ELSEVIER

Bone xx (2006) xxx–xxx

BONE

www.elsevier.com/locate/bone

Identification and characterization of the *Mustang* promoter: Regulation by AP-1 during myogenic differentiation

Cheng Liu, Michael Hadjiargyrou *

Department of Biomedical Engineering, Stony Brook University, Psychology A Building, Room 338, Stony Brook, NY 11794-2580, USA

Received 15 March 2006; revised 4 April 2006; accepted 8 April 2006

Abstract

We previously identified *Mustang* (*mu*sculo*s*keletal *t*emporally *a*ctivated *n*ovel *g*ene) with expression exclusively in the musculoskeletal system. Although its expression is almost undetectable in intact bone, it is robustly upregulated during bone regeneration. It is also abundantly expressed in adult skeletal muscle and tendon. As such, *Mustang* represents a marker for these cells and thus identifying its promoter would enable us to characterize its transcriptional regulation. To this end, we have isolated and characterized a 1512-bp mouse genomic clone representing the *Mustang* 5′-flanking region and identified a transcription start site, a TATA box, and multiple putative transcription factor binding sites (including AP-1 and AP-2). The activity of this promoter was detected in musculoskeletal cells and embryonic fibroblasts, even exceeding levels (145%) of the control SV40 promoter (in C2C12 cells). Further, the contribution of specific AP-1 and AP-2 sites was determined with serially deleted and mutated promoter constructs. Results indicate that one of the four AP-1 sites is required for substantial transcriptional activation, as its specific deletion or mutation decreases promoter activity by 32% and 40%, respectively. In contrast, deletion of both identified AP-2 sites results in only a 12% decrease in promoter activity. We further characterized the key AP-1 site by EMSA and determined that in both proliferating and differentiating C2C12 cells, only c-Fos, Fra-2 and JunD were required for transcriptional activation. *Mustang*'s restricted tissue specificity and strong promoter makes this gene an ideal candidate for utilization in cell lineage studies that could unveil cellular/molecular mechanisms responsible for musculoskeletal development and regeneration.

Keywords: *Mustang*; Promoter; AP-1; C2C12; Differentiation; Musculoskeletal

Introduction

Mustang was initially discovered through our expression profiling experiments of bone regeneration [1]. Previous studies have also identified the spatiotemporal expression of *Mustang* in embryonic mesenchymal cells, in osteoprogenitor cells of the periosteum, chondrocytes of articular cartilage, skeletal muscle, tendon, as well as in fracture callus of a healing bone [2]. More specifically, *Mustang* is dramatically upregulated during the early stages of fracture repair especially at PF (postfracture) Day 5 where its expression peaks with a 55-fold increase as compared to intact bone. Its expression then gradually decreases throughout the rest of the healing process (up to 3 weeks). Although much is known about *Mustang*'s restricted expression in the musculoskeletal system, little is known about its regulation or function. With the exception of a nuclear localization signal, no other structural motif appears in the *Mustang* open reading frame (ORF), thus making it very difficult to place it in a specific protein family. However, *Mustang*'s specificity to the various cell types of the musculoskeletal system and its expression patterns suggests a multifunctional role.

As a first attempt to probe the tissue-specific expression of *Mustang*, we focused on its transcriptional regulation by isolating and characterizing its promoter in myogenic C2C12 cells. C2C12 is a well-established and characterized pluripotent cell line that can be induced to differentiate into various cell types, including myocytes, osteoblasts and adipocytes [3–5]. The natural transition of these cells from their proliferation phase into differentiated myocytes (upon confluency in low-serum environment) has been widely used to study skeletal

* Corresponding author. Fax: +1 631 632 8577.
E-mail address: michael.hadjiargyrou@sunysb.edu (M. Hadjiargyrou).

8756-3282/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.bone.2006.04.002

muscle formation in vitro. We decided to utilize C2C12 cells in order to study the regulation of *Mustang* because of its strong expression in skeletal muscle and in the hope that identifying the regulatory factors involved in myoblast differentiation will facilitate understanding of *Mustang*'s role in this process and provide further clues regarding its function.

Herein, we report that the expression of *Mustang* increases as C2C12 cells undergo differentiation into myocytes and that this mRNA upregulation parallels the increase of skeletal muscle-specific transcription factors, MyoD and myogenin. We further report on the genomic organization of *Mustang* as well as the identification and cloning of a 1512-bp genomic sequence upstream of the gene representing its promoter. The activity of this putative promoter was tested via luciferase activity assays in various cells lines known to express *Mustang*, as well as some that do not. In addition, the activity of the promoter was fully characterized in C2C12 via serial deletions and mutations that show that the *Mustang* promoter represents a very strong promoter, even inducing higher transcription levels than the SV40 promoter (used as a positive control in these experiments). Further, we have identified a number of AP-1 and AP-2 transcription factor binding sites present within the *Mustang* promoter. Specifically, a single AP-1 site was found to be critical for essentially 40% of the promoter activity. Further, electrophoretic gel shift assays (EMSA) identified only three members (JunD, c-Fos and Fra-2) of the AP-1 family that bind to the *Mustang* promoter during both myoblast proliferation and differentiation. Collectively, the data presented herein further support a role for *Mustang* in skeletal muscle differentiation, as was suggested previously for osteoblasts and chondrocytes [2], and suggest that the *Mustang* promoter can be used in conjunction with reporter genes to analyze specific cell lineages (chondrogenic, osteogenic, myogenic, tendogenic) of the musculoskeletal system.

Materials and methods

Materials

[γ-32P]-ATP (3000 Ci/mmol) was purchased from Amersham Pharmacia Biotech. Synthetic oligonucleotides were synthesized by Invitrogen Inc. Anti-c-Jun, anti-JunB, anti-JunD, anti-c-Fos, anti-FosB and anti-Fra-2 polyclonal antibodies were purchased from Active Motif. Anti-cyclophilin antibody was purchased from Upstate. Goat-anti-rabbit IgG HRP monoclonal antibody was purchased from Chemicon.

Cell Culture

C2C12, NIH3T3, COS-1 and HeLa cells were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. RCJ3.1C5 [6] cells were maintained in DMEM supplemented with 15% FBS and 1% penicillin/streptomycin. MC3T3 cells were maintained in minimum essential medium alpha medium (α-MEM) supplemented with 10% FBS and 1% penicillin/streptomycin. All cell lines were cultured at 37°C and 5% CO_2. To induce myogenic differentiation of C2C12 cells, we replaced the medium with DMEM supplemented with 2% horse serum once the cells were confluent.

Quantitative real-time RT-PCR (qRT-PCR)

Proliferating (Day −1) and differentiating (Days 2, 4, 6 and 8) C2C12 cells were collected and total RNA was isolated with RNeasy Mini Kit (QIAGEN) and treated with DNase I (QIAGEN) to remove any traces of DNA. The concentration of each RNA sample was determined by RiboGreen RNA Quantitation Kit (Molecular Probes) following manufacturer's protocol. qRT-PCR was carried out with QuantiTect SYBR Green RT-PCR Kit (QIAGEN) on LightCycler system (Roche) as previously described [7]. Temporal expression levels of *Mustang*, myogenin and MyoD were determined by qRT-PCR and normalized to that of GAPDH. Primer sequences used in these analyses are listed in Table 1. Each experiment was performed in triplicate in order to calculate the standard deviation.

Table 1

Target gene	Accession number	Primer sequence	Amplicon size (bp)	T_m (°C)
A. Primers used for qRT-PCR				
GAPDH	NM_017008	*Forward*: AATGGGGTGATGCTGGTG *Reverse*: GGAAGGGGCGGAGATG	119	60
Mustang	NM_181390	*Forward*: TGCCCAATGTCCCCAAC *Reverse*: TTCCCTGTCCCACCTCA	115	60
Myogenin	NM_031189	*Forward*: GGAAGTCTGTGTCGGTGGAC *Reverse*: CGCTGCGCAGGATCTCCAC	150	60
MyoD	NM_010866	*Forward*: GCCTGAGCAAAGTGAATGAG *Reverse*: GGTCCAGGTGCGTAGAAGG	184	60
B. Primers used for RT-PCR				
Actin (M, R, H, P[a])	M: NM_007393R: NM_031144H: NM_001101P: NM_001009945	*Forward*: AGACCTTCAACACCCCAG *Reverse*: AGGTCCAGACGCAGGATG	166	60
Mustang (H, P)	H: NM_205853P: NW_104868	*Forward*: AAGAAGAAGCGCCCCCCT *Reverse*: CTTTGGGCTTCTCAAAGAC	190	60
Mustang (M)	M: NM_181390	*Forward*: AAGAAGAAGCGGCCCCCT *Reverse*: CTTTGGGCTTCTCAAAGAC	190	60
Mustang (R)	R: NM_181368	*Forward*: AAGAAGAAGCGCCCCCCT *Reverse*: GTCTTCGAGAAGCCCAAAG	190	60

[a] M, *Mus musculus*; R, *Rattus norvegicus*; H, *Homo sapiens*; P, *Pan troglodytes*.

Cloning and construction of the Mustang promoter constructs

To clone the *Mustang* promoter, primers (F(−1447) and R(+65); Table 2A) were designed to amplify the 1512-bp 5′-*Mustang*-flanking mouse genomic region using genomic DNA isolated from mouse tail. The PCR amplicon was cloned into pGL3-Basic luciferase vector following manufacturer's instructions (Promega). Following verification by sequencing, the 1512-bp sequence was analyzed using Alibaba 2.1 (http://www.gene-regulation.de) in order to identify putative transcription factor binding sites. Based on the regulatory sequences identified by this bioinformatic analysis, we created deleted fragments of the promoter and cloned them into the pGL3-Basic vector. The primers used to generate the various deletion constructs are listed in Table 2A. Finally, all genomic clones were verified by sequencing.

Site-directed mutagenesis

Specific mutations and deletion of the AP-1 site between −1151 and −1161 was accomplished by site-directed mutagenesis using the QuikChange II Site-Directed Mutagenesis Kit following the manufacturer's protocol (Stratagene). Specifically, to generate the desired mutation and deletion, we designed a pair of complementary primers for each construct in a way that the sequence of interest is centered and each flanking region contains 13–15 extra oligonucleotides that are identical to the template sequence. Both constructs were then amplified with *PfuUltra* High-Fidelity DNA polymerase using Construct A (Table 2A) as template. Primers used to create these two constructs are listed in Table 2B. Lastly, both mutated and deleted clones were verified by sequencing.

Luciferase activity assay

Cells used in this study were transfected with each of the *Mustang* promoter construct, pGL3-Basic vector (promoterless, negative control) and pGL3-Promoter (SV40 promoter, positive control) (Promega). In a given assay, 8000 cells from each cell line were plated in designated wells of a 96-well plate in triplicate. Twenty-four hours later, the cells were transiently transfected using Fugene 6 Reagent (Roche) according to manufacturer's protocol. Forty-eight hours post-transfection, the cells were lysed within the wells and processed using the Steady-Glo Luciferase assay system according to manufacturer's protocol (Promega). Luminescence from each well was measured on a Tropix TR717 Microplate Luminometer (PE Applied Biosystems). For Fig. 2, normalized data (relative luciferase activity) are presented as percentage based on the value of the *Mustang* promoter construct as compared to the value obtained with positive control (pGL3-Promoter). For Fig. 4, normalized data (relative luciferase activity) are presented as fold increase based on the value of each construct as compared to the value obtained with the negative control (pGL3-Basic vector).

Electrophoretic mobility shift assays (EMSA)

Nuclear extracts from proliferating (Day −1) and differentiating (Day 6) C2C12 cells were isolated using the nuclear extracts kit (Active Motif) following the manufacturer's protocol. EMSA reactions were prepared by adding the following components: (a) ~5–10 μg nuclear extracts from either proliferating (Day −1) or differentiating (Day 6) C2C12 cells; (b) binding buffer (10% glycerol, 1 mM $MgCl_2$, 0.5 mM dithiothreitol, 50 mM KCl, 10 mM Tris–HCl, pH 7.4); (c) 5 ng/μl poly(dIdC); (d) 50× wild-type unlabeled AP-1 sequence (*Forward*: TTGTCC*TAGTCAGCCA*GCTGTG, underlined represents wild-type AP-1 site; *Reverse*: CACAGCTGGCTGACTAGGACAA); (e) 50× mutated unlabeled AP-1 sequence (*Forward*: TTGTCCTGGTTCGACCGCTGTG; *Reverse*: CACAGCGGTCGAACCAGGACAA); and (f) individual antibodies specific to AP-1 family members. All reactions were incubated on ice for 1 h. Lastly, 1× 32P-ATP-labeled wild-type probes were added and the reactions were further incubated on ice for 30 min. The protein–DNA complexes were resolved on 4% non-denaturing polyacrylamide gels in 0.25× TBE buffer at 4°C and visualized by autoradiography. Antibodies used in this assay were included in the Nushift AP-1 Family Kit (Active Motif).

Western blotting

Equal amount of protein extracts (same as those used in EMSA analyses) were resolved on 12% SDS–PAGE gels and transferred to nitrocellulose membranes. After blocking with 5% non-fat milk in 1× TBST buffer for 1 h at room temperature, membranes were incubated with each specific antibody (diluted 1:500 in 1× TBST) overnight at 4°C. The blots were then washed in 1× TBST buffer and probed with goat-anti-rabbit HRP-conjugated IgG secondary antibody (diluted 1:500 in 1× TBST) for 1 h at room temperature. Antibody binding was visualized with ECL chemiluminescence reagent (Pierce) and then exposed on X-ray film (Kodak). After stripping, the blots were re-probed with anti-cyclophilin antibody (1:1000, Upstate) following the same procedure.

Table 2

A. Primers used for the *Mustang* promoter deletions construction

Target construct	Primer name	Primer sequence	Target size (bp)	T_m (°C)
A	F(−1447)	*Forward*: GTCGCTCGAGATGGTGTACTTCCATT	1512	60
B	F(−1250)	*Forward*: ATTACTCGAGCCTAGCGTGGTCTA	1315	60
C	F(−1187)	*Forward*: ATTACTCGAGCTGGGCATCCCTTATC	1252	60
D	F(−1132)	*Forward*: ATTACTCGAGGCATGGCCTGGCCT	1197	60
E	F(−366)	*Forward*: TAGTCTCGAGCATCCACCCTTGTTCA	431	60
F	F(−200)	*Forward*: ATTACTCGAGTAAGCAGCTGTCCCCA	265	60
G	F(−121)	*Forward*: GGTCCTCGAGAATAAACTCCAGCTAG	186	60
H	F(−60)	*Forward*: TAATACTCGAGTGACTACCCAGGACG	125	60
I	F(+1)	*Forward*: ACTACTCGAGATCCTTTCCTGTGGCT	65	60
J	R(−117)	*Reverse*: CAGTGAATTCTTGGCGATGATGGGCA	388	60
	F(−73)	*Forward*: TCAGGAATTCCAAAGGAGGGGAGT		
K	R(−286)	*Reverse*: ACCTGAATTCTGCAAGAACCCATCCC	1381	60
	F(−154)	*Forward*: TACAGAATTCCTCTCACCAGGGCA		
A–K	R(+65)	*Reverse*: TTAGCCATGGTGGATGCCAAGCAA	–	60

B. Primers used for the site-directed mutagenesis

Construct	Application	Primer sequence
L	AP-1 mutation	*Forward*: CCTTATCCTTGTCCGCACTAGCCAGCTGTGGG
		Reverse: CCCACAGCTGGCTAGTGCGGACAAGGATAAGG
M	AP-1 deletion	*Forward*: CATCCCTTATCCTTGTCCTGGGTACTCCTCACAAGG
		Reverse: CCTTGTGAGGAGTACCCAGGACAAGGATAAGGGATG

Results

Identification, cloning and characterization of the Mustang promoter

Using the Basic Local Alignment Search Tool (BLAST) in conjunction with the *Mustang* coding region (accession no.: NM_181390), we were able to identify *Mustang*'s genomic organization. Alignment of the *Mustang* coding sequence with the mouse genome database revealed that *Mustang* resides on chromosome 14. Specifically, *Mustang*'s ORF is comprised of three exons and two introns (Fig. 1A). A 1512-bp fragment corresponding to the 5′-flanking region of the gene was chosen to represent the putative *Mustang* promoter based on bioinformatic analysis that revealed that this sequence contained a translation start codon (ATG), a TATA box, and multiple transcription factor binding sites (Alibaba 2.1; http://www.gene-regulation.de) (Fig. 1). Notable among them are sequences for activator proteins AP-1 (4 sites) and AP-2 (2 sites) (Fig. 1). It is not surprising that we found these sites because it is well established that AP-1 and AP-2 factors are known to be key regulators of several genes specific to the musculoskeletal system [8–12]. In addition, other transcription factor binding sites were also predicted by the program (data not shown).

Transcriptional activity of the Mustang promoter in various cell lines

To study the transcriptional activity of the *Mustang* promoter, we designed primers (Tables 2A and B), performed PCR and isolated the 1512-bp putative promoter sequence from

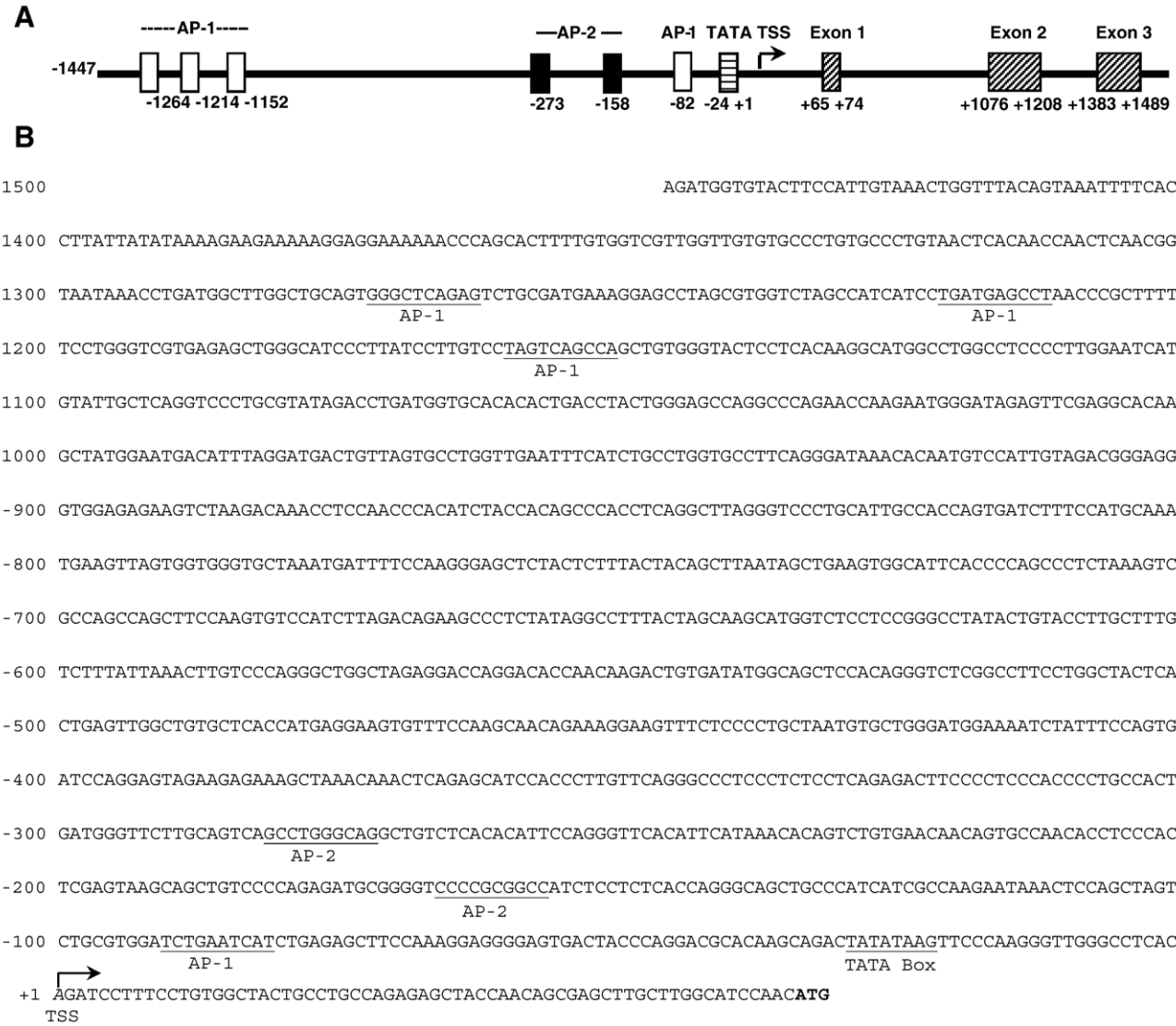

Fig. 1. Genomic organization of murine *Mustang*. (A) Genomic organization of the *Mustang* gene and schematic distribution of the transcription factor binding sites of interest. Numbers below each box represent the start position of the site. (B) 5′ upstream sequence of the *Mustang* gene and the putative transcriptional factor binding sites. The sequence spans from −1447 to +65 (1512 bp in total excluding the translation start codon). Indicated are the AP-1 and AP-2 sites, a TATA box and a transcriptional start site (TSS). The specific sequences corresponding to each transcription factor binding site are all underlined. The TSS is italicized and the translation start codon is in boldface.

mouse chromosomal DNA. This fragment was subcloned upstream of the firefly luciferase gene in the pGL3-Basic vector (indicated as pGL3-Mus1512 construct) (Fig. 2B). pGL3-Mus1512 construct along with the empty promoterless vector (pGL3-Basic, negative control) and an SV40-driven luciferase gene construct (pGL3-Promoter, positive control) were used to transfect the following 6 cell lines: C2C12, RCJ3.1C5, MC3T3, NIH3T3, COS-1 and HeLa. We chose these cell lines because they represent cells where *Mustang* is known to be expressed (C2C12, myogenic; RCJ3.1C5, chondrogenic; MC3T3, osteogenic; and NIH3T3, embryonic fibroblasts) or not (COS-1 and HeLa), as measured by RT-PCR (Fig. 2A). The relative luciferase activity observed from each cell line was normalized to the positive control and presented as a percentage (Fig. 2B). All three lines representing cells of the musculoskeletal system (C2C12, RCJ3.1C5 and MC3T3) and NIH3T3 showed high levels of promoter activity (145%, 83%, 28% and 35%, respectively, as compared to positive control, pGL3-Promoter) consistent with *Mustang* expression. In contrast, the levels in the other two non-*Mustang* expressing cell lines (COS-1 and HeLa) were equivalent to those of the negative control (pGL3-Basic) (Fig. 2). Thus, the promoter activity correlated identically with the expression of *Mustang*; present in C2C12, RCJ3.1C5, MC3T3 and NIH3T3 but absent in COS-1 and HeLa (Fig. 2).

Further, we observed that the highest levels of luciferase activity in C2C12 cells corresponded with the higher levels of *Mustang* expression (Fig. 2), consistent also with our previous results showing that *Mustang* is highly expressed in adult skeletal muscle [2]. Specifically, luciferase activity driven by the 1512-bp *Mustang* promoter was detected at higher (145%) or almost equal levels (83%) in the myogenic C2C12 and chondrogenic RCJ3.1C5 cells, respectively, as compared to the positive control (pGL3-Promoter) driven by the strong SV40 viral promoter (Fig. 2B).

Temporal Mustang expression during C2C12 myogenic differentiation

Based on the aforementioned results, we decided to further characterize the *Mustang* promoter in C2C12 cells. Thus, we initially analyzed *Mustang* expression in order to investigate whether there is a direct correlation between its temporal expression and myogenic differentiation. To monitor C2C12 myogenic differentiation, we chose to measure *Mustang* expression levels and those of two myogenesis-specific genes, MyoD and myogenin, as well as cell morphology. Both MyoD and myogenin are well-established transcription factors and serve as markers for early and late myogenic differentiation, respectively [13,14].

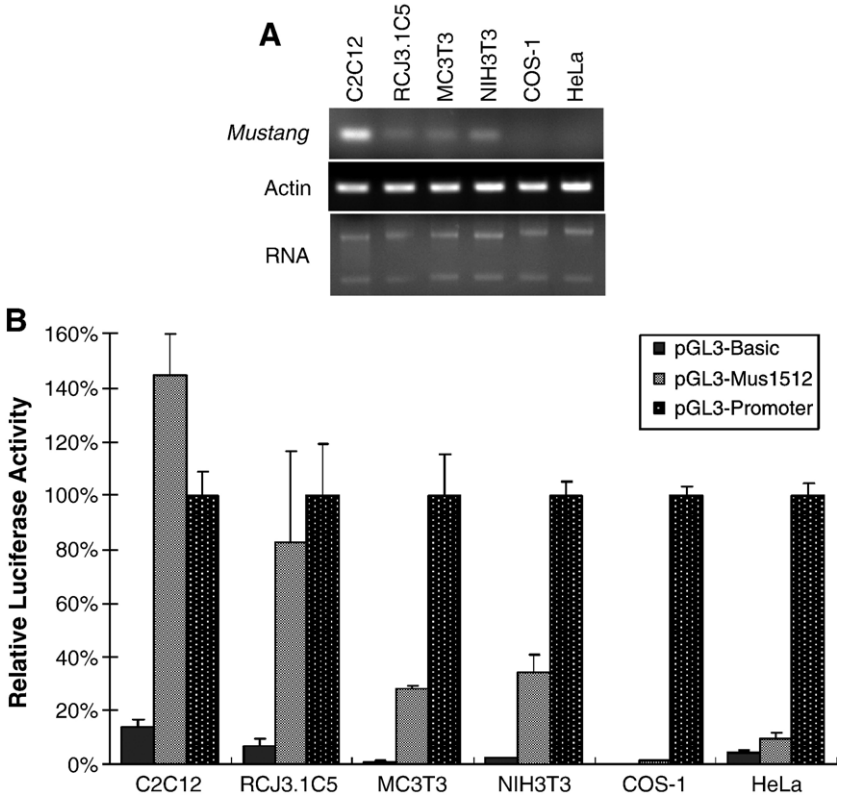

Fig. 2. *Mustang* expression in various cell lines and comparison of *Mustang* promoter activity. Cell lines included in these experiments are C2C12 (myogenic), RCJ3.1C5 (chondrogenic), MC3T3 (osteogenic), NIH3T3 (embryonic fibroblasts), COS-1 (kidney fibroblasts) and HeLa (carcinoma). (A) *Mustang* expression via RT-PCR. Equal amount of RNA was used for each reaction. The cycling was controlled so that all reactions were terminated at the log phase using the same cycle numbers. (B) Luciferase activity assay of the 1512-bp *Mustang* promoter luciferase gene construct in each cell line. pGL3-Basic is the empty vector only (negative control), whereas pGL3-Promoter contains the viral SV40 promoter (positive control). All values were normalized to that of the positive control of each individual cell line and presented as percentage. Error bars indicate standard deviation of triplicate values.

Specifically, using qRT-PCR, we analyzed *Mustang*, MyoD and myogenin expression using RNA isolated from Day −1 (representing cell proliferation) and Days 2, 4, 6 and 8 (representing various differentiation stages) (Fig. 3A). Results indicate that *Mustang* expression was temporally regulated with the highest levels detected at the later stages of differentiation (Days 6 and 8) where the cells have formed distinct multinucleated myotubes (Figs. 3B and C). In contrast, MyoD and myogenic expression, although also temporally regulated, peaked at earlier time points, at Days 4 and 6, respectively, and then declined (Fig. 3A).

Deletion and mutational analysis of Mustang promoter

To determine the minimal sequence required for *Mustang* promoter activity and to define the *cis*-elements responsible for transcriptional activation, we created serially deleted promoter fragments, as well as specific deletion constructs based on the distribution of the AP-1 and AP-2 sites within the 1512-bp promoter region (Fig. 4A). Again, we emphasize that we focused on AP-1 and AP-2 because of their involvement in regulating several other musculoskeletal-specific genes [8–12]. Eight serially deleted fragments (Constructs B–I) and two specific deleted fragments (Constructs J and K) were cloned into pGL3-Basic vector

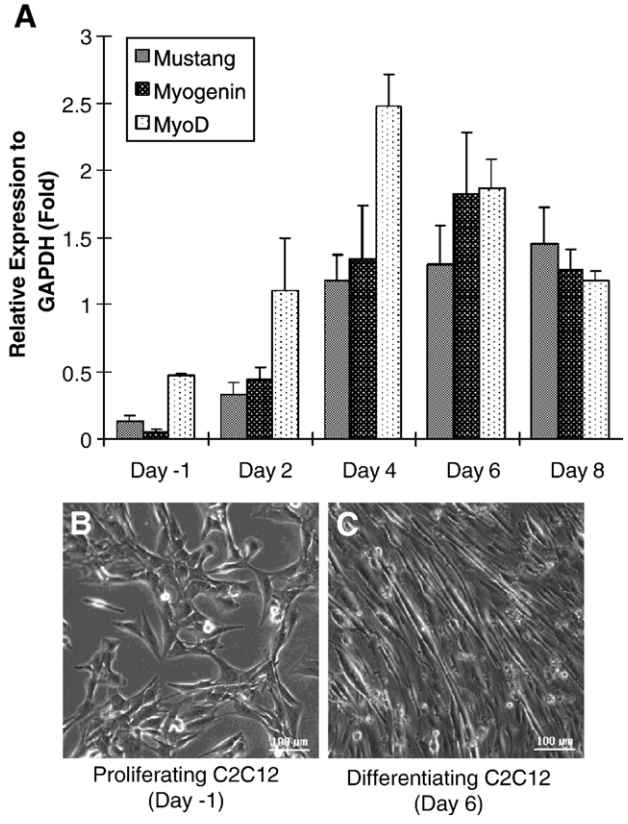

Fig. 3. Temporal expression of *Mustang* during C2C12 differentiation. (A) qRT-PCR of *Mustang*, myogenin and MyoD using RNA from proliferating (Day −1) and differentiating (Days 2, 4, 6 and 8) cells. (B and C) Actively proliferating (Day −1) and differentiating (Day 6) C2C12 cells, respectively. (C) Myotubes, indicative of late differentiation, are also clearly seen.

(Construct A was the same as pGL3-Mus1512) (Fig. 4). Transient transfection of all constructs followed by luciferase activity assay in C2C12 cells was performed and luciferase activity was measured 48 h later coinciding with the early stage of myogenic differentiation and *Mustang*'s upregulation (see Fig. 3A).

Results revealed maximum luciferase activity with Constructs A–C, indicating the contribution of the three AP-1 sites to the transcriptional activation of luciferase (Fig. 4A). In addition, we observe a 40% increase in luciferase activity when the first and second AP-1 sites are deleted (compare Constructs A–C), suggesting the presence of an inhibitory site within these deleted sequences (Fig. 4A). More importantly, when the third AP-1 site at − 1151 is deleted (Construct D), luciferase activity decreased by 64.5% (compare Constructs C and D; Fig. 4A). The other AP-1 and AP-2 sites showed much lower effect on luciferase activity (Constructs E–G). Further, if all binding sites are deleted (Construct H) as well as the TATA box (Construct I), then luciferase activity decreased to the level obtained with the negative control (pGL3-Basic) (Fig. 4A). In addition, we generated two constructs that represent deletions of all AP-1 (Construct J) and AP-2 sites (Construct K). It is clear from these two constructs that deleting both AP-2 sites only decreases transcriptional activity of the promoter by 12%. In contrast, deleting all four AP-1 sites (Construct J) reduced the luciferase activity by 73.5%, thus indicating that the AP-1 sites are the predominant transcription factor binding sites required for maximal promoter activity (Fig. 4A).

Because we observed a dramatic decrease in luciferase activity when the third AP-1 site at − 1151 along with its adjacent sequences were deleted (Construct D; Fig. 4A), we decided to further analyze this sequence using site-directed mutagenesis (Construct L) and specific deletion (Construct M). Additionally, we wanted to rule out the possibility that there may be other sequences flanking this AP-1 site that could also contribute to the induction of luciferase. Hence, the activity of each construct was compared with the wild-type AP-1 (Construct A), as well as those of the negative and positive controls. Results show distinct decreases in luciferase activity with both mutation (40%) and deletion (32%) of the AP-1 site (Fig. 4B). Combined with the previous serial deletion analyses, these results clearly indicate that AP-1 transcription factors are likely required for the activation of *Mustang* in differentiating C2C12 myoblasts.

Confirmation of AP-1 binding by EMSA

Because the AP-1 site at − 1151 was identified as the critical regulatory element for *Mustang* promoter activity, we decided to identify which members of the AP-1 family bind to this site. This was accomplished by using oligonucleotides containing the wild-type AP-1 binding sequence and a mutated sequence in conjunction with nuclear proteins isolated from both proliferating (Day −1) and differentiating (Day 6) C2C12 cells (it has been previously reported that different AP-1 family members are involved in the transition

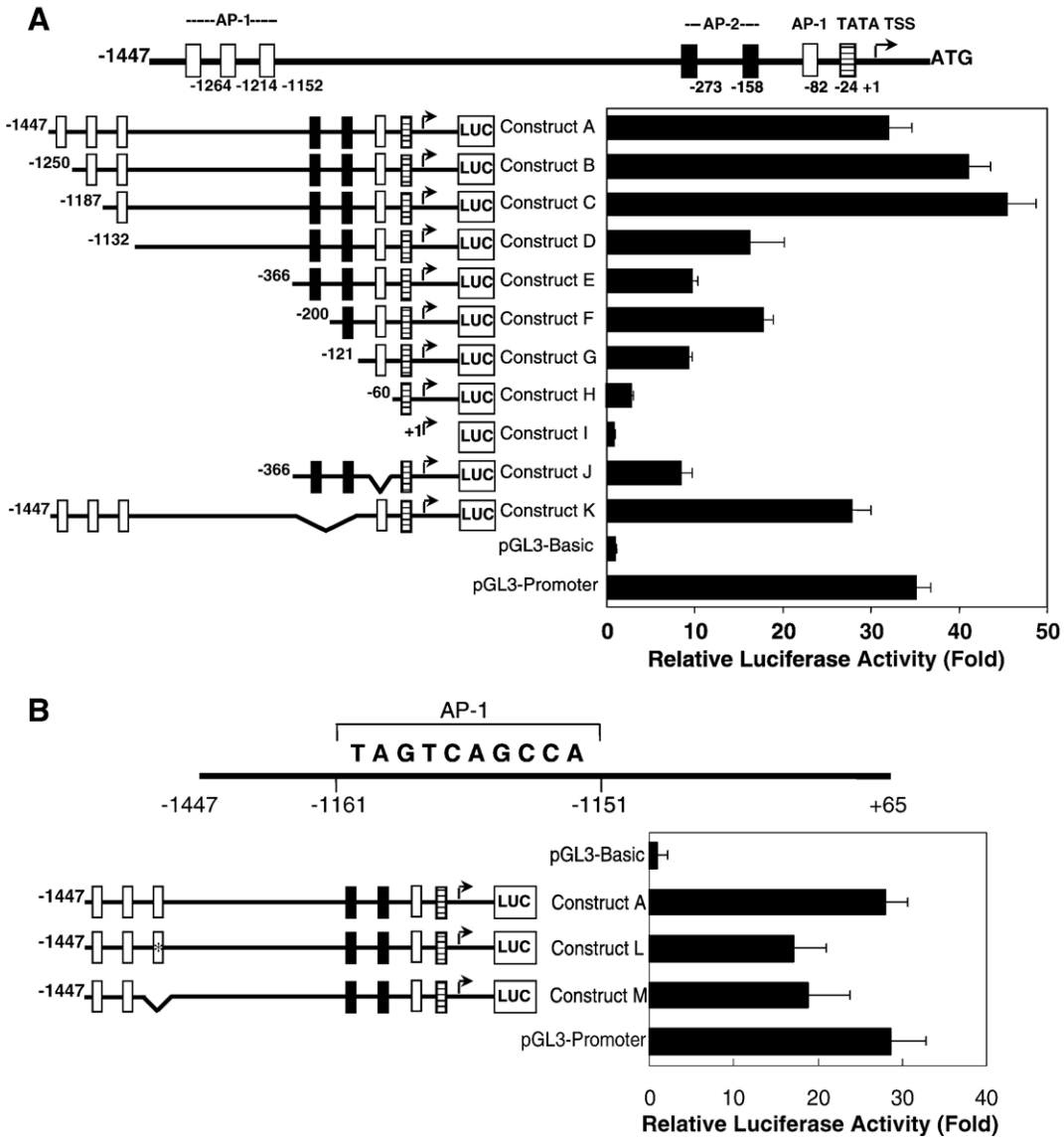

Fig. 4. Deletion/mutation analyses of the *Mustang* promoter. (A) Activity of *Mustang* promoter serially deleted constructs (A–I) and other specifically deleted constructs (J: AP-1 deleted; K: AP-2 deleted). C2C12 cells were transfected with equal amount of plasmids containing Constructs A–K. Luciferase activity was measured 2 days after transfection. Promoter activity is reported as fold of each construct over that of the negative control (pGL3-Basic). Error bars indicate standard deviation of triplicate values. (B) Mutation/deletion analysis of the AP-1 site between −1161 and −1151. Mutated AP-1 was created by randomly switching the first 5 oligonucleotides to their opposite type (i.e., from purine to pyrimidine or from pyrimidine to purine). Both mutation and deletion was generated by site-directed mutagenesis. Luciferase activity was measured and represented as described for panel A.

from proliferation to myogenic differentiation [9]), as well as monoclonal antibodies specific to the following AP-1 family members: c-Jun, JunB, JunD, c-Fos, FosB and Fra-2. Results from these EMSA experiments revealed that AP-1 members bind effectively to the wild-type radioactive probes (AP-1 binding site) as indicated in Lane 2 (Figs. 5A and B, indicated by band labeled as "shift"). Further, this binding was blocked by the addition of excess (50×) wild-type unlabeled probes (Lane 3), whereas it was not affected by the addition of the same amount of the mutated unlabeled probes (Lane 4). A strong non-specific binding band was also observed, as indicated (Figs. 5A and B). Next, in order to identify which AP-1 family members participated in binding to this site, we applied antibodies against c-Jun, JunB, JunD, c-Fos, FosB and Fra-2. Supershifts (Lanes 7, 8 and 10) indicated that JunD, c-Fos and Fra-2 are the only AP-1 members that bind to this AP-1 site (Figs. 5A and B). Lastly, no differences were detected between proliferating and differentiating C2C12 cells (Figs. 5A and B).

Western blot analysis of the AP-1 family members

To further verify the link between specific members of the AP-1 family and promoter activity, we performed Western blot analysis to determine the protein expression levels of c-Jun, JunB, JunD, c-Fos, FosB and Fra-2 using Day − 1 and Day 6 nuclear protein extracts. Results from these analyses showed that c-Jun, JunB, JunD, c-Fos and Fra-2 are expressed during

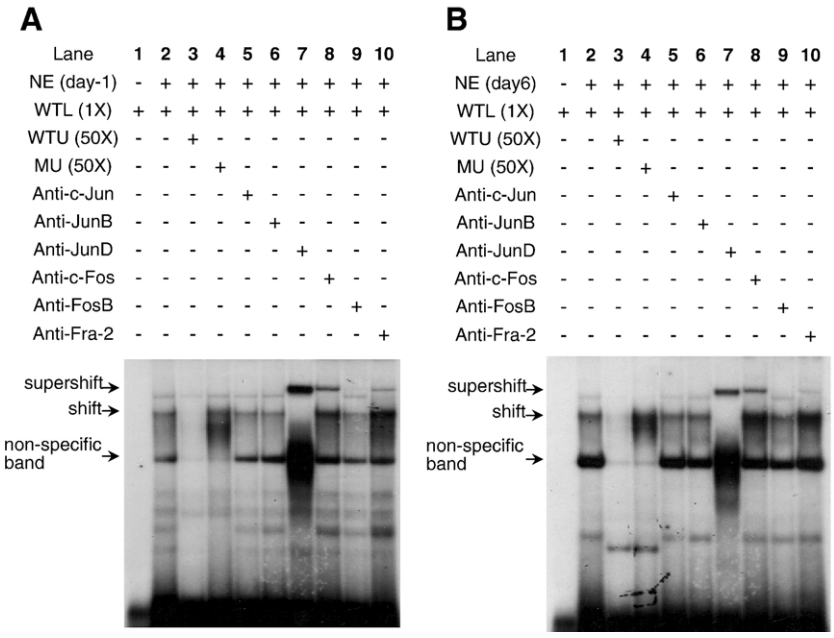

Fig. 5. Verification of AP-1 binding and identification of specific family members. EMSA was performed using C2C12 nuclear extracts (NE) from (A) proliferating cells (Day −1) and (B) differentiating cells (Day 6). Same amount of wild-type-labeled (WTL) probes was added to each lane. For competition analysis, 50× wild-type unlabeled (WTU) probes or mutated unlabeled (MU) probes were also added. Bands that are absent with the addition of excess WTU probe but remain visible with the addition of excess MU probes are shifts caused by AP-1 binding. Supershifts are bands that result from the binding of each specific antibody targeting a particular AP-1 family member. Non-specific bands were also detected.

C2C12 proliferation (Day −1), but in differentiating cells only expression of JunB, JunD, c-Fos and Fra-2 is detected. In contrast, c-Jun expression is completely abolished by Day 6 (Fig. 6). No expression was detected for FosB at either time point (Fig. 6). These results are also consistent with those obtained from the EMSA analyses in that they show that the three AP-1 members (JunD, c-Fos and Fra-2) that induced supershifts are also expressed in both proliferating and differentiating cells (Figs. 5A, B and Fig. 6). Additionally, two closely positioned bands were observed when anti-JunD antibody was applied against Day −1 nuclear proteins whereas at Day 6 only one band was detected (Fig. 6).

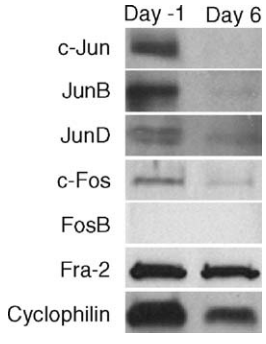

Fig. 6. Western blot analyses of the AP-1 family members. Western analyses as described in Materials and methods were used with equal amounts of nuclear extracts from proliferating (Day −1) and differentiating (Day 6) C2C12 cells. For detection of each AP-1 family member, the identical antibodies utilized in EMSA analyses (see Fig. 5) were also used for these Western blots. The lower levels of cyclophilin detected are indicative of a general decrease in cellular proliferation, as expected.

Discussion

Mustang is a novel gene that possesses a unique spatiotemporal expression pattern but whose function remains unknown. Previous analyses showed that its expression was restricted to fracture callus, skeletal muscle, as well as tendon [1]. A temporal analysis of its expression in regenerating fracture calluses indicated that *Mustang* was upregulated 55-fold at PF Day 5 when compared to intact unfractured bone. As the callus proceeded to heal, *Mustang* expression gradually decreased accordingly and suggested the importance of *Mustang* during the earlier phases of the repair process. More specifically, in situ hybridization revealed that *Mustang* is only strongly expressed in osteoprogenitor cells, young osteoblasts and proliferating chondrocytes of the healing callus. Consistent with the notion that fracture repair recapitulates skeletal development [15], *Mustang* expression was also detected during embryogenesis, especially in mesenchymal condensations of limbs, vertebral perichondrium and mesenchymal cells of the intervertebral discs [2]. Provided with these data, we strongly believe that *Mustang* represents a novel musculoskeletal marker that may play a crucial role in both bone development and regeneration. Thus, identification and characterization of the *Mustang* promoter will enable us to study the gene's transcriptional regulation.

Based on our bioinformatic analyses, the genomic organization of *Mustang* reveals that it contains 3 exons separated by 2 introns. A 1512-bp region upstream of the 5′ end of the first exon was chosen to represent the *Mustang* promoter due to the presence of classical features, including a transcriptional start site, a TATA box and multiple transcription factor binding sites.

Some of these sites included AP-1 sites, which were previously reported to represent regulatory binding sites in many other musculoskeletal-specific genes (i.e., PTHrP, MyoD, beta-catenin, osteocalcin and collagenase-3) [16–20]. AP-2 sites were also found and included in our investigation for their reported activity during embryonic development [12]. Other transcription factor binding sites were also identified through our bioinformatic analyses; however, because our goal was to decipher the link between *Mustang* and the musculoskeletal system, we chose to focus on the AP-1 and AP-2 sites for the aforementioned reasons.

Luciferase activity assays of the full-length *Mustang* promoter showed that *Mustang* expression is highly cell type specific. That is, luciferase activation correlates perfectly with known expression of *Mustang* in myogenic (C2C12), chondrogenic (RCJ3.1C5), osteogenic (MC3T3) and fibroblastic (NIH3T3) cells. In contrast, in other cell types where *Mustang* is not expressed (i.e., COS-1 and HeLa), luciferase activation was not detectable. In addition, these luciferase activity assays also indicated that the *Mustang* promoter represents a very strong promoter, as its ability to activate luciferase in chondrogenic RCJ3.1C5 and myogenic C2C12 cells equal to or exceeded that of the strong viral SV40 promoter, respectively. The strength of the *Mustang* promoter makes it unique among other isolated mammalian promoters, especially in the context of *Mustang*'s highly restricted expression to the musculoskeletal system.

The high level of *Mustang* expression during C2C12 myogenic differentiation provides an ideal system for studying its regulation. Results from our experiments showed that *Mustang* underwent prominent and sustained upregulation during C2C12 differentiation. However, unlike MyoD and myogenin, two well-known myogenic molecular markers whose expression peaked and then decreased steadily, *Mustang* expression remains high (at least up to 8 days after cell confluency). Based on these data, as well as the fact that *Mustang* is highly expressed in terminally differentiated adult skeletal muscle [2], we speculate that *Mustang* is a significant player in myogenic differentiation and may serve as a late differentiation marker for skeletal muscle cells.

In order to unveil the transcriptional regulation of *Mustang* expression, we also tried to decipher the information embedded within its promoter. Based on our promoter deletion/mutation data, as well as EMSAs, we determined that the dominant transcriptional regulators for *Mustang* are specific members of the AP-1 family. This is not surprisingly because AP-1 has been reported to be a versatile transcription factor that regulates numerous genes involved in a variety of cell types and cellular processes [21]. Interestingly, AP-1 and its family members (Fos family: c-Fos, FosB, Fra-1 and Fra-2; Jun Family: c-Jun, JunB and JunD) are key regulators of specific genes of the musculoskeletal system and loss-of-function studies showed different degrees of deficiency when specific members were knocked out individually [8].

Selective pairing between the Fos and Jun family members was previously suggested as one source of AP-1's binding versatility [9]. Further, the ATF family proteins that are able to dimerize with certain Fos or Jun proteins added more complexity to this composition-dependent specificity [22]. Another common mechanism of regulating AP-1's specificity is post-translational modification such as phosphorylation and dephosphorylation [23,24]. To date, the full regulatory activity of AP-1 and its family members is still not completely elucidated and our data add *Mustang* to the list of AP-1 target genes by showing that *Mustang* is activated by the binding of c-Fos, JunD and Fra-2 to its promoter.

No compositional change of AP-1 family members was observed between proliferating and differentiating C2C12 cells. It is noteworthy to mention that our Western blots revealed a double band when anti-JunD was applied to the nuclear extracts from proliferating cells, whereas only a single band was detected with nuclear extracts from differentiating cells. This may indicate that during C2C12 differentiation, dephosphorylation of JunD occurs and may have significant functional consequences. Consistent with this idea, a previous study showed downregulation of both phosphorylated and dephosphorylated JunD upon C2C12 cell entry in differentiation [9]. Hence, it is likely that the upregulation of *Mustang* expression during the latter phases of C2C12 cell differentiation is connected to JunD dephosphorylation. Further, because Fos family proteins are unable to dimerize with each other unlike the Jun family proteins [9], JunD is then likely to serve as the active and determining component of the AP-1 complex that binds to the *Mustang* promoter (because JunD was the only member of Jun family detected by EMSA). However, further experiments are required before we can determine conclusively the exact composition of the AP-1 transcriptional complex responsible for regulating *Mustang* expression during both cell proliferation and differentiation.

Whereas *Mustang* activation depends on the binding of specific AP-1 family proteins, MyoD is negatively regulated by AP-1 through downregulation of c-Fos and c-Jun [17]. Further, a dual role of c-Jun (stimulates/represses myoblast differentiation) in regulating myogenin expression has also been reported [25]. Hence, although evidence has shown that *Mustang* is tightly linked to myoblast differentiation, its transcriptional regulation could be completely different from that of MyoD and myogenin. Provided that the process of myogenesis is heavily linked to the MyoD-associated signaling pathway, elucidation of *Mustang* function could certainly place it within this pathway.

In summary, we have identified and characterized the *Mustang* promoter. This 1512-bp DNA sequence contains multiple AP-1 binding sites that activate gene transcription very strongly, especially in cells of the musculoskeletal system. Given that *Mustang* expression is restricted to the musculoskeletal system, coupled with its high level of expression during development and regeneration, makes its promoter ideal for future studies. Specifically, using *Mustang* promoter GFP transgenic mice will enable us to characterize the spatiotemporal expression of *Mustang*, as well as to perform lineage mapping analyses more comprehensively during musculoskeletal development and regeneration.

Acknowledgments

We would like to thank Dr. Anil Dhundale for providing the luciferase vector and critically reading the manuscript; Dr. Michael Frohman for the NIH3T3 cells; Emily Huang and Dr. David Komatsu for technical assistance and helpful discussions; and Colleen Michaels for administrative support. This work was kindly supported by a grant from NASA (MH).

References

[1] Hadjiargyrou M, Lombardo F, Zhao S, Ahrens W, Joo J, Ahn H, et al. Transcriptional profiling of bone regeneration. Insight into the molecular complexity of wound repair. J Biol Chem 2002; 277:30177–82.

[2] Lombardo F, Komatsu D, Hadjiargyrou M. Molecular cloning and characterization of *Mustang*, a novel nuclear protein expressed during skeletal development and regeneration. FASEB J 2004; 18:52–61.

[3] Tomczak KK, Marinescu VD, Ramoni MF, Sanoudou D, Montanaro F, Han M, et al. Expression profiling and identification of novel genes involved in myogenic differentiation. FASEB J 2004;18: 403–405.

[4] Balint E, Lapointe D, Drissi H, van der Meijden C, Young DW, van Wijnen AJ, et al. Phenotype discovery by gene expression profiling: mapping of biological processes linked to BMP-2-mediated osteoblast differentiation. J Cell Biochem 2003;89:401–26.

[5] Teboul L, Gaillard D, Staccini L, Inadera H, Amri EZ, Grimaldi PA. Thiazolidinediones and fatty acids convert myogenic cells into adipose-like cells. J Biol Chem 1995;270:28183–7.

[6] Grigoriadis AE, Heersche JN, Aubin JE. Continuously growing bipotential and monopotential myogenic, adipogenic, and chondrogenic subclones isolated from the multipotential RCJ 3.1 clonal cell line. Dev Biol 1990;142:313–8.

[7] Komatsu DE, Hadjiargyrou M. Activation of the transcription factor HIF-1 and its target genes, VEGF, HO-1, iNOS, during fracture repair. Bone 2004;34:680–8.

[8] Wagner EF. Functions of AP1 (Fos/Jun) in bone development. Ann Rheum Dis 2002;61(Suppl 2):40–2.

[9] Andreucci JJ, Grant D, Cox DM, Tomc LK, Prywes R, Goldhamer DJ, et al. Composition and function of AP-1 transcription complexes during muscle cell differentiation. J Biol Chem 2002; 277:16426–32.

[10] Palcy S, Bolivar I, Goltzman D. Role of activator protein 1 transcriptional activity in the regulation of gene expression by transforming growth factor beta1 and bone morphogenetic protein 2 in ROS 17/2.8 osteoblast-like cells. J Bone Miner Res 2000;15:2352–61.

[11] Karreth F, Hoebertz A, Scheuch H, Eferl R, Wagner EF. The AP1 transcription factor Fra2 is required for efficient cartilage development. Development 2004;131:5717–25.

[12] Eckert D, Buhl S, Weber S, Jager R, Schorle H. The AP-2 family of transcription factors. Genome Biol 2005;6:246.

[13] Berkes CA, Tapscott SJ. MyoD and the transcriptional control of myogenesis. Semin Cell Dev Biol 2005;16:585–95.

[14] Knapp JR, Davie JK, Myer A, Meadows E, Olson EN, Klein WH. Loss of myogenin in postnatal life leads to normal skeletal muscle but reduced body size. Development 2006;133:601–10.

[15] Gersch R, Lombardo F, McGovern S, Hadjiargyrou M. Re-activation of Hox gene expression during bone regeneration. J Orthop Res 2005; 23:882–90.

[16] Ionescu AM, Schwarz EM, Vinson C, Puzas JE, Rosier R, Reynolds PR, et al. PTHrP modulates chondrocyte differentiation through AP-1 and CREB signaling. J Biol Chem 2001;276:11639–47.

[17] Pedraza-Alva G, Zingg JM, Jost JP. AP-1 binds to a putative cAMP response element of the MyoD1 promoter and negatively modulates MyoD1 expression in dividing myoblasts. J Biol Chem 1994; 269:6978–85.

[18] Li Q, Dashwood WM, Zhong X, Al-Fageeh M, Dashwood RH. Cloning of the rat beta-catenin gene (Ctnnb1) promoter and its functional analysis compared with the Catnb and CTNNB1 promoters. Genomics 2004; 83:231–42.

[19] Goldberg D, Polly P, Eisman JA, Morrison NA. Identification of an osteocalcin gene promoter sequence that binds AP1. J Cell Biochem 1996;60:447–57.

[20] Winchester SK, Selvamurugan N, D'Alonzo RC, Partridge NC. Developmental regulation of collagenase-3 mRNA in normal, differentiating osteoblasts through the activator protein-1 and the runt domain binding sites. J Biol Chem 2000;275:23310–8.

[21] Hess J, Angel P, Schorpp-Kistner M. AP-1 subunits: quarrel and harmony among siblings. J Cell Sci 2004;117:5965–73.

[22] Bakiri L, Matsuo K, Wisniewska M, Wagner EF, Yaniv M. Promoter specificity and biological activity of tethered AP-1 dimers. Mol Cell Biol 2002;22:4952–64.

[23] Au YP, Dobrowolska G, Morris DR, Clowes AW. Heparin decreases activator protein-1 binding to DNA in part by posttranslational modification of Jun B. Circ Res 2002;75:15–22.

[24] Le NH, van der Bent P, Huls G, van de Wetering M, Loghman-Adham M, Ong AC, et al. Aberrant polycystin-1 expression results in modification of activator protein-1 activity, whereas Wnt signaling remains unaffected. J Biol Chem 2004;279:27472–81.

[25] Daury L, Busson M, Tourkine N, Casas F, Cassar-Malek I, Wrutniak-Cabello C, et al. Opposing functions of ATF2 and Fos-like transcription factors in c-Jun-mediated myogenin expression and terminal differentiation of avian myoblasts. Oncogene 2001;20: 7998–8008.