since the MyoD1 protein could be detected on western blots using protein extracts from transiently transfected fibroblasts (data not shown), artifacts caused by the absence of expression of MyoD1 in fibroblasts can be excluded. However, the SV40 promoter present in the plasmid pCH110 into which the MyoD1 gene was cloned is only about half as active in fibroblasts as compared to myoblasts. This level of MyoD1 expression may not be sufficient to titrate negative regulatory proteins, such as Id, present in fibroblasts. Since the endogenous MyoD1 protein is expressed in myoblasts and the level of Id is lower than in fibroblasts the threshold level required for autoregulation may be more easily reached than in fibroblasts (36,37). Indeed, for transactivation of muscle genes in fibroblasts it was critical that either the Id level was reduced by serum removal or that the MyoD1 and the E47 proteins were linked thus preventing the formation of inactive MyoD1/Id and/or E47/Id heterodimers (38).

The SV40 enhancer stimulates the MyoD1 promoter both in myoblasts and in fibroblasts

To evaluate whether the absence of MyoD1 promoter autoregulation in fibroblasts is due to a non-functional MyoD1 core promoter, the MyoD1 deletion constructs were linked to the SV40 enhancer. As shown in Figure 2, the SV40 enhancer was able to induce basal MyoD1 promoter activity both in myoblasts as well as in fibroblasts (pMO54/56). Deletion up to the PvuII site did not significantly lower enhancer action (pMO66/67, pMO68/69 and pMO60/61). However, further deletion to the SmaI site nearly abolished any effect obtained from the enhancer (pMO58/59). Therefore, regulatory elements in the PvuII/SmaI fragment are mostly responsible for the stimulatory effect of the SV40 enhancer. The PvuII/SmaI fragment harbours an AP2 site, two SP1 sites, a muscle CAAT-box and a CAAT box (28). It is very likely that the SV40 enhancer interacts with the AP2 site and the two SP1 sites in a manner similar to interactions with the SV40 early promoter (39). The stimulation of the MyoD1 core promoter by the SV40 enhancer in fibroblasts indicates that the core promoter is functional and that the absence of autoregulation is most likely due to negative regulatory factors interfering with autoregulation.

The proximal two E-boxes P1 and P2 do not act as a MyoD1 stimulatory element when inserted at a distance from the MyoD1 core promoter

As indicated above, the AccI/PvuII fragment containing the two proximal E-boxes P1 and P2 is sufficient for the

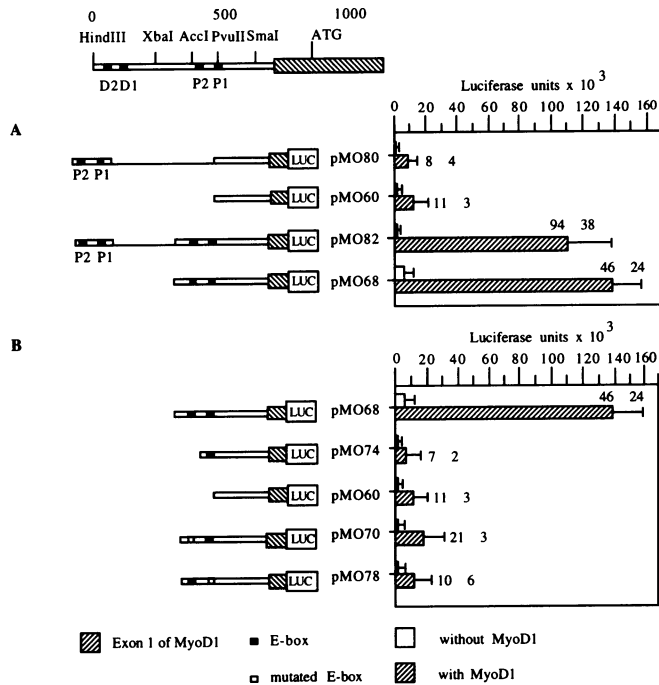

Figure 3. (**A**) Effect of the position of the two E-boxes P1 and P2 on the autoregulation by MyoD1 in the presence or absence of the two original proximal E-boxes. The MyoD1 deletion constructs were transfected with or without the MyoD1 expression vector pMO19 into G8 myoblasts as outlined in Materials and methods. The numbers on the side of the bar diagrams are the fold increase in luciferase activity (with the standard deviation) obtained from transfection experiments with or without cotransfected MyoD1 expression vector pMO19. (**B**) Effect of deletion or mutation of individual E-boxes P1 or P2. The constructs with deletion or mutations were made as described in methods and the experiments were performed as described above.

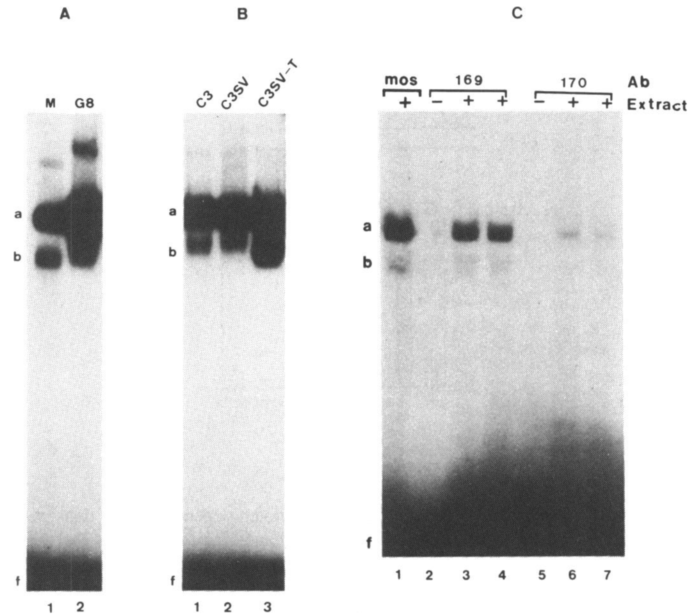

Figure 4. MyoD1 binds to the E-box P1 located in the MyoD1 promoter. Gel mobility shift assay with an oligonucleotide containing the proximal E-box P1. The upper band (a) corresponds to the E12/E47 heterodimer (38) and the lower band (b) to the MyoD1/E12 or MyoD1/E47 heterodimer (38). (**A**) Lane 1: hindlimb skeletal muscle nuclear extract (M), lane 2: G8 myoblasts nuclear extract (G8). (**B**) Lane 1: C3H/10T1/2 fibroblasts nuclear extract (C3); lane 2: nuclear extract from C3H/10T1/2 fibroblasts stable transfected with the MyoD1 expression vector pMO19 (C3SV); lane 3: as lane 2 but grown in differentiation medium (DMEM containing 2% horse serum) (C3SV-T). (**C**) Gel mobility shift competition assays with specific MyoD1 antibodies as competitors. Lane 1: control with mos antibody (mos); Lane 2 and 5; control with incubation of antibody and probe alone; Lane 3 and 4: nuclear extracts were preincubated with antibody 169 at two different dilutions (1:50 and 1:100, respectively); Lane 6 and 7: nuclear extracts were incubated with antibody 170 at two different dilutions (1:50 and 1:100, respectively). f is the free DNA.

MyoD1-dependent stimulation whereas removal of the two distal sites D1 and D2 had no effect (Fig.1A). To evaluate whether the two proximal E-boxes contain a specific sequence that allows autoregulation of the MyoD1 promoter in myoblasts or whether the distance of the two E-boxes from the core promoter is important, an oligonucleotide containing the two proximal E-boxes P1 and P2 was inserted into the BamHI site at about 2.9 kb upstream of the MyoD1 promoter. As shown in Figure 3A, the AccI/PvuII fragment containing the two E-boxes P1 and P2 placed 2.9 kb upstream of the promoter did not stimulate MyoD1 promoter autoregulation (pMO80, pMO60) in the absence of the two proximal E-boxes in the MyoD1 core promoter. Furthermore, when the two proximal E-boxes (pMO82, pMO68) were present in the MyoD1 promoter, the upstream insertion of the AccI/PvuII fragment did not increase the absolute level of MyoD1 promoter activity (Fig.3A). Since the original two distal E-boxes D1 and D2 did not result in MyoD1 dependent promoter stimulation (Fig.1A, pMO54 and pMO66) and the proximal E-boxes P1 and P2 only stimulated MyoD1 promoter activity when located close to the core promoter, the distance of E-boxes from the MyoD1 core promoter is of paramout importance to achieve the autoregulation by MyoD1.

The two proximal E-boxes P1 and P2 cooperatively autoregulate the MyoD1 promoter

To evaluate whether the two proximal E-boxes alone or together are involved in MyoD1 promoter autoregulation, further deletion and point mutation analyses were carried out. Oligonucleotides with either the E-box P1 or P2 deleted or mutated were inserted into the promoter test plasmid and transfected with and without MyoD1 expression vector pMO19. As shown in Figure 3B, deletion of the E-box P2 (pMO74) significantly reduced the autoactivation potential whereas further deletion of the second E-box P1 (pMO60) did not further reduce the autoactivation by MyoD1. Similar results were obtained when either E-box P1 (pMO78) or P2 (pMO70) was specifically mutated (Fig.3B), thus corroborating that E-box P1 or P2 alone is not sufficient for autoregulation and that sequences between the two E-boxes do not contribute to autoregulation of MyoD1.

The MyoD1 protein binds to the proximal two E-boxes of the MyoD1 promoter

To further evaluate whether it is indeed the MyoD1 protein that binds to the two proximal E-boxes P1 and P2 and not another protein that is indirectly induced by MyoD1 expression, the specificity of binding was analysed by gel mobility shift assay. As shown in Figure 4A, two major protein complexes were observed (a and b) in gel mobility shift assays with an oligonucleotide (O7) containing the proximal E-box P1 using nuclear extracts of hindlimb skeletal muscle (lane 1, M) and proliferating G8 myoblasts (lane 2, G8). These complexes could be specifically competed with non-labelled oligonucleotides

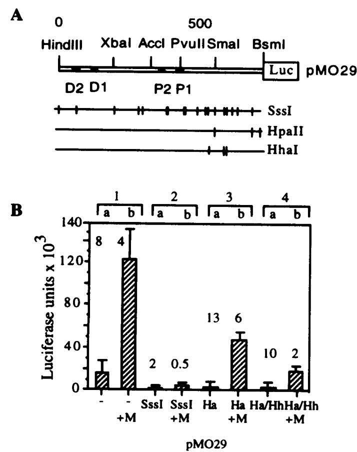

Figure 5. Autoregulation of the MyoD1 promoter depends on the density of methylated CpGs. **(A)** Map of plasmid pMO29 and location of CpGs that were methylated by the DNA methylases SssI, HpaII or HhaI. **(B)** Plasmid pMO29 (see map in A) containing the HindIII/BsmI MyoD1 promoter fragment was transfected into G8 myoblasts and luciferase activity measured as outlined in Materials and methods. The plasmid was methylated by either CpG methylase (SssI), HpaII methylase (Ha) or by HpaII and HhaI methylase (Ha/Hh). The numbers above the bar diagrams are the fold increase in luciferase activity (with the standard deviation) obtained from transfection experiments with (+M) or without cotransfection of the MyoD1 expression vector pMO19.

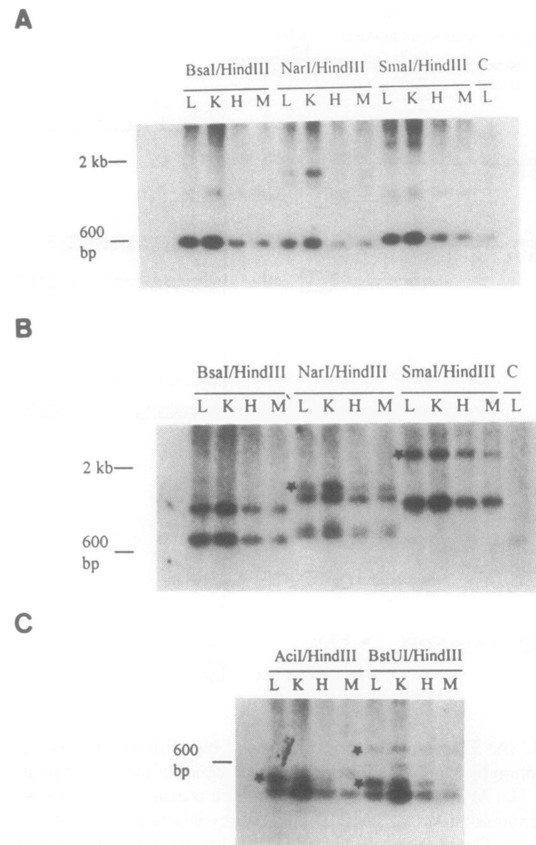

Figure 6. State of methylation of the mouse MyoD1 gene. Southern blot of genomic DNA isolated from various tissues and digested with methyl-CpG sensitive restriction enzymes. For hybridisation the probes indicated in Figure 7 were used; probe X (for Figure 6A and 6C) or probe Y (for Figure 6B). Partially methylated sites are marked with a star. The numbers in brackets behind the enzyme name indicate the location of the restriction site as shown in Figure 7. **(A)** Sites that are in all tissues unmethylated: BsaHI (4), NarI (7) and SmaI (5). **(B)** Sites that are in all tissues unmethylated: BsaHI (11); or that are in all tissues partially methylated: NarI (12) and SmaI (13). **(C)** Sites that are specifically demethylated in skeletal muscle: AciI (1) and BstUI (1 and 2). L: liver; K: kidney, H: heart, and M: skeletal muscle genomic DNA. C: control DNA (liver DNA digested with the enzyme MspI).

containing the E-boxes P1, P2 and the muscle creatine enhancer (data not shown). To define which complex was due to MyoD1 binding, nuclear extracts were prepared from C3H/10T1/2 fibroblasts and from fibroblasts stably transfected with the MyoD1 expression vector pMO19. As shown in Figure 4B, extracts from C3H/10T1/2 fibroblasts (C3) and extracts from fibroblasts stably transfected with the MyoD1 expression vector pMO19 (C3SV) did not give a specifically shifted band b. When the serum was removed from C3SV fibroblasts (C3SV-T), a strong increase of complex b formation was observed, possibly indicating that the Id protein is downregulated by serum removal (9,36,37) and the free MyoD1 protein can bind. The specificity of binding to the oligonucleotide was further assayed by gel mobility shift competition analysis with both E-boxes P1 and P2 (data not shown) and by immunobandshift analysis with antibodies derived against different peptides of the MyoD1 protein (Fig.4C). Complex b could weakly be competed by using a polyclonal antibody against the aminoterminal end of the MyoD1 protein (Fig.4C, lane 3 and 4), but not by an anti-c-mos antibody (lane1). An antibody against the basic domain of the MyoD1 protein (antibody 170) strongly competed in addition to complex b also with complex a (Fig.4C, lane 6 and 7), possibly indicating cross-reactivity with the basic domains of the E12/E47 proteins. All together our data indicate that the MyoD1 protein binds directly to the E-boxes P1 and P2 and is most likely directly responsible for the observed stimulation of the MyoD1 promoter.

The density of DNA methylation determines the susceptibility of the MyoD1 promoter to autoregulation

It has been shown that transfection of the MyoD1 gene into C3H/10T1/2 fibroblasts induces the endogenous MyoD1 gene suggesting that MyoD1 does not only autoregulate but also autoactivates its own expression (10,17). The endogenous gene in C3H/10T1/2 fibroblasts is methylated and not expressed and therefore, the MyoD1 protein must be able to neutralise the inactivating effect of DNA methylation (40). This hypothesis was tested by transfecting the MyoD1 promoter, methylated to various degrees with SssI CpG methylase or site-specific methylases (HpaII and/or HhaI), with and without the MyoD1 expression vector pMO19. Figure 5A shows the plasmid pMO29 and the sites methylated in the MyoD1 promoter by SssI, HpaII and HhaI methylases. As shown in Figure 5B, the unmethylated plasmid pMO29 is induced 8 fold by MyoD1 expression (lane1a/b). Methylation of all CpGs reduced MyoD1 promoter activity nearly to background levels and expression of MyoD1 only very weakly reactivated the MyoD1 promoter (lane 2a/b). However, when plasmid pMO29 was methylated only at specific sites either with HpaII alone (lane3a/b) or with HpaII and HhaI (lane4a/b), the MyoD1 promoter still could be partially activated by MyoD1 expression. These results indicate that the MyoD1 protein is able to autoactivate its own expression provided the density of methylated CpGs is not too high.

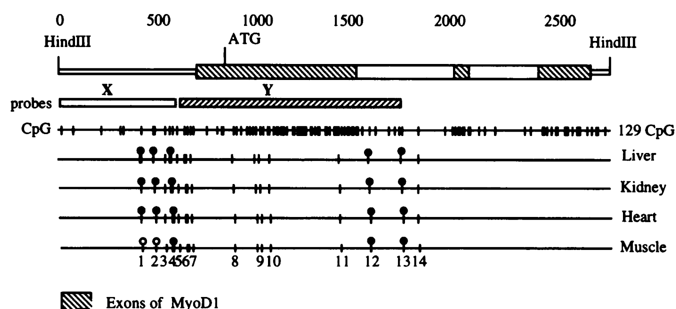

Figure 7. Summary of DNA methylation pattern in genomic DNA of liver, kidney, heart and skeletal muscle. The DNA was first digested with HindIII and then with one of the following enzymes; (1) AciI or BstUI; (2) AciI or BstUI; (3) AciI; (4) AciI, BsaHI, HhaI, NarI; (5) HpaII, SmaI; (6) BstUI, HhaI, NarI; (7) BsaHI, HpaII; (8) AciI, HpaII; (9) BstUI, HhaI; (10) HpaII; (11) BsaHI; (12) NarI; (13) HpaII, SmaI; (14) BstUI, HhaI. The Southern blots were hybridised with either probe X or probe Y. Sites that are partially methylated in all tissues tested are marked by a filled circle and the corresponding enzymes are underlined in the legend. Sites that are demethylated in a tissue-specific way are marked by an open circle. Sites that are unmethylated in all tissues are shown by a bar.

In vivo the MyoD1 promoter is only partially methylated but is specifically demethylated in skeletal muscle

As shown in Figure 5, DNA methylation has the potential to reduce or completely inactivate the MyoD1 promoter. Since promoter inactivation was observed in cell lines of various tissue origin (G8 myoblasts, C3H/T101/2 fibroblasts and Nb neuroblastoma cells, data not shown), promoter inactivation is probably not mediated by cell type-specific regulatory factors but rather a result of ubiquitous methyl-DNA binding proteins or inactive chromatin formation. It was therefore of interest to test whether the tissue-specific *in vivo* MyoD1 expression pattern is reflected also by a tissue-specific DNA methylation pattern, which may enforce inactive chromatin formation in non-expressing tissues and prevent the expression of the MyoD1 gene in inappropriate tissues.

Genomic DNA from liver, kidney, heart and skeletal muscle was isolated, digested with various methyl-CpG sensitive restriction enzymes and the MyoD1 gene detected by Southern blotting. As shown in Figure 6A, most sites detected were unmethylated in all four tissues, some sites were partially methylated in all tissues tested (Fig.6B) and some partially methylated sites were specifically demethylated in skeletal muscle tissues (Fig.6C). A summary of all sites tested is shown in Figure 7. It is unknown whether and how the partially methylated sites are involved in the *in vivo* tissue-specific expression pattern of the MyoD1 gene. However, the fact that some sites are specifically demethylated in skeletal muscle tissue may reflect that this subset of sites, possibly together with other sites, is involved in the tissue-specific silencing of the MyoD1 gene in non-expressing tissues.

DISCUSSION

Although the MyoD1 gene is not essential for *in vivo* muscle determination and differentiation, its skeletal muscle-specific expression makes it a good model to study the mechanisms responsible for cell type- and tissue-specific expression (41). As discussed by Blau (42), the mechanisms for tissue-specific gene expression may be actively regulated by a network of interacting positive or negative ubiquitous and tissue-specific regulatory factors that ultimately result in the cell type-specific expression pattern of a gene. In addition, the expression state of a gene might be influenced by the specific chromatin structure present on intact chromosomes which is possibly established by distant elements like Locus Control Regions and/or cell type-specific DNA methylation patterns (43–50).

Genes encoding transcriptional regulatory factors such as the MyoD1 gene may harbour a further level of regulation which is mediated by autoregulation (10,17,51–53). The advantage of autoregulation with respect to cell type-specific gene expression may lie in the possibility to generate a regulatory switch, such that unspecific low level expression of a gene does not lead to activation whereas an increased cell type-specific expression activates the autoregulatory loop and increases its amplitude. In this way, genes encoding regulatory factors stabilise their own expression by autoregulation and/or as in the case of the myogenic family, stabilise a whole network of regulatory genes by auto- and crossregulation (10,17,54). In this model it is open to debate as to what kind of mechanism initially activates the gene to a low level. Possibly, a combination of the activity and stoichiometry of positive and negative regulatory factors and a gene-specific opening of chromatin and demethylation is involved in reaching the threshold level and activating the autoregulatory loop. In the case of MyoD1 autoregulation, negative regulatory factors such as the proteins Id or c-jun, destabilisation of the MyoD1 mRNA or protein, modification of the MyoD1 protein and/or the density of DNA methylation as described here may interfere with autoregulation and prevent the switching of the MyoD1 gene expression states.

We show here that autoregulation of the MyoD1 gene is dependent on the cell type and very likely requires mechanisms and/or factors in addition to the presence of the MyoD1 protein. Since fibroblasts convert into fully functional myoblasts and myotubes by stable transfection of the MyoD1 gene (28), these additional mechanisms and regulatory factors are possibly controlled by MyoD1 itself but need for their appearance a longer period of exposure to the MyoD1 protein than our transient transfection protocol allows. A simpler explanation would be, that in fibroblasts the amount of negative regulatory factors such as Id or jun is higher than in myoblasts and that the level of the transiently transfected MyoD1 gene is not sufficient to titrate these factors (36,37). The absence of autoregulation in fibroblasts could also be explained by the lack of a coactivator of MyoD1 or the absence of a specific modification of the MyoD1 protein. However, since autoregulation occurs in stably transfected fibroblasts (28), the absence of autoregulation can be easiest explained by an insufficient accumulation of the MyoD1 protein to titrate negatively acting factors such as the protein Id in transiently transfected fibroblasts.

Using deletion and point mutation analyses we show that the E-boxes P1 or P2 alone are not sufficient for autoregulation. Rather, autoregulation is a result of the two E-boxes P1 and P2 together, which cooperatively mediate MyoD1 dependent autoregulation in ways similar to activation of the acetylcholine receptor delta subunit gene promoter (55), the myosin light chain promoter (56) or the muscle creatine kinase promoter (57). When the two E-boxes P1 and P2 were removed from their location close to the MyoD1 core promoter and inserted at a distance, MyoD1 autoregulation was impaired. Therefore, the fragment

containing the two E-boxes does not contain a bona fide enhancer since its stimulatory effect was dependent on the distance from the MyoD1 core promoter. Using gel mobility shift competition assays with MyoD1 antibodies we show that it is indeed the MyoD1 protein which binds to the proximal E-boxes. Further studies will be required to evaluate the specific characteristics of the fragment containing the E-boxes P1 and P2 that enable it to stimulate MyoD1 dependent cooperative autoregulation. When the SV40 enhancer was inserted at a distance, promoter stimulation was independent of the two E-boxes, but was dependent on the MyoD1 core promoter fragment containing the two SP1 boxes. The SV40 enhancer also stimulated the MyoD1 promoter in fibroblasts, which makes it unlikely that fibroblasts express a repressor protein binding to the MyoD1 core promoter thus preventing its activation.

DNA methylation has been implicated in the stabilisation of the inactive expression state of many genes (43–50,58). Whether activation of a gene precedes demethylation or the opposite is still under debate. However, during the activation of an inactive and methylated promoter sequence transcription factors must gain access to the promoter sequence which is possibly sequestered by specific methyl-DNA binding proteins such as MDBP-1/-2 (59,60) and MeCP2 (61,62) or by inactive chromatin formation (63). Some transcription factors have the ability to recognise their cognate target sequence when complexed in chromatin, others have not and require the opening and removal of chromatin (64). We show here that the MyoD1 protein activates a methylated MyoD1 promoter, provided not all the CpGs are methylated. The MyoD1 protein, therefore, recognises the MyoD1 promoter sequence only if there are gaps between the methyl-DNA-binding-proteins or nucleosomes. Therefore, in myoblasts and muscle tissues only genes with a low degree of methylation are accessible to the regulation by MyoD1. Similarly, the myogenic potential of the MyoD1 protein could be increased by treatment of keratinocytes transfected with a MyoD1 expression vector with the demethylating agent 5-azacytidine (65). In this way the gene expression pattern and identity of a cell is determined not only by the presence or absence of regulatory proteins in the nucleus, but also by the degree of DNA methylation and accessibility of a gene. Similar results have been described by Boyes and Bird (66) and reviewed by Bird (67), suggesting that the degree of inactivation is a result of the density of methylated CpGs in the promoter and of the promoter strength.

When the *in vivo* DNA methylation pattern of the MyoD1 gene was examined, only partially methylated sites were so far detected. Partially methylated sites either reflect the heterogeneity of cell types in a given tissue, chromosome-specific demethylations or chromosome-specific modifications. However, the observed skeletal muscle-specific demethylation of partially methylated sites could indicate a role of these sites in the tissue-specific expression of the MyoD1 gene. Partially methylated autoregulatory genes could show a level of expression insufficient to activate a positive feedback loop. If high expression of a gene is dependent on autoregulation, such a basal level of gene expression from a partially methylated gene may be required to switch to autoregulation as soon as the negative constraints such as negative regulatory proteins and/or methylated CpGs are removed. Demethylation may increase the basal level expression of the MyoD1 gene to a threshold level where autoregulation can occur. Although it is difficult to evaluate the role and origin of partially methylated sites, they have been found by other groups both in intact animals and in cell culture (68,69).

In conclusion the differentiation state of a cell may well be determined by an active network of cross- and autoregulatory proteins, that is, however, limited in its action by a gene and cell type-specific DNA methylation pattern. The threshold to activate gene expression may be increased by the number of methylated sites as well as by factors which negatively interfere with the autoregulatory or crossregulatory loop. During development the position and environment of a cell may specifically alleviate the negative constraints interfering with autoregulation and allow feedback regulation to occur. The mechanisms of setting up cell type-specific gene expression and DNA methylation patterns and their significance for cell identity are only starting to be elucidated.

ACKNOWLEDGEMENTS

We thank Drs P.Caroni, A.Raizis, L.Perez and W.M.Rideout III for suggestions and critically reading the manuscript.

REFERENCES

1. Rosenthal, N. (1989) *Curr. Op. Cell Biol.*, **1**, 1094–1101.
2. Emerson, C.P. (1990) *Curr. Op. Cell Biol.*, **2**, 1065–1075.
3. Stockdale, F.E. (1992) *Dev. Biol.*, **154**, 284–298.
4. Buckingham, M. (1992) *TIG*, **8**, 144–149.
5. Weintraub, H., Davis, R., Tapscott, S., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., Zhuang, Y., and Lassar, A. (1991) *Science*, **251**, 761–766.
6. Chen, J. and Jones, P.A. (1989) *Curr. Op. Cell Biol.*, **1**, 1075–1080.
7. Kingston, R.E. (1989) *Curr. Op. Cell Biol.*, **1**, 1081–1087.
8. Olson, E.N. (1990) *Genes Dev.*, **4**, 1454–1461.
9. Davis, R.L., Weintraub, H., and Lassar, A.B. (1987) *Cell*, **51**, 987–1000.
10. Thayer, M.J., Tapscott, S.J., Davis, R.L., Wright, W.E., Lassar, A.B., and Weintraub, H. (1989) *Cell*, **58**, 242–248.
11. Yutzey, K.E., Rhodes, S.J., and Konieczny, S.F. (1990) *Mol. Cell. Biol.*, **10**, 3934–3944.
12. Braun, T., Rudnicki, M.A., Arnold, H.H., and Jaenisch, R. (1992) *Cell*, **71**, 369–382.
13. Fujisawa-Sehara, A., Nabeshima, Y., Komiya, T., Uetsuki, T., and Asukara, A. (1992) *J. Biol. Chem.*, **267**, 10031–10038.
14. Li, H. and Capetanaki, Y. (1993) *Nucleic Acids Res.*, **21**, 335–343.
15. Cserjesi, P., Lilly, B., Bryson, L., Wang, Y., Sassoon, D.A., and Olson, E.N. (1992) *Dev.*, **115**, 1087–1101.
16. Piette, J., Bessereau, J.L., Huchet, M., and Changeux, J.P. (1990) *Nature*, **345**, 353–355.
17. Weintraub, H., Tapscott, S.J., Davis, R.J., Thayer, M.J., Adam, M.A., Lassar, A.B., and Miller, A.D. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5434–5438.
18. Blau, H.M. (1993) *TIG*, **5**, 268–272.
19. Schaefer, B.W., Blakely, B.T., Darlington, G.J., and Blau, H.M. (1990) *Nature*, **344**, 454–458.
20. Olson, E. N. (1992) *Dev. Biol.*, **154**, 261–272.
21. Li, L. and Olson, E. N. (1992) *Adv. Cancer Res.*, **58**, 95–119.
22. Tapscott, S.J., Davis, R.L., Thayer, M.J., Cheng, P.F., Weintraub, H., and Lassar, A.B. (1988) *Science*, **242**, 405–411.
23. Li, L., Zhou, J., James, G., Heller-Harrison, R., Czech, M.P., and Olson, E.N. (1992) *Cell*, **71**, 1181–1194.
24. 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. (1989) *Cell*, 58, 537–544.
25. Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. (1990) *Cell*, **61**, 49–59.
26. Li, L., Chambard, J.C., Karin, M., and Olson, E.N. (1992) *Genes Dev.*, **6**, 676–689.
27. Gu, W., Schneider, J.W., Condorelli, G., Kaushal, S., Mahdavi, V., and Nadal-Ginard, B. (1993) *Cell*, **72**, 309–324.
28. Zingg, J.M., Pedraza Alva, G., and Jost, J.P. (1991) *Nucleic Acids Res.*, **19**, 6433–6439.
29. Artelt, P., Grannemann, R., Stocking, C., Friel, J., Bartsch, J., and Hauser, H. (1991) *Gene*, **99**, 249–254.

30. De Wet, J.R., Wood, K.V., DeLuca, M., Helsinki, D.R., and Subramani, S. (1987) *Mol. and Cell. Biol.*, **7**, 725–737.
31. Lucibello, F.C. and Mueller, R. (1989) *Methods Mol. Cell. Biol.*, **1**, 9–18.
32. Pedraza, G., Zingg, J. M. and Jost, J. P. (1994) *J. Biol. Chem.*, **269**, 6978–6985.
33. Saluz, H.P. and Jost, J.P. (1993) In: *Biomethods: A Laboratory Guide for In Vivo Studies of DNA Methylation and Protein/DNA Interactions*. Birkhäuser Verlag Basel, **3**.
34. Church, G. M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**,1991–1995.
35. Lassar, A.B., Davis, R.L., Wright, W.E., Kadesch, T., Murre, C., Voronova, A., Baltimore, D., and Weintraub, H. (1991) *Cell*, **66**, 305–315.
36. Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. (1990) *Cell*, **61**, 49–59.
37. Christy, B.A., Sanders, L.K., Lau, L.F., Copeland, N.G., Jenkins, N.A., and Nathans, D. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1815–1819.
38. Neuhold, L. A. and Wold, B. (1993) *Cell*, **74**, 1033–1042.
39. Mitchell, P.J., Wang, C., and Tjian, R. (1987) *Cell*, **50**, 847–861.
40. Michalowsky, L.A. and Jones, P.A. (1989) *Mol. Cell. Biol.*, **9**, 885–892.
41. Rudnicki, M.A., Braun, T., Hinuma, S., and Jaenisch, R. (1992) *Cell*, **71**, 383–390.
42. Blau, H. M. (1992) *Annu. Rev. Biochem.*, **61**, 1213–1230.
43. Dillon, N. and Grosveld, F. (1993) *TIG*, **9**, 134–137.
44. Peterson, K. R. and Stamatoyannopoulos, G. (1993) *Mol. Cell. Biol.*, **13**, 4836–4843.
45. Bestor, T.H., Hellewell, S.B., and Ingram, V.M. (1984) *Mol. Cell. Biol.*, **4**, 1800–1806.
46. Young, P.R. and Tilghman, S.M. (1984) *Mol. Cell. Biol.*, **4**, 898–907.
47. Ysraeli, J., Adelstein, R.S., Melloul, D., Nudel, U., Yaffe, D., and Cedar, H. (1986) *Cell*, **46**, 409–416.
48. Albini, A., Toffenetti, J., Zhu, Z., Chader, G.J., and Noonan, D.M. (1990) *Nucleic Acids Res.*, **18**, 5181–5187.
49. Sasaki, T., Hansen, R.S., and Gartler, S.M. (1992) *Mol. Cell. Biol.*, **12**, 3819–3826.
50. Li, E., Beard, C. and Jaenisch, R. (1993) *Nature*, **366**, 362–365.
51. Serfling, E. (1989) *TIG*, **5**, 131–133.
52. Chen, R., Ingraham, H. A., Treacy, M. N., Albert, V. R., Wilson, L., and Rosenfeld, M. G. (1990) *Nature*, **346**, 583–586.
53. Schier, A. F. and Gehring, W. J. (1992) *Nature*, 356, 804–807.
54. Edmondson, D. G., Brennan, T. J., and Olson, E. N. (1991) *J. Biol. Chem.*, **266**, 21343–21346.
55. Simon, A. M. and Burden, S. J. (1993) *Mol. Cell. Biol.*, **13**, 5133–5140.
56. Wentworth, B.M., Donoghue, M., Engert, J.C., Berglund, E.B., and Rosenthal, N. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 1242–1246.
57. Weintraub, H., Davis, R., Lockshon, D., and Lassar, A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5623–5627.
58. Opdecamp, K., Riviere, M., Molne, M., Szpirer, J., and Szpirer, C. (1992) *Nucleic Acids Res.*, **20**, 171–178.
59. Ehrlich, M. and Ehrlich, K. C. (1993) In Saluz H. P. and Jost, J. P. (ed.), *DNA Methylation: Molecular Biology and Biological Significance*, Birkhäuser Verlag, Basel, 145–168.
60. Jost, J.P. and Hofsteenge, J. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 9499–9503.
61. Lewis, J.D., Meehan, R,R., Henzal, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992) *Cell*, **69**, 905–914.
62. Meehan, R.R., Lewis, J.D., and Bird, A.P. (1992) *Nucleic Acids Res.*, **20**, 5085–5092.
63. Fedor, M. J. (1992) *Curr. Op. Cell Biol.*, **4**, 436–443.
64. Hayes, J.J. and Wolffe, A.P. (1993) *Bioessays*, **14**, 597–603.
65. Chen, J. and Jones P. A. (1990) *Cell Growth Differentiation*, **1**, 383–392.
66. Boyes, J. and Bird, A. (1992) *EMBO*, **11**, 327–333.
67. Bird, A. (1992) *Cell*, **70**, 5–8.
68. Turker, M.S., Swisshelm, K., Smith, A.C., and Martin, G.M. (1989) *J. Biol. Chem.*, **264**, 11632–11636.
69. Yeivin, A. and Razin, A. (1993) Gene methylation patterns and expression. In Saluz, H. P. and Jost, J. P. (eds), *DNA Methylation: Molecular Biology and Biological Sigificance*, Birkhäuser Verlag Basel, 523–568.