



MyoD and E-protein heterodimers switch rhabdomyosarcoma cells from an arrested myoblast phase to a differentiated state

Zhihong Yang, Kyle L. MacQuarrie, Erwin Analau, et al.

Genes Dev. 2009 23: 694-707

Access the most recent version at doi:[10.1101/gad.1765109](https://doi.org/10.1101/gad.1765109)

**Supplemental
Material**

<http://genesdev.cshlp.org/content/suppl/2009/03/19/23.6.694.DC2.html>

References

This article cites 59 articles, 32 of which can be accessed free at:
<http://genesdev.cshlp.org/content/23/6/694.full.html#ref-list-1>

**Email alerting
service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

To subscribe to *Genes & Development* go to:
<http://genesdev.cshlp.org/subscriptions>

MyoD and E-protein heterodimers switch rhabdomyosarcoma cells from an arrested myoblast phase to a differentiated state

Zhihong Yang,¹ Kyle L. MacQuarrie,^{1,2} Erwin Analau,¹ Ashlee E. Tyler,¹ F. Jeffery Dilworth,³ Yi Cao,¹ Scott J. Diede,¹ and Stephen J. Tapscott^{1,4,5}

¹Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA; ²Medical Scientist Training Program, University of Washington, Seattle, Washington 98105, USA; ³Sprott Center for Stem Cell Research, Ottawa Health Research Institute, Ottawa, Ontario K1Y 4E9, Canada; ⁴Department of Neurology, University of Washington, Seattle, Washington 98105, USA

Rhabdomyosarcomas are characterized by expression of myogenic specification genes, such as *MyoD* and/or *Myf5*, and some muscle structural genes in a population of cells that continues to replicate. Because *MyoD* is sufficient to induce terminal differentiation in a variety of cell types, we have sought to determine the molecular mechanisms that prevent *MyoD* activity in human embryonal rhabdomyosarcoma cells. In this study, we show that a combination of inhibitory *Musculin*:E-protein complexes and a novel splice form of *E2A* compete with *MyoD* for the generation of active full-length E-protein:MyoD heterodimers. A forced heterodimer between *MyoD* and the full-length *E12* robustly restores differentiation in rhabdomyosarcoma cells and broadly suppresses multiple inhibitory pathways. Our studies indicate that rhabdomyosarcomas represent an arrested progress through a normal transitional state that is regulated by the relative abundance of heterodimers between *MyoD* and the full-length *E2A* proteins. The demonstration that multiple inhibitory mechanisms can be suppressed and myogenic differentiation can be induced in the RD rhabdomyosarcomas by increasing the abundance of *MyoD*:E-protein heterodimers suggests a central integrating function that can be targeted to force differentiation in muscle cancer cells.

[*Keywords:* *E2A*; *Musculin*; *MyoD*; myogenesis; rhabdomyosarcoma]

Supplemental material is available at <http://www.genesdev.org>.

Received November 20, 2008; revised version accepted February 9, 2009.

The myogenic basic helix–loop–helix (bHLH) transcription factors—*MyoD*, *Myf5*, *Myogenin*, and *MRF4*—together regulate skeletal muscle specification and differentiation (Bergstrom et al. 2002; Berkes and Tapscott 2005). *MyoD* and *Myf5* are expressed in the replicating myoblasts, and the molecular mechanisms of maintaining proliferative growth in the presence of myogenic bHLH proteins is only partly understood and is likely to be complex. For example, *MyoD* and *Myf5* function as transcriptional activators as heterodimers with E-proteins, a subfamily of bHLH proteins consisting of *E2-2/TCF4*, *HEB/TCF12*, and *E2A/TCF3*. In contrast, the HLH protein *Id* and the bHLH protein *Musculin* (*MSC*)/*MyoR* can prevent *MyoD* activity by forming antagonistic dimers with E-proteins (Benezra et al. 1990; Lu et al. 1999; for review, see Massari and Murre 2000; Berkes and Tapscott 2005). In addition, notch signaling induces expression of the hairy enhancer

of split (*HES*) family of bHLH proteins, which can directly inhibit gene expression by recruiting corepressors, or through interaction with E-proteins (Kuroda et al. 1999; Nofziger et al. 1999). Despite these and many other distinct mechanisms of inhibiting the *MyoD*-mediated myogenic program of terminal differentiation, the switch from a proliferative myoblast to a terminally differentiated myocyte appears to be a highly regulated transition and *MyoD*/*Myf5* are nodal points that integrate the balance between multiple growth and differentiation signals (Weintraub et al. 1991).

One model system for studying the molecular events that regulate the transition between myoblast growth and myocyte differentiation is the pediatric tumor, rhabdomyosarcoma (RMS). RMS cells are characterized by the expression of myogenic specification genes, such as *MyoD* and *Myf5*, yet they do not complete a normal program of myogenesis (for review, see Merlino and Helman 1999). Multiple genetic defects are associated with RMSs (Merlino and Helman 1999 and references therein; Graf Finckenstein et al. 2008), and numerous studies have

⁵Corresponding author.

E-MAIL stapscot@fhcrc.org; FAX (206) 667-6524.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1765109>.

identified different individual pathways that inhibit myogenesis (Hahn et al. 2000; Puri et al. 2000; Sharp et al. 2002; Langenau et al. 2007; Wang et al. 2008), making it difficult to design a single curative therapy. The molecular mechanism behind the development of human RMSs, especially the embryonal RMS, is largely unknown.

In our initial study on the differentiation defect in RMSs (Tapscott et al. 1993), we demonstrated that MyoD was present as a heterodimer with E-proteins and was able to bind an E-box in gel shift assays *in vitro*, but was a poor transcriptional activator *in vivo*, despite the fact that the activation domain of MyoD fused to the DNA-binding domain of Gal4 was very active. In this present study, we further examined the molecular mechanisms that prevent MyoD activity in RMSs, focusing on the RD RMS cell line derived from an embryonal RMS (McAllister et al. 1969). We demonstrate that MSC and an inhibitory splice form of E2A compete with MyoD for dimerization with the full-length functional E-proteins, and that establishing a dimer between MyoD and the full-length E-protein is sufficient to broadly suppress inhibitory factors in RMSs and induce terminal differentiation. Because these inhibitory factors are also expressed in regulative growth phases of normal myogenesis, our data demonstrate that a regulated balance between activation and repression of myogenic genes is integrated by specific E-protein function in the context of a heterodimer with either MyoD or other antagonistic bHLH factors. The demonstration that multiple inhibitory mechanisms can be suppressed and myogenic differentiation can be induced in the RD RMSs by increasing the abundance of MyoD:E-protein heterodimers suggests a central integrating function that can be targeted to force differentiation in muscle cancer cells and to further improve existing therapies.

Results

MSC competes with MyoD for a limiting amount of E2A

Our prior studies indicated that MyoD is an inefficient transcriptional activator of E-box-driven reporters in the RD RMS cell line, and several additional embryonal and alveolar cell lines, whereas the activation domain of MyoD maintained function when fused to the heterologous Gal4 DNA-binding domain when tested in RMS cell lines (Tapscott et al. 1993). Based on these observations, we decided to investigate the dimerization partners of MyoD and E-proteins in RD cells. We used LC MS/MS (liquid chromatography and tandem mass spectrometry) to separate and identify protein complexes purified through a modified tandem affinity purification (TAP) procedure at a physiologic salt concentration (Supplemental Fig. S1; Rigaut et al. 1999; Gingras et al. 2005). The E2A gene encodes two alternatively spliced transcripts, E12 and E47, which differ in the exon encoding the bHLH region. Both E12 and E47 dimerize with MyoD and the other myogenic bHLH proteins and we chose to TAP-tag E12 for our protein interaction studies. We established stable RD cell lines expressing either TAP tag alone,

a C-terminal TAP-tagged MyoD, or a C-terminal TAP-tagged E12, and verified that each fusion protein was active in transfection assays (data not shown). To isolate intact, DNA-bound protein complexes, we carried out the purification procedure in the presence of fragmented chromosomal DNA. Nuclear extract was prepared by treating isolated nuclei with micrococcal nuclease to fragment the chromosomal DNA to mostly tri-, bi-, and mononucleosomal size prior to the affinity purification (Materials and Methods; Supplemental Fig. S1). Affinity purified MyoD and E12, along with their associated proteins, were subjected to trypsin digestion, separation, and analyzed by ThermoFinnigan LTQ FT mass spectrometry. For this study we focused on bHLH dimerization partners of MyoD and E12 identified in the purified complexes. TAP-tagged MyoD coelutes with the known E-protein heterodimer partners: E2A (E12 and E47), E2.2, and HEB, and TAP-tagged E12 interacts with the myogenic bHLH proteins MyoD, Myogenin, and Myf5, but also copurifies with the inhibitory bHLH protein MSC (Table 1). MSC, also called MyoR, has been shown to form heterodimers with E proteins, bind to similar E-boxes as MyoD:E-protein heterodimers, and act as a repressor of MyoD-induced gene expression (Lu et al. 1999).

Western analysis verified that MSC and E12 copurify from RD cells, whereas MyoD copurifies with E12 but not MSC, demonstrated by either TAP purification (Fig. 1A) or coimmunoprecipitation (Fig. 1B). Gel shift assays using *in vitro* transcribed and translated MSC, E12 and MyoD showed that MSC:E12 heterodimers bind the "right" E box from the CKM (muscle creatine kinase) enhancer (Fig. 1C, left panel). Gel shift assays with nuclear extracts from RD cells show E-box-binding protein complexes that contain MyoD:E-protein and MSC:E-protein, based on supershift with antisera to each protein (Fig. 1C, right panel; data not shown). Supplementing the nuclear extract with additional *in vitro* translated E12 enhanced the abundance of the MyoD:E-protein and the MSC:E-protein bands, indicating that the amount of E-protein in the RD nuclear extracts was the limiting factor for the formation of both MyoD and MSC heterodimers. Therefore, the data indicate that MyoD and MSC compete for a limiting pool of E-proteins in RD cells. Similar to previous studies in other cell types, transient transfection of MSC strongly antagonizes the activity of MyoD:E12 heterodimer on the CKM-luciferase reporter in RD cells (Fig. 1D). It is interesting, however, that while transfection of MyoD or E12 alone showed very little activation of the CKM-luciferase, cotransfection of MyoD and E12 strongly induces the CKM-luciferase reporter, suggesting that both factors are in limited abundance endogenously, which can be compensated by stoichiometrically increasing the amount of both proteins.

An E2A splice variant in RD and other RMS cells activates less on myogenic promoters

The genomic locus encoding E2A is shown to be a hot spot for generating translocation and fusion proteins such as E2A-PBX and E2A-HLF, which cause human lymphocytic

Yang et al.

Table 1. The bHLH factors identified by TAP of MyoD-cTAP (bottom panel) and E12-cTAP (top panel) and detected by LC MS/MS from RD cells

bHLH proteins of E12-cTAP	Number of unique peptides identified	Number of total peptides	Percent of sequence coverage
Myoblast determination protein 1 (MyoD)	10	12	43%
Myogenic factor 5 (Myf5)	1	1	5%
Myogenin (Mgn)	5	5	25%
Transcription factor 12 (HEB)	13	16	21%
^a Transcription factor 4 (E2-2/SEF2)	9	9	37%
^b Transcription factor 3 (E2A)	20	28	43%
bHLH protein scleraxis	3	3	15%
Musculin/MyoR (MSC)	9	12	52%
bHLH proteins of MyoD-cTAP	Number of unique peptides identified	Number of total peptides	Percent of sequence coverage
Myoblast determination protein 1 (MyoD)	6	10	16%
Transcription factor 12 (HEB)	11	12	28%
^c Transcription factor 4 (E2-2/SEF2)	8	12	22%
^d Transcription factor 3 (E2A)	11	14	33%

^aTranscription factor 4 (E2-2): Peptides belong to the common sequence shared by either SEF2-1A or 1B. For SEF2-1A, it misses the N-terminal 183 residues of SEF2-1B. For SEF2-1B-only peptide, five unique peptides out of six total peptides identified.

^bTranscription factor 3 (E2A): Peptides belong to the common sequence shared by either E12 or E47. For E47-only peptides, five unique peptides out of five total peptides identified. For E12-only peptides, seven unique peptides out of 13 total peptide identified.

^cTranscription factor 4 (E2-2): Peptides belong to the common sequence shared by either SEF2-1A or 1B. For SEF2-1A, it misses the N-terminal 183 residues of SEF2-1B. For SEF2-1B-only peptide, five unique peptides out of six total peptides identified.

^dTranscription factor 3 (E2A): Peptides belong to the common sequence shared by either E12 or E47. For E47-only peptides, two unique peptides out of two total peptides identified. For E12-only peptides, one unique peptide out of one total peptide identified.

leukemia (Inaba et al. 1992; Lu et al. 1994; see the Discussion). To investigate whether there is any mutation in the coding sequence of the full-length E2A, we cloned and sequenced the E2A transcripts from the RD cell line. We identified the known wild-type E2A transcripts and an in-frame splice variant that removes exons 3 and 4 (Fig. 2A), part of the activation domain 1 (AD1) region of E2A (Markus et al. 2002). Database searches identify a transcript, missing exon 3 and 4 of the full-length E2A/TCF3, reported in a collection of mammalian cDNAs with a NCBI accession number of AAI10581 (Strausberg et al. 2002). In addition, we found several ESTs with this splice variant (see Discussion), and we will refer to this splice variant as E2A-2/5. RT-PCR of cDNA from four additional RMS cell lines demonstrated the expression of MSC, E2A, and the E2A-2/5 splice form (Fig. 2B, lanes 2–5); and RNAs from six different primary RMS showed some expression of both MSC and E2A-2/5 (Fig. 2B, lanes 6–11). In addition, low amounts of E2A-2/5 were identified in cultured primary human myoblasts (Fig. 2B, lanes 12–14), which also expressed MSC RNA (Supplemental Fig. 2). Reanalysis of the peptide mass spectral data from both MyoD and E12-associated proteins detects the tryptic peptide generated from the joining of exon 2 and 5 with a high confidence score (Supplemental Fig. S3A), and coimmunoprecipitation using a flag-tagged E12 showed the interaction between E12 and E12-2/5 in RD cells (Supplemental Fig. S3B). Therefore, the E2A-2/5 protein is copurified with both MyoD and E2A in RD cells. Finally, we determined that both E12 and E47 exist as the E2A-2/5 splice form in RD cells (data not shown).

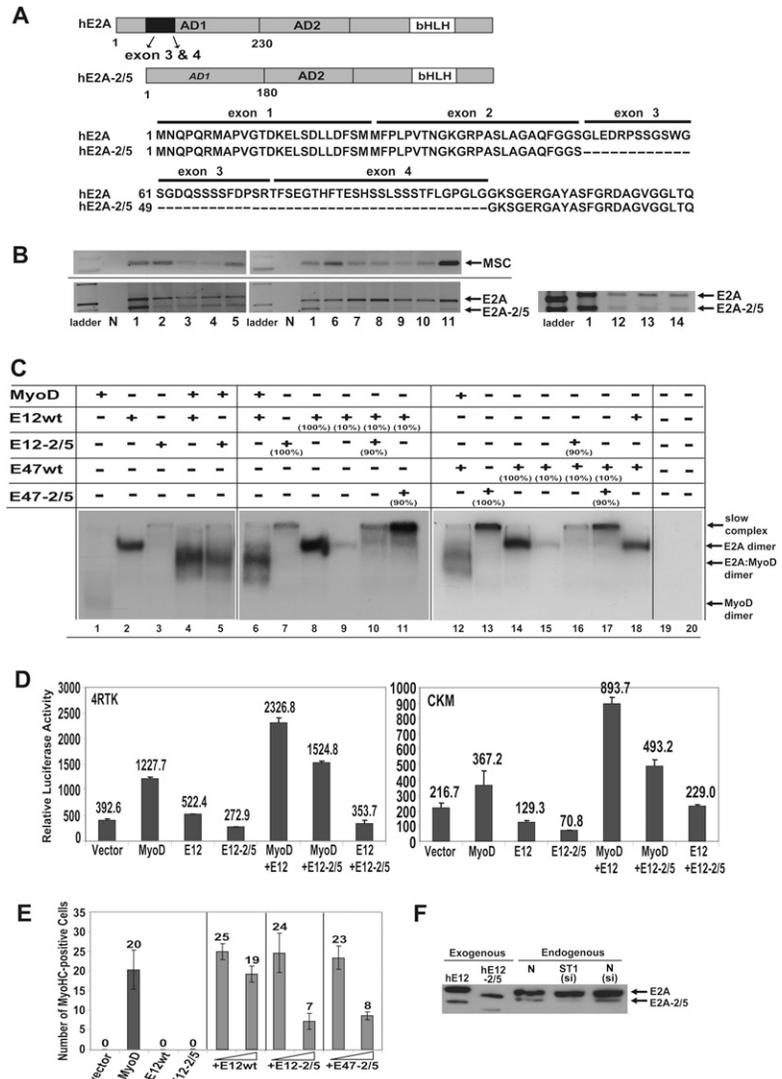
Gel shift assays revealed that E12-2/5 forms a heterodimer with MyoD and binds to the “right” E box from the

CKM enhancer similar to the full-length E12:MyoD heterodimer (Fig. 2C, lanes 4,5). The full-length E12 itself binds the E-box as a homodimer (Fig. 2C, lane 2), whereas the E12-2/5 forms a weak, slowly migrating complex suggesting a multimer, or other higher-order protein complex, with some residual DNA interaction (Fig. 2C, cf. lanes 2 and 3). In addition to forming heterodimers with MyoD, E12-2/5 and E47-2/5 also interact with the full-length E12 and E47, respectively, and partly preclude the formation of the full-length E2A homodimer complex, presumably by incorporating the full-length proteins into a higher-order complex (Fig. 2C, cf. lanes 9 and 10,11, cf. lanes 15 and 16,17; see also Supplemental Fig. 4). Competition assays demonstrate that the slowly migrating E47-2/5 and E12-2/5 complexes are competed with a cold self E-box but not a mutant E-box, indicating that this higher-order complex maintains a specific interaction with the E-box sequence (data not shown). It is unclear whether such a higher-order complex exists in vivo, but the in vitro data suggest that the E2A-2/5 splice variant might interfere with the function of the full-length E2A proteins by incorporating them in a multi-protein complex.

Transient transfection studies show that E12-2/5 and E47-2/5 activate transcription of reporter constructs driven by a multimerized E-box (p4R-TK-Luc) and by pCKM-Luc less than the full-length forms, either alone or in combination with MyoD (Fig. 2D; data not shown). Similar transfection results were obtained in 3T3 mouse embryonic fibroblasts (data not shown). Further supporting an inhibitory role of the E2A-2/5 splice form, cotransfection suppressed the ability of MyoD to induce muscle differentiation in 3T3 fibroblasts, as indicated by the

Yang et al.

Figure 2. A novel E2A splice variant, E2A-2/5, isolated from RD cells activates less than the full-length E2A. (A) Sequencing identified a novel E2A splice variant, E2A-2/5, in RD cells missing exons 3 and 4. The *top* panel shows schematic domain representations of E2A (E12/E47) and E2A-2/5 proteins. The *bottom* panel shows the amino acid sequence at the location of exon 3 and 4 in the activation domain 1 of E2A. (B) E2A-2/5 and MSC were detected in multiple independent RMS cell lines, primary tumors, and normal human primary myoblasts. Total RNAs isolated from cell lines 1-RD (e), 2-Rh1 (e), 3-RH18 (e), 4-RHJT (a), 5-RH28 (a); primary tumors 6-RMS60 (e), 7-RMS61 (a), 8-RMS62 (a), 9-RMS63 (a), 10-RMS64 (e), 11-RMS65 (e); and human primary myoblast cell lines 12–14 were subject to RT-PCR to detect expression level of MSC and E2A. The “e” represents the embryonal and the “a” represents the alveolar subtype of RMSs. (C) EMSA using in vitro transcribed and translated (TnT) products showed that E12-2/5 dimerizes with MyoD similar to the full-length E12 (shown in lanes 4,5). The full-length E12 binds the E-box as a homodimer, whereas the E12-2/5 does not show homodimer binding but does have a faint slower migrating species that might be a higher-order multimer (cf. lanes 2 and 3). E2A-2/5 prevents the full-length E2A homodimers from binding to DNA (cf. lanes 10,11 and 9, cf. lanes 16,17 and 15). Numbers in percentage represent how much TnT product was used for each EMSA reaction. Lanes 19 and 20 represent reticulocyte alone and probe alone, respectively. (D) MyoD:E2A-2/5 activates less efficiently in RD cells than MyoD:E2A. Transient transfections of RD cells with MyoD and/or E2A splice variants were done in the presence of 4RTK or CKM luciferase reporter and pCS2-βLacZ as an internal control. (E) In the 3T3 cell, MyoD:E2A-2/5 activates the myogenic program less efficiently than MyoD:E12, estimated by the numbers of myosin heavy chain-positive cells. The amount of MyoD remained constant (0.75 μg) while increasing the amount of E12 or E2A-2/5 as indicated by the gradient triangles from 0.05 to 0.25 μg. (F) Western blot showing relative amounts of E2A and E2A-2/5 in RD cells. Lanes 1 and 2 are cells transfected with hE12 and hE12-2/5 to show migration. (N) Endogenous E-protein in untransfected RD cells; [ST1(si)] RD cells transfected with siRNA to E2A-2/5 (see Supplemental Fig. S5A); [N(si)] control scrambled siRNA.



affinity of MSC with the different E-protein isoforms has not been determined.

In contrast to the partial rescue of myogenesis by expressing a dnMSC, RNAi knockdown of MSC and/or E2A-2/5 was not sufficient to induce muscle gene expression (Supplemental Fig. 5A,B), suggesting that additional factors might limit the formation of functional MyoD:E-protein dimers. However, when MyoD was ectopically expressed in RD cells following retroviral transduction, we did observe a partial rescue of MyoD activity with the knockdown of MSC or E2A-2/5. Compared with MSC knockdown alone, knockdown of MSC in the presence of retroviral expression of MyoD resulted in ~10-fold higher levels of CKM expression. Similarly, compared with E2A-2/5 knockdown alone, knockdown of E2A-2/5 in the presence of retrovirally expressed MyoD resulted in an

approximately eightfold increase in CKM expression (Supplemental Fig. 5C).

Dimerization of MyoD with a full-length E-protein rescues myogenesis

We observed a partial rescue of MyoD activity by retroviral expression of dnMSC, MyoD, or E12, and the RNAi knockdown of dnMSC and/or E2A-2/5 in the presence of overexpressed MyoD. We decided to use a forced dimer in which a peptide linker physically joins MyoD and E protein together (Neuhold and Wold 1993; Dilworth et al. 2004). To determine whether a MyoD~E-protein forced heterodimer could resist competition from bHLH repressors and broadly rescue the myogenic defect in RD cells, we compared the activity of MyoD, E-proteins, and the forced

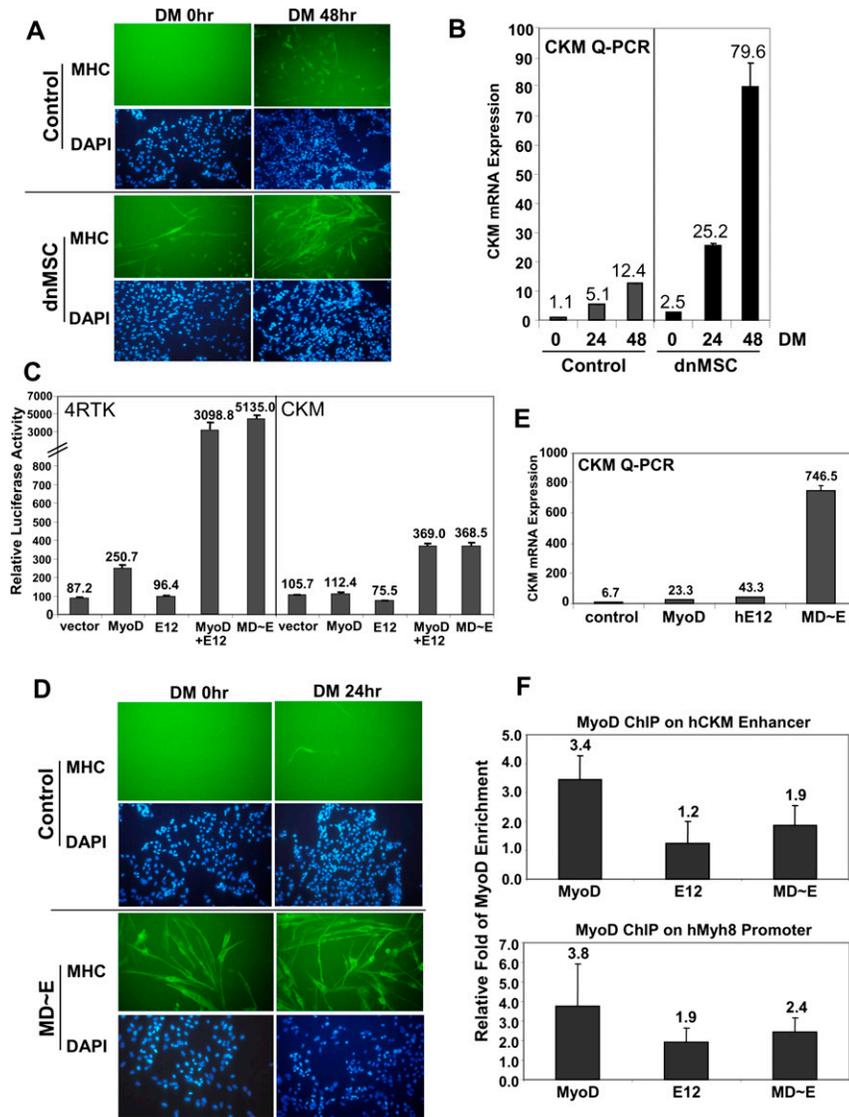


Figure 3. A dnMSC or a MyoD~E12 forced dimer drive RD cells into differentiation. (A) RD cells were infected with retroviruses expressing either vector alone or dnMSC. Infected RD cells were selected in puromycin and differentiated for 0 h, 24 h, and 48 h in insulin transferrin medium. Myogenesis was estimated by immunostaining for myosin heavy chain (MHC) (differentiation at 0 and 48 h shown) and nuclei were stained by DAPI. (B) Corresponding expression of muscle creatine kinase (CKM) was estimated by qPCR. Values were normalized to Timm17b, an RNA that does not change with myogenic differentiation. (C) Luciferase assay was conducted on RD cells transiently transfected with MyoD and/or E12 or MyoD~E12 forced dimer. (D) Myosin heavy chain (MHC) and DAPI staining of RD cells infected by retroviruses expressing either a control or MyoD~E forced dimer. (E) qPCR estimated CKM expression in RD cells infected by retroviruses expressing either vector alone, MyoD, E12, or MyoD~E12 dimer. (F) Relative folds of enrichment of MyoD binding on the Myh8 (myosin heavy chain 8) promoter and the CKM enhancer averaging three independent ChIP experiments from cells infected by retroviruses expressing MyoD, E12, or the MD~E12 forced dimer versus the control virus at 24 h of differentiation.

MyoD~E-protein heterodimer in the RD cell line. The combination of MyoD and E12, or the forced MyoD~E12 heterodimer, robustly activated expression of the p4R-TK-Luc and pCKM-luc reporters (Fig. 3C). In addition, retroviral expression of the forced heterodimer in RD cells strongly induced fusion and expression of myosin heavy chain (Fig. 3D). Affymetrix HG U133 plus 2.0 arrays were used to compare transcript abundance from RD cells infected with the retroviruses expressing the forced heterodimer to vector-alone infected RD cells. Using criteria of a twofold change and a false discovery rate (FDR) <0.05, 1414 features had significantly higher signal and 571 features had lower signal in the RD cells infected with the retroviruses expressing the forced heterodimer compared with vector-alone infected RD cells. The genes activated by the forced heterodimer in RD cells showed a highly significant overlap with genes regulated by MyoD expressed in normal human fibroblasts (Supplemental Fig. S6). Using a twofold criteria for both sets, one-third of the

genes activated by the forced dimer in RD cells are also activated by wild-type MyoD in fibroblasts, and using a fourfold criteria for RD cells one-half of the genes overlap with MyoD-regulated genes in fibroblasts (both with $P < 2.2 \times 10^{-16}$). The genes activated by the forced dimer in RD cells that overlap with MyoD-regulated genes in human fibroblasts are significantly enriched in skeletal muscle genes based on gene ontology (GO) analysis (Supplemental Tables 1, 2), indicating that the forced heterodimer is activating a core program of skeletal muscle gene expression in the RD cells. Representative genes are shown in Table 2, and validation of a subset of genes shown in Supplemental Figure S7. In addition, cell cycle analysis demonstrated a marked decrease in the percentage of cells incorporating BrdU during a 24-h pulse in low serum (67.4% of control cells vs. 26.3% of cells expressing the forced dimer), and FACS showed an increase in G1 (32% of control cells vs. 57% with forced dimer) and decrease in S-phase (46% control cells vs. 23% forced dimer) (Supplemental Fig. 8).

Yang et al.

Table 2. Representative genes regulated by MD~E12 forced dimer in RD cells

GenBank ^a	Unigene ^b	Symbol ^c	Name and function ^d	FC(log2) ^e
<u>Muscle structure/metabolism genes</u>				
M36769	Hs.534028	MYH8	Myosin heavy chain 8, skeletal muscle, perinatal	8.233
BF939176	Hs.381047	MYOZ2	Myozenin 2	6.624
AF020769	Hs.118845	TNNC1	Troponin C type1 (slow)	6.558
AW975934	Hs.654592	TTN	Titin	5.996
NM_001824	Hs.334347	CKM	Creatine kinase, muscle	5.785
NM_002470	Hs.440895	MYH3	Myosin, heavy chain 3, skeletal muscle, embryonic	5.303
NM_003281	Hs.320890	TNNI1	Troponin I type 1 (skeletal, slow)	4.912
AF116676	Hs.463300	MYL4	Myosin, light polypeptide 4, alkali; atrial, embryonic	4.063
AI819052	Hs.658808	MIB1	Mindbomb homolog 1 (<i>Drosophila</i>)	3.928
<u>Transcriptional factors</u>				
NM_006917	Hs.26550	RXR γ	Retinoid X receptor, γ	3.757
AL536517	Hs.654474	MEF2C	Myocyte enhancer factor 2C	3.696
NM_017826	Hs.124519	SOHLH2	Spermatogenesis and oogenesis-specific bHLH2	2.102
NM_014707	Hs.196054	HDAC9	Histone deacetylase 9	1.799
D13889	Hs.504609	ID1	Inhibitor of DNA binding 1	1.241
AI829724	Hs.314327	MEF2D	Myocyte enhancer factor 2D	1.234
NM_002479	Hs.2830	MYOG	Myogenin	1.195
NM_002229	Hs.25292	JUNB	jun B proto-oncogene	-1.114
NM_005586	Hs.520119	MDF1	MyoD family inhibitor (I-mfa)	-1.268
BC002712	Hs.25960	MYCN	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	-1.535
NM_003670	Hs.171825	BHLHB2	Basic HLH containing, class B, 2 (DEC2)	-2.13
BG528420	Hs.643910	SOX4	SRY (sex determination region Y)-box 4	-2.248
<u>Homeobox factors</u>				
H49805	Hs.194725	ONECUT2	One cut homeobox 2	2.458
AW769732	Hs.155644	PAX2	Paired box 2	2.261
AA194168	Hs.42146	PAX3	Paired box gene 3	-1.095
AL049381	Hs.654412	PBX1	Pre-B-cell leukemia txn factor-1	-1.106
AK000445	Hs.658823	HOXC9	Homeobox C9	-1.164
R59304	Hs.128193	MKX	Mohawk homeobox	-1.265
AW137982	Hs.659337	HOXA3	homeobox A3	-1.267
AA040332	Hs.249196	DLX6	Distal-less homeobox 6	-1.489
AI288759	Hs.652666	ZFH3	Zinc finger homeobox 3	-1.742
NM_004527	Hs.438	MEOX1	Mesenchyme homeobox 1	-2.022
<u>Cell cycle machinery</u>				
N33167	Hs.106070	CDKN1C	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	3.897
AL567411	Hs.500015	CDK5R1	Cyclin-dependent kinase 5, reg. subunit 1 (p35)	2.076
NM_000389	Hs.370771	CDKN1A	Cyclin-dependent kinase inhibitor 1A (p21, Cip1)	2.054
AK022122	Hs.657656	MTBP	Mdm2, p53-binding protein	-1.155
BC021222	Hs.696431	CNNM2	Cyclin M2	-1.048
<u>Signaling pathway</u>				
U23460	Hs.351887	CAMK2B	Calcium/calmodulin-dep. protein kinase II β	3.78
AA526844	Hs.556600	MYLK	Myosin, light polypeptide kinase	3.553
NM_006477	Hs.73088	RASL10A	RAS-like, family 10, member A	3.363
BC006107	Hs.437126	ARHGAP9	Rho GTPase-activating protein 9	3.386
L08835	Hs.133379	TGFB2	Cystostrophia myotonic-protein kinase	2.523
M19154	Hs.631596	DMPK	Transforming growth factor, β 2	2.458
NM_005592	Hs.521653	MUSK	Muscle, skeletal, receptor tyrosine kinase	2.118
BF343625	Hs.527021	MRAS	Muscle RAS oncogene homolog	2.109
U34349	Hs.25363	PSEN2	Presenilin 2	2.099
U15979	Hs.533717	DLK1	Delta-like 1 homolog (<i>Drosophila</i>)	1.402
U77914	Hs.224012	JAG1	jagged 1	-1.117
NM_012258	Hs.234434	HEY1	Hairy enhancer of split related with YRPW motif 1	-1.381
AJ224869	Hs.593413	CXCR4	Chemokine (C-X-C motif) receptor 4	-1.414
NM_000435	Hs.8546	NOTCH3	Notch homolog-3	-1.606
BC000055	Hs.269512	FSTL1	Follistatin-like 1	-1.617
AA194149	Hs.471162	RAPH1	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1	-1.695
AF003934	Hs.616962	GDF15	Growth differentiation factor 15	-2.096
S69738	Hs.303649	CCL2	Chemokine (C-C motif) ligand 2	-2.88
NM_020997	Hs.656214	LEFTY1	Left-right determination factor 1	-3.188
AW007532	Hs.635441	IGFBP5	Insulin-like growth factor-binding protein 5	-4.939

^aGeneBank accession number.^bUnigene identifier.^cGene symbol.^dGene name and function inferred from Medline research.^eLog₂ of fold of change of RD cells infected by retroviruses expressing MyoD~E forced dimer relative to cells infected by control viruses at 24 h of differentiation.

To determine whether the difference between MyoD and the forced dimer was due to binding efficiency at genes preferentially activated by the forced dimer, we performed chromatin immunoprecipitation (ChIP) with antisera to MyoD in RD cells. We prepared cross-linked chromatin from RD cells infected with (1) a control retrovirus, (2) a retrovirus expressing the MyoD~E12 forced heterodimer, (3) a retrovirus expressing MyoD, and (4) a retrovirus expressing E12. After three consecutive retroviral infections, cells were grown to confluence and induced in differentiation medium for 24 h prior to formaldehyde cross-linking and qPCR showed that the forced heterodimer-activated CKM expression >10-fold higher than MyoD or E12 alone (Fig. 3E). The E-box-containing regulatory regions for *MYH8* (myosin heavy chain 8, perinatal) and *CKM* are well characterized. ChIP from the cells expressing the monomeric MyoD or E12 resulted in a consistent 1.5-fold to threefold enrichment of MyoD binding relative to the binding of endogenous MyoD in the control cells, and expression of the forced dimer showed enrichment in a similar range (Fig. 3F). Therefore, a simultaneous co-occupancy on DNA plus the synergistic activation by MyoD~E12 heterodimer appear critical for the full activation of the myogenic program. Together, these studies demonstrate that a forced heterodimer of MyoD and E12 increases transcriptional activity of the myogenic genes in RD without significantly increasing MyoD binding to gene regulatory regions.

Forced MyoD~E2A heterodimer broadly suppresses expression of many myogenic inhibitors

As noted in the introduction, many different factors and pathways have been implicated in blocking myogenesis in RMSs, including Id proteins, notch signaling, and kinase signaling pathways, and it is therefore not too surprising that knocking down MSC and E2A-2/5 are not sufficient for a full myogenic differentiation. In contrast, however, the MyoD~E2A forced dimer activates a broad set of myogenic genes, which suggests that it might broadly suppress these multiple inhibitory pathways. Indeed, Western blot analysis shows the expression of MyoD~E2A in RD cells dramatically suppresses the protein levels of MSC and E2A (Fig. 4A), and it is interesting to note that other studies have shown low abundance of E2A in differentiating muscle (Conway et al. 2004). For MSC, the decrease is almost exclusively at the protein level (qPCR data not shown); whereas the MyoD~E2A appears to suppress the RNA of the E2A-2/5 splice form without decreasing the full-length transcript (Fig. 4B). In addition to MSC and E2A-2/5, multiple other inhibitors of myogenesis in the RD cells are suppressed by MyoD~E2A. For example, continued expression of PAX7 has been shown to inhibit myogenesis and Western blot analysis shows that the MyoD~E2A reduced PAX7 protein relative to control cells (Fig. 4C). The expression array data also shows that multiple inhibitors are suppressed at the RNA level (see Table 2; Fig. 4C). For example, DEC1/BHLHB2/SHARP2/Stra13 is a bHLH protein of the HES family that can bind E-boxes and repress

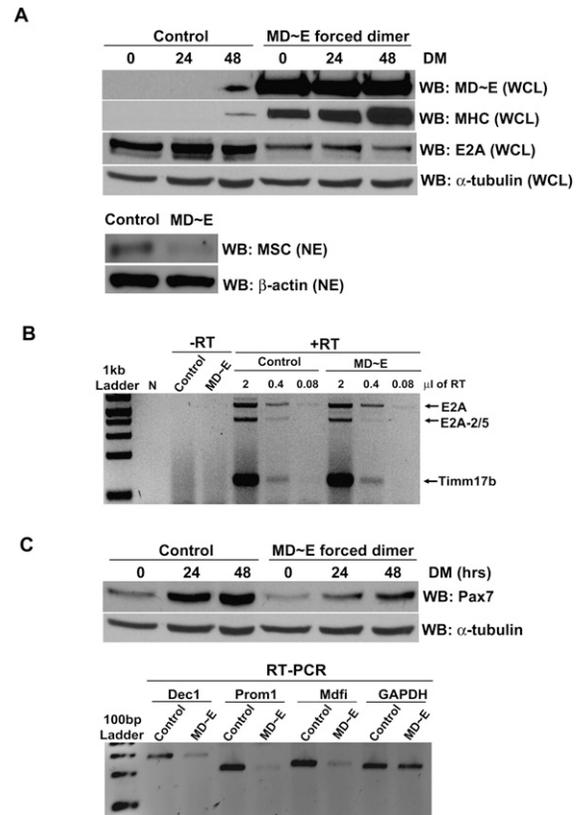


Figure 4. MyoD~E12 forced dimer broadly inhibits myogenic repressors. (A) Reduced E2A and MSC expression in RD cells differentiated by retrovirus-mediated expression of MD~E12 forced dimer. Correspondingly, expression of myosin heavy chain (MHC) is strongly enhanced. Whole-cell lysate (WCL) was immunoblotted for MyoD~E12 dimer, MHC, E2A, and α -tubulin as the loading control. Nuclear extract (NE) was immunoblotted for MSC and β -actin as the loading control. (B) Duplex RT-PCR showed a preferential reduction in E2A-2/5 mRNA level but not the full-length E2A mRNA. Timm17b is a loading control. (C) Immunoblotting showed that expression of Pax7 is reduced in RD cells differentiated by MyoD~E12 forced dimer. RT-PCR showed down-regulation of Dec1, Prominin1, and Mdf1 expression. GAPDH serves as an internal control for RT-PCR.

transcription (St-Pierre et al. 2002). Mdf1/I-mfa is a negative regulator that binds myogenic bHLH proteins and prevents their translocation to the nucleus (Chen et al. 1996; Snider et al. 2001). Mdf1/I-mfa is also known to regulate Hand1 activity by nucleolus sequestration to sustain cell proliferation and renewal of trophoblast cells (Martindill et al. 2007). In addition, our earlier studies on cell cycle regulators demonstrated that a potential role of MDM2 in mediating repression of MyoD-dependent transcription (Fiddler et al. 1996) and subsequently, of p21 in integrating proliferation and myogenic differentiation in RMS cells (Otten et al. 1997). The expression array data shows a down-regulation of MDM2 and an up-regulation of p21 expression when RD cells underwent myogenic differentiation (Table 2).

In recent years, nonhuman RMS models were developed in mouse and zebrafish to address abnormal signaling

events in Sonic hedgehog and Ras signaling pathway, respectively, in RMS development (Hahn et al. 2000; Langenau et al. 2007). We observed a strong down-regulation of IGFBP5 when RD cells were differentiated (Table 2; Supplemental Fig. S7). The down-regulation of IGFBP5 would enhance myogenic differentiation through kinase pathways downstream from IGF (Cobb et al. 2004; Salih et al. 2004). Furthermore, the Notch pathway appears broadly inhibited with down-regulation of NOTCH3, the ligand JAG1, and the Notch-regulated HES family member bHLH gene HEY1, which might function similar to DEC1 in inhibiting myogenesis. It is interesting that the possible formation of inactive Hey1/MyoD was proposed to inhibit MyoD function (Sun et al. 2001), and that Buas et al. (2009) showed that a constitutive expression of Hey 1 blocked early myogenesis in C2C12 cells. Finally, the stem and cancer cell marker PROM1 (Prominin-1 or CD133) (for review, see Mizrak et al. 2008) is also dramatically down-regulated. In summary, the forced MyoD~E2A heterodimer broadly suppresses pathways that promote growth and suppress differentiation and rescues a full myogenic differentiation in RMS cells.

Discussion

Taken together, our results support a new and specific model of RMSs and the mechanisms that regulate the transition from transcriptional repression to activation of myogenic promoters through dimerization equilibrium where the synergistic MyoD:E-protein activity is attenuated by existence of antagonistic DNA-binding dimers or inactive sequestering complexes (Fig. 5). In RD RMSs, MyoD heterodimerizes with an E-protein and that complex binds DNA, both in vitro (Tapscott et al. 1993) and in vivo (present study). We propose and confirm a mecha-

nism of how the myogenic bHLH antagonists could prevent these MyoD complexes from activating a subset of muscle genes and initiating terminal differentiation. First, the bHLH factor MSC competes with MyoD for a limiting amount of E-proteins and binds to the same consensus E-box as the MyoD:E-protein complexes, inhibiting muscle gene activation. In addition, the expression array analysis suggests that DEC1 and HEY1 are also potentially competing for E-proteins and inhibiting MyoD activity, although we did not detect them by mass spectrometry analysis in our E-protein copurification, which could represent variations in protein abundance, complex stability, or DNA-binding characteristics. Second, the limited amount of E-protein in RD cells is partially composed of the E2A-2/5 splice form that lacks exons 3 and 4 encoding part of the first activation domain. Compared with heterodimers of MyoD and the full-length E2A, the MyoD:E2A-2/5 heterodimers show similar in vitro binding characteristics to an E-box but are relatively inefficient transcriptional activators. Furthermore, the E2A-2/5 splice can form a higher-order protein complex, either a multimer or complex with other reticulocyte lysate proteins demonstrated by in vitro EMSA, and interact with the full-length E2A protein, demonstrated by both EMSA and coimmunoprecipitation, thereby providing another mechanism for limiting the amount of the full-length E2A for dimerization with MyoD. Therefore, disruption of a balanced equilibrium between bHLH factors that dimerize with E-proteins could result in an excess of repressive rather than active complexes when the full activation of MyoD is required for terminal differentiation. Coexisting with these bHLH factors are multiple other inhibitors of MyoD activity, Pax3 and Pax7, and activated inhibitory signaling pathways, such as Ras (Marampon et al. 2006; Langenau et al. 2007) and Notch pathways (Kuroda et al. 1999; Nofziger et al. 1999).

The major finding in our study is that the forced MyoD~E2A heterodimer overcomes and suppresses this multiplicity of inhibitory factors and shows that the transcriptional synergy between MyoD and E2A protein is required for a complete myogenic rescue in human embryonal RMS cells. Our studies suggest that MSC, E2A-2/5, and other multiple inhibitors of myogenesis in RD cells are not cancer-specific negative regulators of MyoD, nor an abnormal amalgam of inhibitors co-opted from other processes, but rather that they represent a normal regulative mechanism for impeding a portion of MyoD activity during a normal phase of regulative growth. This insight reveals a new model for studying the multiple factors that regulate the transition from regulative growth to differentiation, and indicates that a fundamental defect in RMSs results in an "arrested myoblast phase" that fails to transition from regulative growth to differentiation. An arrested transition state has been recognized as a classic feature of hematopoietic cancers, such as promyelocytic leukemia. Our studies suggest that a similar transition state arrest occurs in solid tumors such as RMSs.

MSC/MyoR/ABF-1 was first identified as a myogenic repressor in a screen for bHLH proteins in muscle that

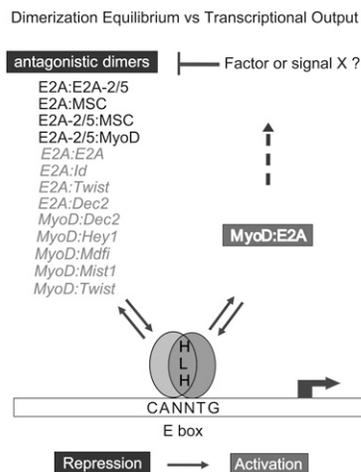


Figure 5. A competitive equilibrium between antagonist, partial agonist, and active bHLH dimers modulate MyoD:E2A activity to maintain a myoblast state, whereas shifting the equilibrium to favor MyoD:E2A heterodimers activates a process (factor X or signal X) that broadly suppresses the inhibitors and locks in a differentiated state.

dimerize with E-proteins (Lu et al. 1999). MSC is expressed during embryonic myogenesis in the dorsomedial lip of the dermomyotome and in limb migratory muscle precursor cells (von Scheven et al. 2006), and in the early phase of adult mouse muscle regeneration (Zhao and Hoffman 2006), consistent with a role in the growth and/or migration phase of myogenesis. MSC is expressed in mouse C2C12 myoblasts and suppressed with muscle cell differentiation (Lu et al. 1999). Despite the molecular characterization of MSC as an inhibitor of the myogenic bHLH proteins, MSC is necessary for the formation of facial skeletal muscle because mice with a genetic disruption of both MSC and a closely related bHLH factor *Tcf21/Capsulin* lack facial muscles (Lu et al. 2002). MSC: E-protein heterodimers should have an overlapping, or possibly identical, set of genomic binding sites as MyoD: E-protein heterodimers, and it is possible that MSC is necessary for an early regulative growth phase of myogenesis and suppresses a late phase of differentiation. Indeed, expression of MSC in MEFs has a complex pattern of activation and suppression of genes regulated by MyoD (S.J. Tapscott, unpubl.), indicating both positive and negative roles in specification and differentiation of muscle cells. This combination of both positive and negative roles might both lock the cell into an early state of differentiation and also might be the reason that knockdown of MSC results in less expression of some muscle markers.

Three genes comprise the E-protein family in vertebrates: *TCF3/E2A*, *TCF4/E2-2*, and *TCF12/HEB*. All three E-proteins are expressed in skeletal muscle and form heterodimers with MyoD and the myogenic bHLH factors. Genetic studies demonstrate a necessary and largely redundant role for each E-protein gene in T- and B-cell development (Zhuang et al. 1996), as well as redundant and region-specific roles in brain development (Flora et al. 2007). All three E-proteins are expressed during skeletal myogenesis and each can heterodimerize with MyoD and the other myogenic bHLH proteins to form functional transcription factor complexes. No specific muscle phenotype has been reported with any of the individual E-protein knockouts, suggesting broadly redundant functions in skeletal myogenesis. The *E2A* gene generates alternatively spliced transcripts, E12 and E47, that differ only in the bHLH region. *E2A* transcripts are expressed in most tissues examined (Massari and Murre 2000). However, the *E2A* proteins have been shown to be post-transcriptionally regulated. For example, E47 degradation is regulated by Notch signaling and MAP kinase activity (King et al. 2007). In lymphocytes, *E2A* proteins regulate proliferation and differentiation (Nutt and Kee 2007), and have been shown to regulate cell cycle factors, such as p57 and CDK6 (Rothschild et al. 2006; Schwartz et al. 2006). Two activation domains, AD1 and AD2, are located at the N-terminal half of the proteins and have both redundant and distinct functions. AD1 can function as either an activation domain or a repression domain depending on cellular context (Markus et al. 2002; Bhalla et al. 2008). A conserved LDFS motif in AD1 recruits the SAGA complex (Massari et al. 1999) and is also the region

that interacts with the repressor protein ETO (Zhang et al. 2004), which might account for the alternative activation and repression functions of AD1. It is interesting to note that the *E2A-2/5* splice maintains the LDFS motif and the potential to interact with either SAGA or ETO.

Several mechanisms have been identified that regulate the activity of E-proteins. The Id proteins are HLH factors that lack the basic DNA-binding domain and competitively heterodimerize with E-proteins, preventing DNA binding to an E-box (Benezra et al. 1990). In addition, E-proteins can heterodimerize with bHLH transcriptional repressors, such as HES proteins (*Dec2*) (Azmi et al. 2004; Fujimoto et al. 2007) or Twist (Kophengnavong et al. 2000) or, as in this study, MSC (see also Lu et al. 1999). We have now identified a splice form of *E2A*, *E2A-2/5*, that removes a portion of the AD1 domain and is less transcriptionally active than the full-length *E2A* protein. This splice form is as abundant as the full-length *E2A* in RD RMS cells, and is also expressed in a total of five independent RMS cell lines and six independent primary RMSs. Although previously not characterized, ESTs representing the *E2A-2/5* splice have been sequenced in libraries generated from human thymus, B cells, HeLa cells, mammary adenocarcinoma cells, and an epithelioid carcinoma. In addition, we could detect this splice form by RT-PCR in RNAs from three different normal human myoblast cell lines. Therefore, although this might not be an abundant transcript in most cell types, it is likely expressed in specific phases of normal development and in cancer cells in addition to RMSs. However, it is important to note that we have not yet detected this splice form in RNAs from mouse embryos, or in C2C12 muscle cells (data not shown). It should be mentioned here that myogenin and some structural genes are expressed in RD cells, indicating the mechanism in activating the early myogenic program remains intact. As our previous studies demonstrated, Pbx-mediated stable recruitment of MyoD facilitates activation of myogenin promoter through binding to the noncanonical E box (Berkes et al. 2004). The repression mediated by MSC or *E2A-2/5* is likely to be exerted through binding to the canonical E box. Therefore, the regulatory usage of E-boxes on a specific promoter could represent diversity and selectivity in dimer binding, and subsequently, the magnitude of gene activation. In this regard, the E-protein appears to be a critical integrator of growth and differentiation signals, possibly mediated by its differential activity in the context of dimerization partners.

It is likely that E-proteins play a central role in the transition between regulative growth and differentiation in other normal and cancer cells as well. For example, disruption of balance between MSC/ABF-1 and E-proteins results in the reprogramming of neoplastic B-cells in Hodgkin lymphoma (Mathas et al. 2006). *E2A* proteins can repress E-cadherin expression in epithelial cells and promote epithelial-mesenchymal transitions (EMT), suggesting a role of *E2A* proteins in tumor progression (for review, see Peinado et al. 2007). The AML1-ETO translocation fusion protein targets E-proteins to modulate their activity as a cause of AML (Zhang et al. 2004), and

Yang et al.

E2A-PBX and E2A-HLF translocations are associated with acute lymphocytic leukemia (Inaba et al. 1992; Lu et al. 1994). Although the fusion proteins might have many different functions, our study suggests that they might promote the growth phase of E-protein function and impede the transition to the differentiation phase.

The mechanism of how human embryonal RMSs develop is still poorly understood. Our studies, for the first time, show that the apparent mechanistic complexity, indicated by the existence of multiple transcriptional antagonists and inhibitory signaling pathways, represents an arrested progress through a transitional state that is normally regulated by the relative abundance of heterodimers between MyoD and the full-length E2A proteins. The demonstration that multiple inhibitory mechanisms can be suppressed and myogenic differentiation can be induced in the RD RMSs by increasing the abundance of MyoD:E-protein heterodimers suggests a central integrating function that can be targeted to improve existing therapies.

Materials and methods

Plasmid construction

pT81-4RTK-luciferase (Weintraub et al. 1991) and pGL3-CKM-luciferase (Berkes et al. 2004) have been described previously. MSC was cloned from RD cells into the pCS2(+) vector and a dnMSC was generated by site-directed mutagenesis of the DNA-binding site (Lu et al. 1999). Human E47-2/5 and E12-2/5 were cloned from RD cells by RT-PCR. All retroviral expression constructs were cloned into the pCLBabe vector (Dilworth et al. 2004). Both pCLBabe-cTAP and pCLBabe-MyoD~E12 forced dimer were kindly provided by F.J. Dilworth (Dilworth et al. 2004).

Cell culture and transient transfections and immunofluorescence

RD cells and NIH 3T3 fibroblasts were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% bovine calf serum (HyClone), penicillin, and streptomycin. Differentiation medium is serum-free DMEM with 10 μ g/mL of insulin and transferrin. Cells were transiently transfected using Superfect (Qiagen) as described previously (Berkes et al. 2004) with pCS2- β -galactosidase as an internal control. Cells were fixed in 2% paraformaldehyde and immunostained for myosin heavy chain using MF20 antibody (supernatant).

Immunoprecipitation and detection

Immunoprecipitation was done in nuclear extract acquired as described below in the presence of antibody against either MSC (F20, Santa Cruz Biotechnologies) or human E2A (G193-86, BD Biosciences). Half of the washed protein G beads were treated with λ protein phosphatase (New England Biolabs) before boiled in SDS loading dye. Immunoblotting was carried out with antibodies against MSC, E2A (G98-271, BD Biosciences), MyoD (5.8A supernatant), Pax7 (kindly provided by Dr. B. Olwin), α -tubulin (Sigma), β -actin (Sigma), or MHC (MF20).

TAP and LC MS/MS

TAP was modified from both the Séraphin and Gingras protocols. Cells were suspended in buffer A (10 mM HEPES, 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT, complete protease inhibitors [Roche]) and cytoplasmic membrane was physically sheared. Nucleus

fraction was resuspended in buffer B (20 mM HEPES, 0.5 mM EDTA, 100 mM KCl, 10% glycerol, 2 mM DTT, 3 mM CaCl₂, 1.5 mM MgCl₂) in the presence of protease and phosphatase inhibitors. Isolated nuclei were treated with micrococcal nuclease and slowly extracted cold. After being incubated with IgG-conjugated sepharose for 2 h, the beads were washed in Buffer B with 150 μ M of KCl and then incubated for TEV protease digestion (TEV digestion buffer: 10 mM HEPES, 150 mM KCl, 10% glycerol, 0.1% NP40, 0.5 mM EDTA, 1 mM DTT). The eluate was further incubated with calmodulin sepharose for 1.5 h (calmodulin-binding buffer: 10 mM HEPES, 150 mM KCl, 1 mM Mg Acetate, 1 mM imidazole, 2 mM CaCl₂, 0.1% NP-40, 10 mM β -mercaptoethanol). Protein complexes were released from the column by chelation (calmodulin rinsing buffer: 50 mM ammonium bicarbonate, 75 mM NaCl, 1 mM Mg acetate, 1 mM imidazole, 2 mM CaCl₂, elution buffer: 50 mM ammonium bicarbonate, 25 mM EGTA). Peptides followed by trypsin digestion were dried down and enriched before loaded onto LC MS/MS. Peptides isolated from affinity purifications were separated by liquid chromatography and analyzed on ThermoFinnigan LTQ FT and searched by X!Comet against 2006 human IPI database.

Retroviral infection, siRNA-mediated expression knockdown, and RT-PCR

All retroviral expression plasmids were packaged in 293T cells with an amphotropic envelope. For cell lines used in TAP, RD cells were first infected by retroviruses and followed by puromycin selection for stable clones. Stable cell lines were switched to DM for 36 h before harvesting. For ChIP and forced differentiation assays, RD cells were infected consecutively for three times by retroviruses expressing either the control, E12, MyoD, MyoD~E12, or dnMSC and followed by 3–4 d of puromycin selection before switched to DM. Cells were treated with HiPerFect/siRNA mixture twice at a final concentration of 40 to 100 nM for a 24-h interval in growth medium (HiPerFect, Qiagen). All siRNAs were generated by Invitrogen Stealth RNAi. Three siRNAs against human MSC are available from Invitrogen. The ST1 siRNA specifically against E2A-2/5 and its randomly shuffled control siRNA, ST1rm, were designed by Invitrogen and the sequences were ST1 (5'-GCGCAGUUCG GAGGUUCAGGCAAGA-3') and ST1rm (5'-GCGUGGAGG CUACGUGACGACUAGA-3'). At 24 h after the second siRNA treatment, cells were switched to DM for 48 h. RNAs were harvested by Trizol method (Invitrogen).

CKM expression is quantified by primers hCKML1 (TCGAT GACCACTTCCTGTTC) and hCKMR1 (TTGTTCATTGTGCC AGATGC) and the probe hCKMP1 (CTGCTGCTGGCCTCAG GCAT). The splice variants of E2A are detected by 5NhE2Aatg (ATGAACCAGCCGCAGAGGATGGCGCC) and 3ChE2A-536 (GTGGGTACACCGAGGATGGAAGAC). The MSC transcript is detected by 5NbcLhMyoR (CGCTGATCAGCCGCCACCA TGTCCACGGGCTCGGTGAG) and 3CbamhMyoR (CGCGG ATCCAGCGGTGGTTCACATAGTCTG), hMSCF1 (GGAAG GCTGCAAGAGGAAG), and hMSCR1 (GCGGAAACTTCTT TGGTGTC) or hMSCR2 (TGGTCTTGAGCCTGGAGAAG). All real-time qPCRs used hTimm17b as the internal control, the primer set AA 94F (GGAGCCTTCACTATGGGTGT) and AA 95R (CACAGCATTGGCACTACCTC), and the probe TCAACCGGTGCCGAATCCA.

ChIP

ChIP was performed as described previously (Sawado et al. 2001) with 15 min formaldehyde fixation. Immunoprecipitations were done with both a rabbit polyclonal anti-MyoD (Tapscott et al.

1988) and a mouse monoclonal 5.8A (Santa Cruz Biotechnologies). The promoters were amplified with the following sets of primers: human Myh8, 5'-CATCAGGCACTGTCAAAGATTC-3' and 5'-AATGTGCTGATGCTGCAACTAC-3'; human CKM enhancer, 5'-CGGTTATAATTAACCAGGACACG-3' and 5'-CTGGAGATTTCCGCTGTG-3'. An intronic region of human Timm17a gene was used as a control for background binding: 5'-TTATGGATCTCCTCCTCATCCT-3' and 5'-AGTAATTCTATGCCCTGGGAGA-3'. The PCR products were separated on 3% low-melt and 1% standard agarose gel and visualized with SYBR Gold staining (Molecular Probes) and a Typhoon scanner (Molecular Dynamics). The intensity of DNA bands was quantified by ImageQuant.

Expression array analysis and target validation

Expression array profiling was done by using Affymetrix HG U133 plus 2.0 arrays and criteria of a twofold change and a FDR <0.05. The perfect-match probe intensities were corrected by robust multichip analysis, normalized by quantile normalization, and summarized by Tukey's median polish procedure using the Affy package of Bioconductor (Gentleman et al. 2004). The comparison of global gene expression profiles was done using the LIMMA package of Bioconductor. The data are accessible through the GEO series accession number GSE14825.

Primers used for array target validation were Dec1F (GACCGATTAACGAGTGCAT), Dec1R (GCTTGGCCAGATACTGAGC), Prom1F (CCTGGGGCTGCTGTTTATTA), Prom1R (TTACCTGGTGATTTGCCACA), MDH1 F (CAGGAAGACTGCTGTGTCCA), MDH1 R (ATGCAGATCTCCAGGCAGTC), GAPDH F (GAAGGTGAAGGTCGGAGTC), GAPDH R (GAAGATGGTGATGGGATTTTC), DesmF (GATCAATCTCCCCATCCAGA), DesmR (CTTCAGGGAGCAGTGAGGAC), Myh8F (AGGGTACGTGAGCTTGAAGG), Myh8R (GACCTGGGACTCAGCAATGT), Mef2C F (TGCATTGACATACACAGCAAAG), Mef2C R (ACAACCAAAGGTGCATTGTG), Myl1F (AGCCATTTCCAAACAAGG), Myl1R (TGTGCTTGACAAAAGCTTCG), Tnn1F (CACTCTGCTCCTGTCCGAGAA), Tnn1R (GGCAGATGGTGTTCACTTT), and IGFBP5F (GAAGAAACCCCATAGCGTCA), IGFBP5R (AGGAGGCCAGAGTCTGTGAA).

EdU or BrdU incorporation

RD cells infected by either the control virus or the virus expressing MyoD~E forced dimer were pulse-labeled by either EdU or BrdU. Flow cytometry assay was carried out as suggested by Invitrogen.

Acknowledgments

This work was supported by NIH-AR045113 (to S.J.T.) and F32CA117622 (to Z.Y.), and FHCRC/Amgen Interdisciplinary Training Grant (to Z.Y.). We thank Z. Yao, A.E. Tyler, L.D. Snider, S. Collins, I. Bernstein, L. Rott, and P. Gafken for experimental support and review of the text.

References

Azmi, S., Ozog, A., and Taneja, R. 2004. Sharp-1/DEC2 inhibits skeletal muscle differentiation through repression of myogenic transcription factors. *J. Biol. Chem.* **279**: 52643–52652.

Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. 1990. The protein Id: A negative regulator of helix-loop-helix DNA-binding protein. *Cell* **61**: 49–59.

Bergstrom, D.A., Penn, B.H., Strand, A., Perry, R.L.S., Rudnicki, M.A., and Tapscott, S.J. 2002. Promoter-specific regulation of MyoD binding and signal transduction cooperate to pattern gene expression. *Mol. Cell* **9**: 587–600.

Berkes, C.A. and Tapscott, S.J. 2005. MyoD and the transcriptional control of myogenesis. *Semin. Cell Dev. Biol.* **16**: 585–595.

Berkes, C.A., Bergstrom, D.A., Penn, B.H., Seaver, K.J., Knoepfler, P.S., and Tapscott, S.J. 2004. Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. *Mol. Cell* **14**: 465–477.

Bhalla, S., Spaulding, C., Brumbaugh, R.L., Zagort, D.E., Massari, M.E., Murre, C., and Kee, B.L. 2008. Differential roles for the E2A activation domains in B lymphocytes and macrophages. *J. Immunol.* **180**: 1694–1703.

Buas, M.F., Kabak, S., and Kadesch, T. 2009. Inhibition of myogenesis by Notch: Evidence for multiple pathways. *J. Cell. Physiol.* **218**: 84–93.

Chen, A.C., Kraut, N., Groudine, M., and Weintraub, H. 1996. I-mf, a novel myogenic repressor, interacts with members of the MyoD family. *Cell* **86**: 731–741.

Cobb, L.J., Salih, D.A., Gonzalez, I., Tripathi, G., Carter, E.J., Lovett, F., Holding, C., and Pell, J.M. 2004. Partitioning of IGFBP-5 actions in myogenesis: IGF-independent anti-apoptotic function. *J. Cell Sci.* **117**: 1737–1746.

Conway, K., Pin, C., Kiernan, J.A., and Merrifield, P. 2004. The E protein HEB is preferentially expressed in developing muscle. *Differentiation* **72**: 327–340.

Dilworth, F.J., Seaver, K.J., Fishburn, A.L., Htet, S.L., and Tapscott, S.J. 2004. In vitro transcription system delineates the distinct roles of the coactivators pCAF and p300 during MyoD/E47-dependent transactivation. *Proc. Natl. Acad. Sci.* **101**: 11593–11598.

Fiddler, T.A., Smith, L., Tapscott, S.J., and Thayer, M.J. 1996. Amplification of MDM2 inhibits MyoD-mediated myogenesis. *Mol. Cell Biol.* **16**: 5048–5057.

Flora, A., Garcia, J.J., Thaller, C., and Zoghbi, H.Y. 2007. The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. *Proc. Natl. Acad. Sci.* **104**: 15382–15387.

Fujimoto, K., Hamaguchi, H., Hashiba, T., Nakamura, T., Kawamoto, T., Sato, F., Noshiro, M., Bhawal, U.K., Suardita, K., and Kato, Y. 2007. Transcriptional repression by the basic helix-loop-helix protein Dec2: Multiple mechanisms through E-box elements. *Int. J. Mol. Med.* **19**: 925–932.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. 2004. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **5**: R80. doi: 10.1186/gb-2004-5-10-r80.

Gingras, A.-C., Aebersold, R., and Raught, B. 2005. Advances in protein complex analysis using mass spectrometry. *J. Physiol.* **563**: 11–21.

Graf Finckenstein, F., Shahbazian, V., Davicioni, E., Ren, Y.-X., and Anderson, M.J. 2008. PAX-FKHR functions as pangene by simultaneously inducing and inhibiting myogenesis. *Oncogene* **27**: 2004–2014.

Hahn, H., Wojnowski, L., Specht, K., Kappler, R., Calzada-Wack, J., Potter, D., Zimmer, A., Muller, U., Samson, E., Quintanilla-Martinez, L., et al. 2000. Patched target Igf2 is indispensable for the formation of medulloblastoma and rhabdomyosarcoma. *J. Biochem.* **27**: 28341–28344.

Inaba, T., Roberts, W.M., Shapiro, L.H., Jolly, K.W., Raimondi, S.C., Smith, S.D., and Look, A.T. 1992. Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia. *Science* **257**: 531–534.

King, A.M., Van der Put, E., Blomberg, B.B., and Riley, R.L. 2007. Accelerated Notch-dependent degradation of E47 proteins in aged B cell precursors is associated with increased ERK MAPK activation. *J. Immunol.* **178**: 3521–3529.

- Kophengnavong, T., Michnowicz, J.E., and Blackwell, T.K. 2000. Establishment of distinct MyoD, E2A, and twist DNA binding specificities by different basic region-DNA conformations. *Mol. Cell Biol.* **20**: 261–272.
- Kuroda, K., Tani, S., Tamura, K., Minoguchi, S., Kurooka, H., and Honjo, T. 1999. Delta-induced Notch signaling mediated by RBP-J inhibits MyoD expression and myogenesis. *J. Biol. Chem.* **274**: 7238–7244.
- Langenau, D.M., Keefe, M.D., Storer, N.Y., Guyon, J.R., Kutok, J.L., Le, X., Goessling, W., Neuberg, D.S., Kunkel, L.M., and Zon, L.I. 2007. Effects of RAS on the genesis of embryonal rhabdomyosarcoma. *Genes & Dev.* **21**: 1382–1395.
- Lu, Q., Wright, D.D., and Kamps, M.P. 1994. Fusion with E2A converts the Pbx1 homeodomain protein into a constitutive transcriptional activator in human leukemias carrying the t(1;19) translocation. *Mol. Cell Biol.* **14**: 3938–3948.
- Lu, J.-R., Webb, R., Richardson, J.A., and Olson, E.N. 1999. MyoR: A muscle-restricted basic helix–loop–helix transcriptional factor that antagonizes the actions of MyoD. *Proc. Natl. Acad. Sci.* **96**: 552–557.
- Lu, J.-R., Bassel-Duby, R., Hawkins, A., Chang, P., Valdez, R., Wu, H., Gan, L., Shelton, J.M., Richardson, J.A., and Olson, E.N. 2002. Control of facial muscle development by MyoR and capsulin. *Science* **298**: 2378–2381.
- Marampon, F., Ciccarelli, C., and Zani, B.M. 2006. Down-regulation of c-Myc following MEK/ERK inhibition halts the expression of malignant phenotype in rhabdomyosarcoma and in non muscle-derived human tumors. *Mol. Cancer* **5**: 31. doi: 10.1186/1476-4598-5-31.
- Markus, M., Du, Z.-M., and Benezra, R. 2002. Enhancer-specific modulation of E protein activity. *J. Biol. Chem.* **277**: 6469–6477.
- Martindill, D.M., Risebro, C.A., Smart, N., Franco-Viseras Mdel, M., Rosario, C.O., Swallow, C.J., Dennis, J.W., and Riley, P.R. 2007. Nucleolar release of Hand1 acts as a molecular switch to determine cell fate. *Nat. Cell Biol.* **9**: 131–141.
- Massari, M.E. and Murre, C. 2000. Helix–loop–helix proteins: Regulation of transcription in eukaryotic organisms. *Mol. Cell Biol.* **20**: 429–440.
- Massari, M.E., Grant, P.A., Pray-Grant, M.G., Berger, S.L., Workman, J.L., and Murre, C. 1999. A conserved motif present in a class of helix–loop–helix proteins activates transcription by direct recruitment of the SAGA complex. *Mol. Cell* **4**: 63–73.
- Mathas, S., Janz, M., Hummel, F., Hummel, M., Wollert-Wulf, B., Lusatis, S., Anagnostopoulos, I., Lietz, A., Sigvardsson, M., Jundt, F., et al. 2006. Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma. *Nat. Immunol.* **7**: 207–215.
- McAllister, R.M., Melnyk, J., Finkelstein, J.Z., Adams, Jr., E.C., and Gardner, M.B. 1969. Cultivation in vitro of cells derived from a human rhabdomyosarcoma. *Cancer* **24**: 520–526.
- Merlino, G. and Helman, L.J. 1999. Rhabdomyosarcoma—Working out the pathways. *Oncogene* **18**: 5340–5348.
- Mizrak, D., Brittan, M., and Alison, M.R. 2008. CD133: Molecule of the moment. *J. Pathol.* **214**: 3–9.
- Neuhold, L.A. and Wold, B. 1993. HLH forced dimers: Tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell* **74**: 1033–1042.
- Nofziger, D., Miyamoto, A., Lyons, K.M., and Weinmaster, G. 1999. Notch signaling imposes two distinct blocks in the differentiation of C2C12 myoblasts. *Development* **12**: 1689–1702.
- Nutt, S.L. and Kee, B.L. 2007. The transcriptional regulation of B cell lineage commitment. *Immunity* **26**: 715–725.
- Otten, A.D., Firpo, E.J., Gerber, A.N., Brody, L.L., Roberts, J.M., and Tapscott, S.J. 1997. Inactivation of MyoD-mediated expression of p21 in tumor cell lines. *Cell Growth Differ.* **8**: 1151–1160.
- Peinado, H., Olmeda, D., and Cano, A. 2007. Snail, ZEB and bHLH factors in tumor progression: An alliance against the epithelial phenotype? *Nat. Rev. Cancer* **7**: 415–428.
- Puri, P.L., Wu, Z., Zhang, P., Wood, L.D., Bhakta, K.S., Han, J., Feramisco, J.R., Karin, M., and Wang, J.Y. 2000. Induction of terminal differentiation by constitutive activation of p38 MAP kinase in human rhabdomyosarcoma cells. *Genes & Dev.* **14**: 574–584.
- Rigaut, G., Schevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**: 1030–1032.
- Rothschild, G., Zhao, X., Iavarone, A., and Lasorella, A. 2006. E proteins and Id2 converge on p57^{Kip2} to regulate cell cycle in neural cells. *Mol. Cell Biol.* **26**: 4351–4361.
- Salih, D.A., Tripathi, G., Holding, C., Szeszak, T.A., Gonzalez, M.I., Carter, E.J., Cobb, L.J., Eiseman, J.E., and Pell, J.M. 2004. Insulin-like growth factor-binding protein 5 (Igfbp5) compromises survival, growth, muscle development, and fertility in mice. *Proc. Natl. Acad. Sci.* **101**: 4314–4319.
- Sawado, T., Igarashi, K., and Groudine, M. 2001. Activation of β -major globin gene transcription is associated with recruitment of NF-E2 to the β -globin LCR and gene promoter. *Proc. Natl. Acad. Sci.* **98**: 10226–10231.
- Schwartz, R., Engel, I., Fallahi-Sichani, M., Petrie, H.T., and Murre, C. 2006. Gene expression patterns define novel roles for E47 in cell cycle progression, cytokine-mediated signaling, and T lineage development. *Proc. Natl. Acad. Sci.* **103**: 9976–9981.
- Sharp, R., Recio, J.A., Jhappan, C., Otsuka, T., Liu, S., Yu, Y., Liu, W., Anver, M., Navid, F., Helman, L.J., et al. 2002. Synergism between INK4a/ARF inactivation and aberrant HGF/SF signaling in rhabdomyosarcomagenesis. *Nat. Med.* **8**: 1276–1280.
- Snider, L., Thirlwell, H., Miller, J.R., Moon, R.T., Groudine, M., and Tapscott, S.J. 2001. Inhibition of Tcf3 binding by I-mfa domain proteins. *Mol. Cell Biol.* **21**: 1866–1873.
- St-Pierre, B., Flock, G., Zacksenhaus, E., and Egan, S.E. 2002. Stral3 homodimers repress transcription through class B E-box elements. *J. Biol. Chem.* **277**: 46544–46551.
- Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., et al. 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequence. *Proc. Natl. Acad. Sci.* **99**: 16899–16903.
- Sun, J., Kamei, C.N., Layne, M.D., Jain, M.K., Liao, J.K., Lee, M.E., and Chin, M.T. 2001. Regulation of myogenic terminal differentiation by the hairy-related transcription factor CHF2. *J. Biol. Chem.* **276**: 18591–18596.
- Tapscott, S.J., Davis, R.L., Thayer, M.J., Cheng, P.F., Weintraub, H., and Lassar, A.B. 1988. MyoD1: A nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* **242**: 405–411.
- Tapscott, S.J., Thayer, M.J., and Weintraub, H. 1993. Deficiency in rhabdomyosarcomas of a factor required for MyoD activity and myogenesis. *Science* **259**: 1450–1453.
- von Scheven, G., Bothe, I., Ahmed, M.U., Alvares, L.E., and Dietrich, S. 2006. Protein and genomic organization of vertebrate MyoR and Capsulin and their expression during avian development. *Gene Expr. Patterns* **6**: 383–393.
- Wang, H., Garzon, R., Sun, H., Ladner, K.J., Singh, R., Dahlman, J., Cheng, A., Hall, B.M., Qualman, S.J., Chandler, D.S., et al.

2008. NF- κ B-YY1-MiR-29 regulatory circuitry in skeletal myogenesis and rhabdomyosarcoma. *Cancer Cell* **14**: 369–381.
- Weintraub, H., Davis, R., Tapscott, S.J., Thayer, M., Krause, M., Benezra, R., Blackwell, T.K., Turner, D., Rupp, R., Hollenberg, S., et al. 1991. The myoD gene family: Nodal point during specification of the muscle cell lineage. *Science* **251**: 761–766.
- Zhang, J., Kalkum, M., Yamamura, S., Chait, B.T., and Roeder, R.G. 2004. E protein silencing by the leukemogenic AML1-ETO fusion protein. *Science* **305**: 1286–1289.
- Zhao, P. and Hoffman, E.P. 2006. Musculin isoforms and repression of MyoD in muscle regeneration. *Biochem. Biophys. Res. Commun.* **342**: 835–842.
- Zhuang, Y., Cheng, P., and Weintraub, H. 1996. B-lymphocyte development is regulated by the combined dosage of three basic helix–loop–helix genes, E2A, E2-2, and HEB. *Mol. Cell. Biol.* **16**: 2898–2905.