uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Vanessa Seale

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Cellular and Molecular Medicine)

GRADE / DEGREE

Department of Cellular and Molecular Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

ZAC1, A ZN Finger Transcription Factor, is a Direct PAX7 Target Gene in Post-Natal Myogenesis

TITRE DE LA THÈSE / TITLE OF THESIS

M. Rudnicki

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

A. Blais

B. McKay

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

ZAC1, A ZN FINGER TRANSCRIPTION FACTOR, IS A DIRECT PAX7 TARGET GENE IN POST-NATAL MYOGENESIS

By

VANESSA SEALE, HONS. B.SC.

This thesis is submitted as a partial fulfillment of the M.Sc. program

in

Cellular Molecular Medicine.

Sept 5th, 2008

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-51857-1
Our file Notre référence
ISBN: 978-0-494-51857-1

Abstract

Pax7 plays an essential role in the commitment, survival and expansion of myogenic progenitors during postnatal muscle growth and regeneration. Recently, our genome-wide expression study in myoblasts identified *Myf5* and *Zac1* (*PlagL1;* encoding a zinc-finger transcription factor) as transcripts regulated by Pax7. *Zac1* has been previously identified as a tumour suppressor showing anti-proliferative properties through regulation of apoptosis and cell cycle arrest. *Zac1* expression paralleled that of *Pax7* during myogenic progression, with both genes being expressed at high levels in proliferating primary myoblasts but being rapidly down-regulated during cell-cycle arrest and terminal differentiation. Genome wide protein-DNA interaction profiling and chromatin immunoprecipitation assays have identified a region bound by Pax7 at +11 kb from the *Zac1* transcriptional start. This region contains a novel Pax7 consensus sequence that is conserved across species and is functionally activated by the Pax7/FKHR fusion protein. Together these data support the assertion that *Zac1* is directly activated by Pax7. Regeneration assays conducted in *Zac1*- null mice indicate a critical role for Zac1 in mediating efficient regenerative myogenesis. Further elucidation of the functional regulation of *Zac1* by *Pax7* will provide important insight into the mechanism by which Pax7 defines the cellular phenotype of myogenic progenitors.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ATCC	American type cell culture
bHLH	Basic helix-loop-helix
BMP	Bone morphogenic protein
BrdU	Bromodeoxyuridine
ChIP	Chromatin immunoprecipitation
CTX	Cardiotoxin
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
EMSA	Electrophoretic mobility shift assay
FBS	Fetal bovine serum
FGF	fibroblast growth factor
FKHR	FoxO1
FORSE-1	forebrain surface embryonic
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GFAP	glial fibrillary acidic protein
HAT	histone acetyl transferase
H3K4	histone 3 lysine 4
HGF	hepatocyte growth factor
HMT	histone methyl transferase
HRP	horseradish peroxidase
HTH	helix-turn-helix
ICR	imprint control region

Igf_2	Insulin Growth Factor 2
LCR	locus control region
MAPK	mitogen-activated protein kinase
MLL	mixed lineage leukemia
MNF	monocyte nuclear factor
MRF	myogenic regulatory factor
NH_2-	amino-
PACAP	pituitary adenylate cyclase-activating polypeptide
PBS	phosphate-buffered saline
PBST	phosphate-buffered saline tween
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PLE/PMQ	proline-leucine-glutamate/proline-methionine-glutamine
Rb	retinoblastoma
RNA	ribonucleic acid
RT	reverse transcription
SDS	sodium dodecyl sulfate
TA	tibialis anterior
TEMED	tetramethylethylenediamine
TGF	transforming growth factor
TK	thymidine kinase

Acknowledgements

I would like to thank Michael for providing me the opportunity to complete my Master's degree in his laboratory. Throughout the years, the support and encouragement I have received from Michael have been essential in the development of my enthusiasm to contribute to scientific research. The skills and knowledge I have gained in Rudnicki Lab will be carried throughout all my future pursuits. I would like to thank all the lab members for all their guidance and technical support throughout my research. In particular, I would like to thank Dr. Jeff Ishibashi who has spent many hours both at the bench and over many scientific discussions guiding me through all aspects of graduate studies. I also wish to acknowledge the members of my advisory committee, Dr. Bruce McKay and Dr. Marjorie Brand , who provided me with much appreciated input and advice. Dr Jenn McCann was instrumental in the revision of my thesis.

I owe this Master's thesis to many people outside of the lab. My brother, Patrick, inspired me to stay focused. The support from my parents has been key in allowing me to see my studies to completion. Finally, I would like to express my appreciation for my husband, Jeff Murphy who endorsed my decision to return to school after many years and made it possible through his support throughout my studies.

CHAPTER 1

Introduction

1.1 Overview

Skeletal muscle tissue is essential for life as it allows for the creation of movement to support breath, posture and locomotion. In addition to understanding the basic structural and molecular mechanisms that underline its basic function, skeletal muscle is an excellent experimental model of regenerative biology due to its inherent ability to regenerate new fibres through activation of a unique population of cells that posses stem cell behaviour (Bischoff, 1994; Kuang et al., 2007; Le Grand and Rudnicki, 2007). Characterized by their positional identity, satellite cells arise late in development and reside, in a quiescent state, adjacent to mature muscle fibres in a region defined as the stem cell niche (Beauchamp et al., 1999; Kuang et al., 2007; Mauro, 1961; Seale et al., 2000). On a molecular level, satellite cells are defined by their expression of Pax7, a paired box transcription factor, which is uniquely required for the commitment, specification and expansion of myogenic progenitor cells (Seale et al., 2000). In response to injury, these cells are activated and eventually undergo terminal differentiation through sequential expression of regulatory factors that together, make up the myogenic program during post-natal muscle regeneration (Charge and Rudnicki, 2004). While the developmental origins of these cells have been established, the exact molecular mechanisms that mediate their regenerative potential remain unclear.

1.2 Adult Skeletal Muscle

Skeletal muscle is a complex organization of muscle, connective, nerve and vascular tissue that together allow movement through contraction. At the anatomical level, each muscle is composed of syncytial muscle fibres that are organized into bundles surrounded by a perimysium layer. These bundles are grouped together by the myomysium, forming a single muscle. Muscle fibres can also be organized with respect to their innervation as each fibre comes in contact with an axon terminus originating from a single motor neuron. Signals activating muscle groups are then transmitted through the central nervous system, inducing contraction of single fibres through motor end plates. Blood vessels surrounding the fibres provide oxygen and nutrients essential for muscle function (Charge and Rudnicki, 2004).

Each muscle fibre is a highly specified, distinctive cell type, characterized by an elongated, striated and multi-nucleated morphology distinguishing it from smooth muscle. The distinctive physical and molecular nature of each fibre allow it to contract following a sliding filament mechanism where organized actin and myosin filaments interact to shorten the fibre (Huxley, 2002). This myofibre is the basic cellular contractile unit of muscle and is essential for life allowing for locomotion, maintenance of posture and breathing.

While adult muscle is normally a stable and post-mitotic tissue with minimal nuclear turn-over (Decary et al., 1997; Schmalbruch and Lewis, 2000), it also possesses the capacity to regenerate and repair damaged fibres in response to injury and trauma to maintain its structure and function (Bischoff, 1994). This ability to regenerate new muscle tissue in the adult is mediated by a

unique population of cells known as satellite cells that arise late in development and take a position beneath the basal lamina of mature fibres (Gros et al., 2005; Kassar-Duchossoy et al., 2005; Schienda et al., 2006). During regeneration, these cells enter into the tightly regulated myogenic pathway, restoring muscle architecture (Parker et al., 2003). While post-natal skeletal myogenesis generally recapitulates embryonic myogenesis, these progenitors are distinct from the embryonic myogenic progenitors that are responsible for muscle formation in the developing embryo (Kuang et al., 2006; Relaix et al., 2006; Ustanina et al., 2007).

1.3 Skeletal Muscle Regeneration

During embryogenesis and early post-natal growth, waves of proliferation and subsequent differentiation of myogenic precursor cells contribute to and build the muscle architecture (Buckingham, 2001). In contrast, during adulthood, muscle tissue remains in a stable state with any minor tears in the plasma membrane being repaired directly (Decary et al., 1997; Schmalbruch and Lewis, 2000). In response to substantial trauma and injury however, this resting state is quickly replaced by a dynamic and highly efficient regenerative process that is marked by a huge immune cell invasion, cellular proliferation, and generation of newly formed multinucleated muscle fibres (Charge and Rudnicki, 2004). The stages of the repair process have been well characterized in muscle damaged by either mechanical or chemical means, such as freeze crush injuries or cardiotoxin (CTX) respectively.

In general, the repair process can be divided into two phases: the degenerative phase, marked by necrosis of the damaged fibres, and the regenerative phase, whereby fibres are repaired (reviewed in Grounds, 1998). Following disruption of the myofibre sarcolemma and induction of necrosis, there is a rapid infiltration of neutrophils at 6 hours post-injury, followed by macrophage mononuclear inflammatory cells at 48 hours post-injury (Fielding et al., 1993; Orimo et al., 1991; Tidball, 2005). Recruitment of these inflammatory cells is thought to occur through chemotactic signals, although systemic response to injury has also been documented (Tidball, 2005). These cells, known to be involved in removing cellular debris resulting from necrosis, may also play a direct role activating the myogenic programme, although this function remains to be determined (Charge and Rudnicki, 2004). The ensuing regenerative phase is characterized by a large influx of myogenic precursor cells known as myoblasts. These cells are derived from the activation and subsequent proliferation of satellite cells, providing the necessary resource of myonuclei required for the generation of new fibres and the repair of existing ones (Bischoff, 1994; Grounds and Yablonka-Reuveni, 1993; McGeachie, 1985; McGeachie and Grounds, 1989; Schultz, 1989; Schultz and Jaryszak, 1985; Schultz and Lipton, 1982). These cells undergo terminal differentiation and eventually fuse (Bischoff, 1994). Newly formed fibres represent a determined myogenic cell population based on their expression of muscle specific transcription factors and structural proteins (Seale and Rudnicki, 2000).

1.4 Developmental Origins of Satellite Cells

Embryonic myogenesis has been well characterized, illustrating the basic structure of skeletal muscle. A complex series of signaling molecules and transcription factor activation define the myogenic cell lineage during embryogenesis leading to the patterning and development of skeletal muscle (Parker et al., 2003). During the early stages of embryogenesis, paraxial mesoderm, which lies adjacent to the neural tube, becomes segmented along a rostral-to-caudal axis to form somites which eventually give rise to myogenic progenitors (Tajbakhsh et al., 1997). The somites undergo a cellular transition to form the mesenchymal sclerotome and the dermomyotome under the influence of Shh, Wnt and BMP signals originating from the notochord, neural tube, surface ectoderm and lateral plate mesoderm (Cossu et al., 1996; Tajbakhsh et al., 1997). Myogenic progenitors located in the epaxial portions of the epithelial dermomyotome of mature somites will delaminate and differentiate to form the myotome immediately under the epithelium (Ben-Yair and Kalcheim, 2005; Gros et al., 2004). In time, these cells will eventually form the musculature of the trunk. Alternatively, hypaxial dermamyotomal cells will migrate to distant sites forming the skeletal musculature of the limbs (Christ and Ordahl, 1995). Myotome formation is dependant on activation of expression of *Myf5* and *MRF4* via signaling from the surrounding tissues and Pax3. Limb bud formation is dependant solely on Pax3 expression which acts directly upstream of *Myf5* to activate its expression (Kassar-Duchossoy et al., 2004; Tajbakhsh et al., 1997). The central region of the dermomyotome expresses *Pax3* and *Pax7* (Buckingham et al., 2003). At E10.5, the epithelial layer disintegrates and these

cells enter the myotome where they proliferate and contribute to skeletal muscle growth (Gros et al., 2005; Kassar-Duchossoy et al., 2005; Relaix et al., 2005). This population of cells gives rise to the satellite cells which begin to take their position beneath the basal lamina of differentiated muscle fibres at E17.5 (Bischoff, 1994; De Angelis et al., 1999). It is this small population of cells that are responsible for both the post-natal growth as well as the regenerative capacity in adult muscle.

1.5 Satellite Cells

The ability to both repair and form new fibres in response to injury is mediated by satellite cells residing in a niche between the basal lamina and the sarcolemma of the mature myofibre (Bischoff, 1994; Grounds and Yablonka-Reuveni, 1993; Seale and Rudnicki, 2000; Seale et al., 2000). These cells, which represent committed myogenic progenitor cells, were originally defined by their location and morphological features through electron microscopy (Mauro, 1961; Seale and Rudnicki, 2000). Their characteristically low cytoplasmic content and small, heterochromatic nuclei reflect their mitotically quiescent state (Armand et al., 1983; Schultz, 1976). Satellite cells represent a total of 30 percent of the sublaminar nuclei along the fibre, but are progressively consumed through post-natal development and represent only 3 to 5 percent of the sublaminar nuclei in adult muscle (Bischoff, 1994).

Upon signals induced by trauma, injury or stretching, these cells are activated, expand in number and undergo terminal differentiation under the

control of myogenic regulatory factors (MRFs) to either repair existing damaged fibres or form new ones (Fig.1). A portion of these cells do not differentiate and instead replenish the satellite cell pool (Baroffio et al., 1996; Moss and Leblond, 1971; Schultz, 1996; Yoshida et al., 1998).

Whereas the morphology of satellite cells has been clearly established, their cellular phenotype remains to be fully characterized. While many studies have demonstrated the presence of certain molecular markers that help to isolate, define and understand the satellite cell, the most defining genetic characteristic of the satellite cell is its expression of the paired box transcription factor, Pax7 (Seale et al., 2000). Discovery of Pax7 expression through representational difference analysis on cDNA derived from cultured myogenic precursor cells by the Rudnicki group was vital in the advancement of understanding satellite cell biology. *Pax7* is uniquely expressed by all satellite cells and is required for their specification during embryogenesis and their expansion throughout post-natal growth and regeneration (Kuang et al., 2006; Seale et al., 2000). *Pax3,* the paralogue of *Pax7* is also expressed in a subset of both quiescent and activated cells (Buckingham et al., 2003), though its function in adult myogenesis is unknown.

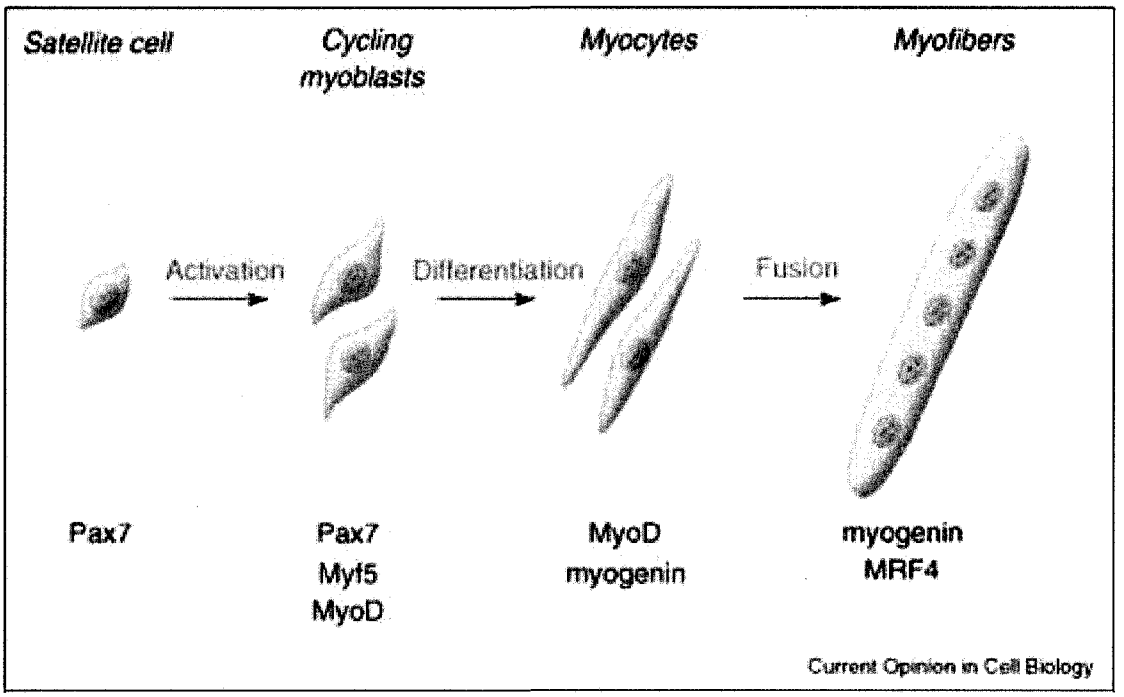


Figure 1. Post-natal myogenesis is mediated by satellite cells. Schematic diagram of satellite cell commitment and terminal differentiation adapted from Legrand, 2007. Satellite cells are activated in response to stress, proliferate and undergo terminal differentiation. Shown is expression of proteins. The myogenic program is controlled by expression of Pax7 and the myogenic regulatory factors (Myf5, MyoD, Myogenin and MRF4)

Expression profiling of both quiescent and activated satellite cells have described a series of molecular markers that are expressed in the quiescent state, the active state or both. Satellite cells express the cell surface markers, Syndecan-3 and Syndecan-4 as well the adhesion molecules, M-Cadherin, VCAM and NCAM (Charge and Rudnicki, 2004). CD34, which is expressed in various cell types including hematopoetic and endothelial progenitors, is also expressed on the cell surface of a subset of satellite cells (Beauchamp et al., 2000; Charge and Rudnicki, 2004; Hawke and Garry, 2001). Integrin-α7 is a muscle-specific laminin receptor involved in signaling and expressed in both satellite cells and committed myogenic precursor cells (Blanco-Bose et al., 2001).

Together, these molecular markers have enabled researchers to identify and isolate subsets of satellite cells through immunostaining and flow cytometry for *in vitro* and transplantation studies (Zammit et al., 2006a). As committed myogenic progenitors, these cells have also been defined on the basis of their expression of myogenic transcription factors. The winged-helix transcription factor, Myocyte Nuclear Factor (MNF), is expressed in the quiescent satellite cell, while an alternative isoform is reciprocally expressed in the activated state (Garry et al., 1997). Importantly, the MRF Myf5 is expressed at low levels in a subset of quiescent cells, whereas MyoD is expressed exclusively in activated satellite cells (Beauchamp et al., 2000; Cornelison and Wold, 1997; Fuchtbauer and Westphal, 1992; Grounds et al., 1992).

The molecular profiles of satellite cells highlight a fundamental feature critical to their function: their heterogeneity (Beauchamp et al., 2000). The

satellite cell population is composed of both committed progenitors and self-renewing stem cells (Collins et al., 2005; Kuang et al., 2008). Heterogeneity was initially described on the basis of their cycling kinetics where 20 percent of cells had a significantly reduced cycling rate (Schultz, 1996). Upon further analysis, satellite cells displayed characteristics of a mixed population of cells when transplantations of cultured myogenic precursors into injured muscle were much less efficient at contributing to muscle regeneration than those conducted with whole fibres containing the intact satellite cell niche (Collins et al., 2005; Montarras et al., 2005; Rando and Blau, 1994). The initial characterization of the satellite cell genetic profile has revealed that while all satellite cells express Pax7, there is also a subset of cells defined by their differential expression of Myf5 (Kuang et al., 2008).

Lineage tracing experiments using a *Myf5* reporter mouse strain have demonstrated that the existence of a small population of satellite cells that have never expressed Myf5 can give rise to both Myf5+ and Myf5- cells through asymmetric divisions occurring in the satellite cells niche (Kuang et al., 2008; Kuang et al., 2007). This finding establishes that satellite cells are a heterogeneous population, including cells that display stem cell properties of self-renewal and asymmetric division.

1.6 Activation of Satellite Cells

The exact mechanism through which satellite cells become activated is not clearly understood. Intrinsic, extrinsic and microenvironmental factors have

been described as playing roles in the transition of cells from a quiescent to an active state. Sphingosine-1-phosphate production, on the inner portion of the plasma membrane from within the cell itself, marks the activation event (Nagata et al., 2006). This signal is necessary for entry into the cell cycle and without it, regeneration is severely impaired (Nagata et al., 2006). Various studies have also demonstrated that hepatocyte growth factor (HGF) plays a similar role in the activation of quiescent satellite cells. HGF is rapidly upregulated in damaged muscle in proportion to the level of tissue damage and increases satellite cell proliferation when injected directly into regenerating muscle (Cornelison and Wold, 1997; Miller et al., 2000; Tatsumi et al., 1998). HGF acts through the c-met receptor which is widely expressed in both quiescent and activated satellite cells (Anastasi et al., 1997). HGF release into the extracellular matrix occurs through the production of both nitric oxide synthase and subsequent nitric oxide following trauma to the myofibril basement membrane. Nitric oxide is believed to initiate the release of HGF bound to heparin sulphate proteoglycans (Wozniak and Anderson, 2007). Interestingly, evidence suggests that nitric oxide also stimulates follistatin release to negatively regulate myostatin, a transforming growth factor-β (TGF-β) family protein that blocks proliferation of satellite cells (Le Grand and Rudnicki, 2007).

Myostatin is thought to function by inhibiting proliferation by blocking cell cycle entry via cyclin-Cdk2 inactivation (Thomas et al., 2000). In addition, fibroblast growth factor (FGF) is released from both non-myogenic and myogenic cells in the microenvironment during the regenerative process, activating satellite

cell proliferation and inhibiting differentiation. It has been suggested that FGF acts through the mitogen-activated protein kinase (MAPK) signaling cascade for satellite cell activation (Campbell et al., 1995). While the exact mechanism of satellite cell activation remains to be elucidated, it is clear that mitogenic signals from growth factors are essential in removing them from their quiescent state and inducing proliferation.

1.7 The Role of Myogenic Regulatory Factors

Myogenic cell expansion and terminal differentiation is orchestrated by the sequential expression of molecular cues that are both finely tuned and elegantly effective for generating myofibres. In the adult, Pax7 is expressed in both quiescent and activated satellite cells (Zammit et al., 2006a). Pax7 expression is dramatically down-regulated during terminal differentiation as the myoblasts exit the cell cycle and fuse to form multi-nucleated myoblasts (Olguin et al., 2007). It has long been established that a family of four basic-loop-helix (bHLH) molecules function as master regulatory switches directing post-natal myogenesis (Parker et al., 2003).

This family is comprised of MyoD, Myf5, Myogenin and MRF4 together known as the myogenic regulatory factors (MRFs). These factors all share a highly conserved 65 amino acid basic helix-loop-helix bHLH domain which allows for the dimerization with E-proteins. Dimerization is required for binding to the E-box sites of the consensus sequence 'CAnnTG' and subsequent activation of muscle specific gene transcription (Puri and Sartorelli, 2000).

MyoD and Myf5 are co-expressed early on during post-natal myogenesis and are therefore considered the primary myogenic factors. In the developing embryo, these factors are able to effectively compensate for the loss of the other whereas in regenerating post-natal muscle, these factors have uniquely specified roles (Braun et al., 1992; Rudnicki et al., 1992). MyoD upregulation is observed immediately following satellite cell activation and it is required for differentiation of myoblasts (Cornelison and Wold, 1997; Megeney et al., 1996; Sabourin et al., 1999; Yablonka-Reuveni and Rivera, 1994). The exact role of Myf5 during regenerative myogenesis remains undefined; however, it has been shown that Myf5 has a distinct function from that of MyoD. Previous studies have suggested that Myf5 may be involved in regulating proliferation of myoblasts as *Myf5*-null myoblasts have deficient cell proliferation (Montarras et al., 2005). MRF4 and myogenin are expressed during the later portion of regenerative myogenesis, with myogenin marking the onset of myogenic terminal differentiation (Berkes and Tapscott, 2005; Smith et al., 1994). While the exact role of MRF4 remains unclear, it is not expressed in satellite cells or myoblasts and as a result, does not determine their cellular fate (Gayraud-Morel et al., 2007).

1.8 Pax3 and Pax7 have Unique Functions in Myogenesis

Mechanisms involved in muscle formation and development in the embryo are also exploited to build new muscle in the adult; however, there are important differences in the genetic program that define these two processes. Skeletal muscle patterning and development in the embryo and in adult regenerative

myogenesis,is dependant on expression of Myf5, MRF4 and MyoD, which together determine the muscle cell lineage (Kassar-Duchossoy et al., 2004). Recent evidence suggests that Myf5 plays a specific regulatory role in adult myogenesis by modulating the proliferation of myogenic precursor cells following injury muscle (Ustanina et al., 2007). The transcription factors Pax3 and Pax7 also have distinct roles in embryonic versus post-natal myogenesis (Kuang et al., 2006).

Whereas Pax3 expression ensures the survival and migration of myogenic progenitor cells during embryonic muscle formation, Pax7 expression is uniquely required for the survival, expansion and specification of the adult muscle progenitors. This has been demonstrated *in vivo* through mouse knockout studies. *Pax3* mutant mice have gross defects in muscle formation and do not survive past birth (Bober et al., 1994; Goulding et al., 1994; Tajbakhsh et al., 1997). In contrast, *Pax7* mutant mice, while severely runted at birth, survive up to three weeks of age, reflective of a functional defect in skeletal muscle growth immediately following birth. Electron microscopy analysis revealed that $Pax7^{-/-}$ post-natal skeletal muscle was completely deficient in satellite cells (Seale et al., 2000). Furthermore, primary cells isolated from this muscle were unable to give rise to the proliferating myoblasts as was seen with their wildtype litter mates (Seale et al., 2000). Histological examination revealed a weaker muscle phenotype as a result of fibre size reduction (Kuang et al., 2006). Finally, $Pax7^{-/-}$ mice had extremely poor regenerative capacity following muscle injury due to a functional loss of satellite cells (Kuang et al., 2006).

The distinct roles of Pax3 and Pax7 are further outlined in experiments where *Pax3* and *Pax7* are knocked into their reciprocal loci. When *Pax3* is replaced by *Pax7* through classical gene targeting techniques, overall muscle patterning in the embryo is retained; however, distant limb formation is compromised and has been attributed to migration defects resulting from misexpression of Pax3 (Relaix et al., 2004). Conversely, when *Pax3* is knocked into *Pax7*, the resulting phenotype mirrors that seen in *Pax7* mutant mice (Keller et al., 2004) . Taken together, these studies suggest that Pax3 may play a role in the migration of myogenic precursors whereas Pax7 has a unique function in ensuring the survival of progenitors throughout post-natal muscle development.

Pax3 expression is down-regulated in the *Pax3/7* progenitors immediately before birth. A subset of satellite cells do express Pax3, however the function of Pax3 expression in adult myogenesis remains to be clearly defined (Kassar-Duchossoy et al., 2005; Montarras et al., 2005; Relaix et al., 2005). Recent studies have identified a subset of Pax3 expressing cells that co-express MyoD in the interstitial environment of skeletal muscle which have the capacity to contribute, although minimally, to adult muscle regeneration (Kuang et al., 2006). As these cells do not express any satellite cell markers, they are not believed to be from satellite cells. Together, these data confirm the distinct roles played by Pax3 and Pax7 in myogenesis.

1.9 Comparative Analysis of Pax-family Gene and Protein Structure

Pax7 and Pax3 belong to a family of nine paired-box transcription factors which are grouped together based on their sequence similarity and their genomic organization. Pax family members all contain the highly conserved paired-box DNA binding domain that includes two helix-turn-helix (HTH) motifs (Stuart et al., 1994). An octapeptide region and a whole or partial homeodomain (which also contains an HTH motif) are also found in subsets of the family (Epstein et al., 1994; Goulding et al., 1991).

Originally identified in *Drosophila*, paired-box genes have been conserved across species throughout evolution suggesting a fundamental role in development (Tremblay and Gruss, 1994). Indeed, *Pax* genes have been shown to orchestrate organogenesis through the activation of a cascade of developmental genes. For example, Pax6 directs eye development in the embryo by acting as a master regulator for gene activation (Ziman et al., 2001). Pax7 expression has been observed in the developing nervous system, particularly in neural crest cells, and is thought to play a role both in neurogenesis and in specifying post-natal myogenic progenitors (Chi, 2002).

Pax proteins are also expressed in complex and overlapping patterns during nervous system development raising the possibility that Pax family members may also exhibit some level of functional redundancy (Relaix et al., 2004). *Pax3* is the most closely related to *Pax7* and these two genes are believed to have arisen from gene duplication (Mansouri et al., 1996). Both genes share similar extended N-terminal sequences, exon-intron boundaries, paired domains and complete homeodomains; however, their C-terminal

sequences diverge (Relaix et al., 2004). Both *Pax3* and *Pax7* can be alternatively spliced at exon 3 in the paired-box domain to include or exclude a glutamine residue (Vogan et al., 1996). *Pax7* has an additional splicing variation in the paired domain at exon 4 where a glycine-leucine dipeptide can be inserted or removed. The consequences of these splice variants are currently under investigation.

The structural similarities of *Pax3* and *Pax7* have led to the hypothesis that they have overlapping roles in the activation of target genes. Both Pax3 and Pax7 are essential in myoblast cell survival. Myoblasts undergo apoptosis in *Pax3*-mutant mice, a phenotype which is rescued by reintroduction of Pax3 expression (Borycki, 1999). Similarly, loss of satellite cells in *Pax7*-mutant mice has been associated with induction of caspase-3 and apoptosis (Relaix et al., 2006). *Pax7* has been demonstrated to compensate for the loss of Pax3 activity in the neural tube, neural crest and somite development; however, *Pax7* does not rescue the defective appendicular muscle formation, suggestive of divergent roles (Relaix et al., 2004). Furthermore, their distinct roles in myogenesis have also been illustrated in mice whose mutations for *Pax3* and *Pax7* result in very different phenotypes (Bober et al., 1994; Goulding et al., 1994; Seale et al., 2000).

While *Pax7* is uniquely required for the specification, expansion and survival of myogenic progenitors in adult muscle, its molecular mechanism of target gene activation remains to be elucidated. As Pax7, like Pax3, is an inherently weak trans-activator due to *cis*-repression (Bennicelli and Barr, 1999),

the identification of downstream target genes has been difficult. Translocation-fusions of either the Pax7 or Pax3 NH_2-terminal DNA-binding domains to the C-terminus of the winged helix transcription factor FoxO1 (FKHR) result in a PAX-FKHR fusion protein that can strongly activate gene transcription relative to wildtype Pax3 and Pax7 gene activation. Expression of these chimeras is associated with the formation of alveolar rhabdomyosarcomas (Keller et al., 2004). Importantly, these soft tissue tumors express various myogenic markers, including the MRFs providing some insight into Pax7 target gene activation (Bennicelli and Barr, 1999).

Pax DNA binding sites were originally described through the use of an *in vitro* PCR-based binding site selection method using purified Pax3 paired domain fusion proteins (Chalepakis and Gruss, 1995). Such a site is found in the promoter of the c-met receptor. In that study, the Pax3 recognition sequence was found to share similarities and differences with other identified paired box proteins. This implies that Pax proteins posses a dual capacity to activate the same target genes as well as other unique target genes (Chalepakis and Gruss, 1995). More recently, this Pax3 consensus sequence was identified in a distal *Myf5* enhancer located 57 kb upstream of its transcriptional start site and was found to be bound and activated by Pax3 in the context of embryonic hypaxial somite specification (Bajard et al., 2006).

Identifying Pax7 target genes has important implications in advancing the understanding of Pax7 function: it provides a way to identify Pax7 binding sites in the hopes of understanding its molecular nature in activating downstream

targets, and it defines the phenotype of satellite cells gaining insight into the exact role played by Pax7. Recently, a comparative microarray analysis in C2C12 cells ectopically expressing Pax7 identified a set of candidate Pax7 target genes (McKinnell et al., 2008). Interestingly *Myf5 was* found to be induced by Pax7 expression in both C2C12 and primary myoblasts. Chromatin immunoprecipitation (ChIP) studies have revealed that Pax7 also interacts with the *Myf5* -57.5 kb region in C2C12 myoblasts suggesting that Pax3 and Pax7 share similar binding sites reflective of their overlapping roles during myogenesis (McKinnell et al., 2008).

1.10 Epigenetics and Regulation of Myogenesis

Epigenetic modifications have emerged as key players in regulating developmental programs of gene expression in the embryo (Kiefer, 2007). Nucleosomes, which are the basic unit of chromatin, are composed of histone octamers wrapped tightly by DNA to form highly structured and compact chromatin that, in its closed state, is inaccessible to proteins involved in gene transcription (Margueron et al., 2005). Coordinated epigenetic modifications of histones are known to play critical roles in maintaining a balance between open and closed chromatin conformations (Park and Pfeifer, 2003). These epigenetic modifications, which include methylation and acetylation of histone amino acid residues, are mediated by large complexes classified as histone-modifying enzymes (McKinsey et al., 2002). The interplay between histone modifications and chromatin structure is complex and classified into a 'histone code' whereby

certain epigenetic marks are associated with active or repressive gene transcription (Margueron et al., 2005). Through genome-wide histone modification mapping, methylation of H3 on lysine residue 4 (H3K4) has been linked to transcriptionally active loci. Notably, high levels of H3K4 trimethylation were observed in 5' regions of transcriptionally active genes, along with high levels of active RNA polymerase II occupancy (Ruthenburg et al., 2007).

In the embryo, the methyl transferase trithorax group proteins (Trx) have been shown to maintain chromatin structure in an open state through methylation of H3K4 to direct developmental program of gene expression (Ruthenburg et al., 2007). This mode of gene regulation has been observed in Drosophila *Hox* genes, where methylated H3K4 clusters have been documented (Ringrose and Paro, 2004; Ruthenburg et al., 2007). Trx proteins contain a SET domain responsible for catalyzing the methyl group transfer, and include the MLL1 (mixed lineage leukemia 1) family of proteins. MLL1 proteins interact with complexes that are organized into conserved subunits comprised of Wdr5, RbBP5 and Ash2L (Ringrose and Paro, 2004; Ruthenburg et al., 2007). Whereas Wdr5 is thought to be involved in recognizing and binding to the H3 amino tail (Iberg et al., 2008), the three subunits together are thought to provide a structural platform that can recruit MLL1 proteins. In fact, studies have demonstrated that the core complex of the three subunits associated with an MLL1 protein is sufficient to trimethylate H3K4 (Dou et al., 2006; Milne et al., 2005; Steward et al., 2006).

Due to its weak transactivation properties, it has been hypothesized that Pax7 recruits proteins to form a complex that allows for transcriptional activation of its target genes (McKinnell et al., 2008). Pax7-interacting proteins were purified using a tandem affinity purification techniques with epitope-tagged Pax7-His-Flag expressed in C2C12 myoblasts. Interestingly, analysis by mass spectrometry technology revealed the presence of Wdr5, a protein known to interact with Mll complexes. This interaction is suggestive of an association of Pax7 with a histone methyltransferase (HMT) complex (McKinnell et al., 2008). Furthermore, reciprocal co-immunoprecipitation analysis revealed an interaction between Pax7 and other core components of the HMT complex supporting the notion that Pax7 activates gene transcription through epigenetic modifications (McKinnell et al., 2008). This hypothesis parallels the previous findings that epigenetic modifications are key determinants in directing cell fate.

Identification of putative Pax7 target genes has allowed us to gain insight into the function of Pax7 expression during post-natal myogenesis. A direct link between Pax7 expression and activation of *Myf5* through microarray analysis has been recently confirmed through the mapping of histone modifications on the *Myf5* locus following induction by Pax7 (McKinnell et al., 2008). The finding that Pax7 acts directly upstream of *Myf5* establishes a novel genetic hierarchy during post-natal myogenesis and underlines the importance of epigenetic modifications regulating the post-natal myogenic programme.

In addition to *Myf5*, Pax7 strongly activates a set of genes with varying functions. Importantly, Pax3 induction did not activate expression of these genes,

supporting the finding that Pax7 and Pax3 have both distinct roles in adult regenerative myogenesis. In particular, *Zac1 (PlagL1)* was strongly activated following induction by Pax7 in both C2C12 myoblasts and 10T1/2 fibroblasts (Fig. 2, J. Ishibashi, unpublished). These data suggest that *Zac1* could potentially be a direct transcriptional target of Pax7.

Putative Pax7 binding regions have been identified -42 kb and +11 kb from the *Zac1* locus using a genome-wide binding site mapping approach (Punch, unpublished). Large scale immunoprecipitation assays combined with high-throughput sequencing on immunoenriched DNA were conducted on primary myoblasts ectopically expressing Pax7. This technique (ChIPseq) is a powerful tool used to elucidate *in vivo* protein-DNA interactions on a global scale and allows the identification of both canonical and non-canonical binding motifs (Johnson et al., 2007). Taken together, these data suggest that *Zac1* is a direct Pax7 target gene.

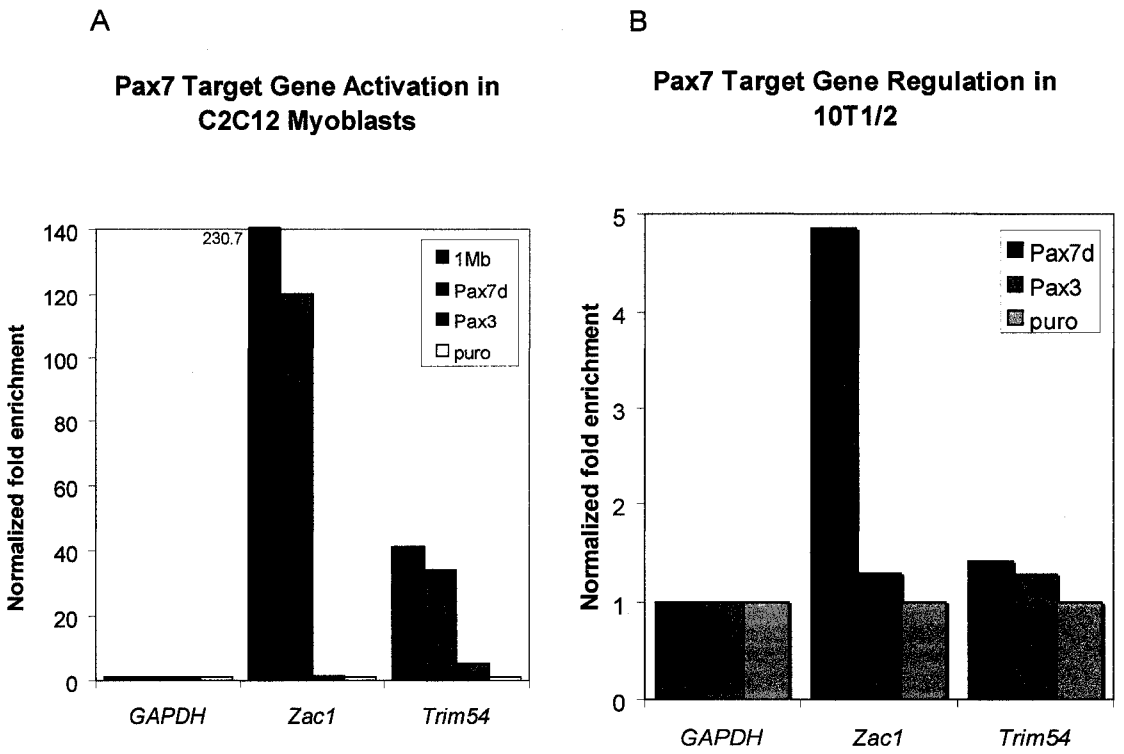


Figure 2. *Zac1* is strongly activated by ectopic expression of Pax7. (A) C2C12 myoblasts were stably infected with Pax7 retrovirus. Zac1 expression in primary myoblasts (1 Mb) is shown for comparison. (B). 10T1/2 fibroblast cell line was stably infected with Pax7 retrovirus. Fold induction of *Trim54*, a candidate Pax7 target gene identified in the microarray analysis was included as a comparison for Pax7 induced expression levels. Transcript levels of *Zac1* were measured by real-time PCR and normalized to *GAPDH* levels. Pax7 expressing cells were compared to empty vector control (Ishibashi, unpublished).

1.11 Zac1 Biology

Zac1 is a member of the small pleiomorphic adenoma gene (*Plag*) family which are classified together based on their seven canonical C2H2 zinc-finger motifs (Abdollahi, 2007). *Zac1* was serendipitously cloned along with p53 in a screen based on its ability to activate the type I PACAP receptor in proliferating mouse corticotropic tumor cell line (Spengler et al., 1997). Mouse *Zac1* is located on chromosome 10 and maps to human chromosome 6q24, a region associated with various neoplasms including breast cancer (Varrault et al., 1998). Reporter assays and RT-PCR on pituitary cells determined that the transcriptional start site is located in exon 7 (Spengler et al., 1997). Human *ZAC* shares 73 percent homology with mouse *Zac1* and many splice variants have been identified revealing its complex transcriptional nature (Bilanges et al., 2001; Smith et al., 2002). While the possibility that mouse *Zac1* undergoes similar alternative splicing, to date, only two isoforms have been reported (Huang and Stallcup, 2000). The basic structure of the protein consists of the DNA binding zinc-finger domain in the N-terminus, a linker region and a central proline repeat region (PME/ PLQ) suggestive of a transactivation domain unique to the mouse (Huang and Stallcup, 2000; Spengler et al., 1997).

Zac1 activates gene transcription by coordinately binding DNA with its zinc-finger domain and using its C-terminal region to recruit p300 to activate histone acetyl transferase (HAT) activity (Hoffman, 2006). Intriguingly, activation of gene expression by Zac1 is dependant on homo-dimerization and differential binding of a palindromic consensus motif (Hensen et al., 2002). Dimer binding to

a double consensus motif confers activation of gene transcription whereas monomer binding to one consensus sequence results in repression. Appropriate spacing of binding sites also impacts gene transcription (Hoffmann et al., 2003).

Zac1 expression is associated with anti-proliferative properties, functionally classifying it as a putative tumor suppressor (Spengler et al., 1997). Indeed, Zac1 was found to inhibit tumour formation in nude mice through induction of G1 cell cycle arrest and apoptosis. *In vitro*, Zac1 displayed the same functional characteristics as p53 in tumor cells based on bromodeoxyuridine (BrdU) incorporation, growth kinetics and fragmentation of DNA upon Zac1 induction in tumour cells (Spengler et al., 1997). Therefore, although structurally unrelated, these two proteins posses the unique capacity to induce both anti-tumour pathways. Moreover, later studies revealed that Zac1 enhances p53 activity in co-transfection reporter assays (Huang et al., 2001).

Striking a balance between cell division, growth arrest, and programmed cell death are fundamental to the myogenic pathway. For example, the regenerative capacity of muscle is tightly coupled to cell cycle progression. The activity of the MRFs has been directly associated to cell cycle regulation in directing the transition from highly proliferative myoblasts to terminally differentiated cells in cell cycle arrest. The role of the retinoblastoma protein (Rb) during myogenic differentiation has been extensively studied and has been shown to be critically involved in the transition from proliferation to differentiation (Huh et al., 2004). Rb transcription is activated by MyoD and induces cell cycle arrest through interaction with E2F transcription factor that together repress

transcription of genes required for entry into S-phase (Huh et al., 2004). Zac1, like Rb, has been classified as a tumour suppressor due to its ability to induce cell cycle arrest (Bilanges et al., 2001; Pagotto et al., 1999; Spengler et al., 1997). While it is known that Zac1 activity proceeds independently of other characterized cell cycle proteins (Rb, p21, p27, and p16) (Spengler et al., 1997), the exact molecular mechanisms by which it exerts its effects remain undefined. A recent study has suggested that Zac1 interacts with the cytokine TGFβII through a negative feedback loop to regulate cell numbers during retinal development (Ma et al., 2007). Though no studies to date have investigated Zac1 function in skeletal muscle development and regeneration, activation of Zac1 could play a role in directing terminal differentiation through cell cycle regulation.

Expression profiling in the embryo through *in situ* hybridization demonstrated that *Zac1* was expressed in the highly proliferative progenitor cells of many tissues throughout development as revealed by overlapping expression of stem cell markers including Nestin, PCNA, GFAP, FORSE-1 and BrdU (Piras et al., 2000; Valente et al., 2005). Concordantly, high levels of *Zac1* transcripts were observed during somite formation, in migrating muscle precursor cells and during waves of fibre formation late in development. Lower levels of *Zac1* transcripts were observed in skeletal muscle post-natally (Valente et al., 2005). Based on expression patterns of *Zac1*, the authors hypothesized that Zac1 may play a role in directing cell fate decisions as cells transition from a mitotic to a post-mitotic state during tissue specification.

More recently, *Zac1* has been identified as a maternally imprinted gene (Piras et al., 2000; Varrault et al., 2006). Through a novel screening technique, a differentially methylated region that maps to 60 kb upstream of *Zac1* and controls its expression, was detected and defined as an imprinted control region (ICR) (Varrault et al., 2006). Aberrant expression of imprinted genes has been correlated to developmental defects. In support of this theory, hypomethylation of the *Zac1* ICR is found in transient neonatal diabetes (Smith et al., 2002). Furthermore, it was recently identified as a member of an imprinted gene network that controls embryonic growth and differentiation (Varrault et al., 2006). Interestingly, in this study Zac1 was found to alter expression of *Peg3*, *Mest*, and *Cdkn1*, which were also identified in our Pax7 microarray analysis (McKinnell et al., 2008). In addition to these activated imprinted genes, Zac1 was also shown to directly bind to the H19 3' enhancers, resulting in the transactivation of the Igf2 and H19 promoters. Igf2 is a mitogenic growth factor that is involved in modulating muscle growth and differentiation (Pedone et al., 1994).

While insight into Zac1 function in tumour suppression, embryonic development and imprinting is becoming clearer, its role in post-natal myogenesis is unknown. In this study, we have identified a novel Pax7 consensus sequence located +11 kb downstream of *Zac1*'s transcriptional start site, shedding new light on the mechanisms of Pax7 regulation in post-natal myogenesis. The elucidation of the Pax7 binding site is a critical step in defining the cellular phenotype of satellite cells. Furthermore, preliminary results indicate that Zac1 plays an essential role in regenerative myogenesis. Further

understanding of the function of Pax7-activated expression of Zac1 will provide insight into the mechanisms of satellite cell mediated repair.

1.12 Statement of Intent

The identification of putative Pax7 target genes through microarray studies has provided us with the opportunity to gain insight into the mechanisms of satellite cell mediated myogenesis. In particular, *Zac1* was strongly induced by ectopic expression of Pax7 in C2C12 myoblasts and was therefore an intriguing candidate target gene with which to study the molecular mechanism of Pax7 function. Based on its expression profile in differentiating myoblasts as well as its activation following Pax7 induction, we hypothesized that *Zac1* was directly regulated by Pax7. We therefore defined the region of *Zac1* bound by Pax7 through a multi-pronged approach using a series of classical chromosomal mapping techniques. Unexpectedly, our studies revealed a novel, non-canonical Pax7 DNA-binding motif that was unrelated to previously identified Pax consensus motifs. Furthermore, we described the functional regulation of this binding site by Pax7, therefore supporting the assertion that Zac1 is a direct Pax7 target gene in myogenesis. Finally, we sought to characterize the role played by Zac1 in post-natal myogenesis through over expression studies *in vitro* as well as regeneration assays *in vivo*. Preliminary results suggest a critical function for Zac1 in mediating efficient regenerative myogenesis. Studies to identify Zac1 downstream target genes in the myogenic context will further our

understanding of the cellular phenotype defining satellite cell activity in post-natal myogenesis.

CHAPTER 2

METHODS and MATERIALS

2.1 Cloning and Constructs

The *Zac1 cDNA* was isolated from cultured wildtype proliferating primary myoblasts derived from Balb/c mice by purifying RNA using the RNeasy kit (Qiagen) and subsequent reverse-transcription with Superscript II reverse transcriptase (Invitrogen) following the manufacturers' protocols. The *Zac1* coding sequence was amplified in two segments using 4 primer pairs to minimize the level of misincorporation from PCR amplification. High fidelity polymerase mix containing proofreading activity was also used to optimize sequence accuracy (Expand, Roche). An EcoRI restriction site was included in the 5' forward primer and an XhoI restriction site was added to the 3' reverse primer to allow for directional cloning into the pBRIT retroviral vector. The 3' reverse primer also removed the stop codon to allow for in-frame translation of the 3X Flag epitope tag located in the retroviral backbone. Forward and reverse primers mapping to the middle of *Zac1* overlapped and contained an NdeI restriction site to allow for three-way ligation into the backbone. Amplicons were initially cloned into pCR II plasmids through TA cloning (Invitrogen) and screened by EcoRI restriction enzyme digest. Positive clones were sequenced (Stemcore Laboratories) and correct inserts then subcloned into the pBRIT retroviral expression vector via an XhoI, NdeI and EcoRI three-way directional ligation.

pBRIT expression vectors were used to generate stable cell lines by retroviral infection, and to express Zac1 in luciferase reporter assays by transfection. The pBRIT backbone is based on the pHAN retroviral expression vector whereby Pax7, Pax7/FKHR or *Zac1* are expressed with 3X carboxy-

terminus Flag epitope tags and puromycin resistance is driven by an independent SV40 promoter.

Luciferase reporter vectors. The 300 bp probe containing the footprinted region +11 kb upstream of *Zac1* transcriptional start site was amplified by PCR using template derived from wildtype mouse genomic DNA. Primers contained BamHI (forward) and BglII (reverse) restriction sites. Amplicons were digested with BamHI and BglII restriction enzymes and subcloned into the multiple cloning site of pGL4.23 luciferase reporter plasmids (Promega) containing the TK minimal promoter. Ligations were conducted overnight at 16°C with T4 DNA ligase (Invitrogen) following the conditions outlined by Invitrogen. Positive clones were screened by BglII restriction digest followed by sequencing (Stemcore Laboratories). All plasmids were transformed into chemically competent TOP 10 cells (Invitrogen) and plated on LB agar medium with 25 µg/mL ampicillin for selection. Bacterial cultures were grown overnight at 37 °C. Large-scale preparations of plasmid used for all transfections were purified by cesium chloride density gradient ultracentrifugation (Maniatis, 1989).

The pGL4.74 reporter plasmid (Promega) was co-transfected in reporter assays to express Renilla luciferase as an internal control.

2.2 Cell Culture, Infection, and Transfection

C2C12 myoblasts were obtained from American Type Cell Culture (ATCC) and cultured in growth conditions in DMEM with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (100 U/mL & 100 µg/mL respectively; Invitrogen). Primary myoblasts were isolated from the hind limbs of 3-6 week old male Balb/c mice. Dissociation of the muscle was carried out in collagenase/dispase (2 mg/mL dispase II (Roche), 1 mg/mL collagenase B (Roche), 0.25 µM $CaCl_2$ (Sigma) in sterile PBS) for two rounds of 12 minutes at 37°C (5% CO_2), followed by trituration. Digestion was stopped by addition of 5% FBS in DMEM and the suspension was filtered through nylon mesh filters (74 µm pore size, Fisher). Cells were then resuspended in Ham's F-10 medium with 20% FBS, 1% penicillin/streptomycin and 12.5 ng/µl of recombinant human bFGF (invitrogen) and pre-plated on 10 cm plastic culture dish for 2.5 hrs. Non-adherent cells were transferred to rat-tail collagen-coated tissue culture dishes (Roche, Costar). Contaminating fibroblasts were removed through a series of pre-plating techniques that involved passaging cells with 2.5% trypsin for 3 minutes, and plating on 10 cm plastic dishes for 20 minutes. Non-adherent cells were then plated on collagen-coated plates and maintained in the selective myoblast media. Myoblasts were maintained in growth conditions whereby cells were passaged at 70% confluence to a maximum of 12 passages.

Stable cell lines ectopically expressing Pax7-Flag were established in both primary myoblasts and C2C12 cells through retrovirus made from Pax7 pBRIT retroviral expression plasmids (Ishibashi et al., 2005). Empty vector pBRIT cell lines were established as controls. Ecotropic retrovirus was produced in the

stable Phoenix-ECO viral packaging cell line (Invitrogen) derived from 293T cells, maintained in DMEM with 10% FBS, 1% penicillin/streptomycin. Viral supernatant was filtered through a 0.45 µm syringe filter (Millipore) then applied directly added to C2C12 and primary myoblasts at a 1:2 dilution with 8 µg/mL of polybrene in DMEM with 10% FBS. Selection was carried out beginning 48 hours after infection using 1.5 µg/mL puromycin for 10 days.

Luciferase assays were conducted in C2C12 myoblasts grown under growth conditions. Cells were seeded onto 12-well culture plates (Costar) 24 hours prior to transfection at a density of 1×10^4. Co-transfections were carried out using the lipid-based transfection reagent Lipofectamine (Invitrogen) according to the manufacturer's protocol. Cells were lysed 24 hours following transfection by addition of 100 µl of passive lysis buffer (Promega) for 15 minutes at room temperature. Both firefly and Renilla luciferase reporter activities were measured with an automated 96-well luminometer (Beckman Coulter) whereby 200 µl of firefly luciferase assay substrate was injected into 10 µl of protein lysate, followed by injection of 10 µi quencher and Renilla luciferase substrate (Stop and Glo reagent, Dual Luciferase System; Promega).

2.3 Binding Site Search

Genomic sequence from LocusLink (NCBI) were searched with the Pax3 paired-box consensus sequence 'tcgtcacrchya' (Chalepakis and Gruss, 1995).

2.4 ChIP and Western Blotting

Chromatin immunoprecipitation (ChIP) experiments were conducted on C2C12 myoblasts that were ectopically expressing Pax7-Flag, and compared to the empty vector puro control. Cells were cultured on 10 cm plates in growth conditions in DMEM, 10% FBS and 1.5 µg/mL puromyin and directly fixed in a 1% formaldehyde fixation buffer (1 mM EDTA, 1mM EGTA, 50 mM Hepes, 100 mM NaCl). Lysis was carried out in 40 mM Tris (pH 8.0), 1% Triton-X, 4 mM EDTA, 300 mM NaCl for 30 minutes at 4° C. Nuclei were sonicated in a water bath with an optimized protocol of 14 cycles of 30 second pulses plus 2 minutes on ice to yield fragments with an average length of approximately 300 bp. Approximately 2 x 10^6 cells (2 x 10 cm plates) were resuspended in 200 µl of lysis buffer per tube, in duplicate for a total of 4 x 10 cm plates per ChIP sample. Cellular debris was removed by centrifugation for 10 minutes at 13000 rpm. Protein concentration was quantified by Bradford assay (Biorad) and using BSA as a standard. Immunopreciptions were carried out using 1500 µg of total chromatin. Input lysates were precleared with protein-A Sepharose beads for 1 hour at 4°C and then immunoprecipitated with 50 µl of monoclonal anti-Flag (M2) antibody directly conjugated to agarose beads (Sigma) overnight at 4°C on a rotating platform. Elution and reversal of cross links were carried out as per the protocol outlined in the ChIP assay kit (Upstate Biotechnology). Immunoprecipitated DNA was analyzed by real-time PCR using a SYBR green universal mix (BioRad). Primers were designed to amplify genomic regions -42 kb (upstream) and +11 kb (downstream) of the *Zac1* transcriptional start site.

Fold-enrichment was measured relative to the empty vector control and normalized to -22.9 kb from *MyoD* or -3 kb from *GAPDH.*

Western blot analysis was conducted on direct and immunoprecipitated protein extracts from cultured cells. Treated cells were harvested in PBS and then lysed in RIPA buffer for protein extraction at 4°C for 30 minutes (Maniatis, 1989). Protein was quantified by Bradford assay. A portion of both input lysates and immunoprecipitated lysates were added to sample buffer. Samples were denatured at 100°C for 5 minutes in 2X Laemmli sample buffer (100 mM Tris-HCL pH 6.8, 4% SDS, 200 mM DTT, 20 % glycerol and 0.1% bromophenol blue) and cooled on ice. Samples were electrophoresced through an 8% SDS-polyacrylamide gel (375 mM Tris-HCl pH 8.8, 0.1% SDS, 0.05 % APS, 0.05% TEMED). Proteins were transferred overnight at room temperature to an Immobilon P membrane (Millipore) at 20 V. Blots were incubated in primary antibody plus 5% skim milk powder in PBS for 1 hour at room temperature. Antibodies for Pax7 was derived from Hybridoma supernatant (ATCC) and used at a dilution of 1:15 in blocking buffer. Polyclonal anti-Flag antibody (M2; Sigma) was used at a 1:5000 dilution. Following washing in 0.2% PBST, membranes were blotted with an appropriate anti-mouse or anti-rabbit secondary antibody conjugated to horseradish peroxidase (BioRad) for 1 hour at room temperature in 5% skim milk/PBS. Signal was detected by with enhanced chemiluminescence reagent (Upstate Biotechnology) and exposed on XAR film (Kodak).

2.5 DNase I Footprinting

Probes that were approximately 350 bp in length were amplified from genomic DNA isolated from primary mouse myoblasts and spanned the 1 kb region 11 kb downstream of the *Zac1* transcriptional start site identified in the CHIPseq experiments. Primers amplifying the sense strand were labeled with FAM fluorescent dye (Invitrogen). Amplicons were then purified using micro spin columns (Qiagen) in a final volume of 10 µl. The concentration of the probe was approximately 50 ng/µl based on comparison to lambda-HinDIII digested DNA ladder. Approximately 100 ng of probe was used for each binding reaction, which were carried out in a buffer consisting of 150 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 8% glycerol and 10 mM Tris-HCl. Both 1 µg of *in vitro*-translated Pax7 protein and purified recombinant Pax7 protein were used in the binding reactions which were carried out at room temperature for 20 minutes. Protein-DNA complexes were then digested with DNase I (1 µg/mL) (Cedarlane) at a dilution of 1:10 for 5 minutes and stopped by addition of 0.25 M EDTA. Digestion conditions were determined following a series of optimization experiments where time and concentration variables were adjusted to obtain an even distribution of fragment sizes. Binding reactions with labeled probe and BSA were used as negative controls. Reactions were purified using a concentrating spin column (Qiagen) and fragment sizes were resolved using an ABI 3730 capillary-based automated DNA sequencer (Applied Biosystems) to generate an electropherogram representing the footprint. Template probes were then sequenced and the chromatogram superimposed on the electropherogram to determine the exact sequence of regions that were bound by Pax7.

2.6 Electromobility Shift Assays (EMSA)

Oligonucleotides were designed to represent the region protected in the Zac1 footprint experiments that contained a 10 bp core sequence similar to the *Myf5* Pax7 protected region. The 5' phosphate group of the forward strand was labeled through an addition of radiolabeled ^{32}P mediated by T4 DNA kinase (Invitrogen). The forward labeling reaction was carried out on 3 pmol of forward oligonucleotide for 30 minutes at 37°C in 70 mM Tris-HCl pH7.6, 10 mM $MgCl_2$, 100 mM KCl, 1 mM beta-mercaptoethanol and 0.75 µCi γ-[^{32}P]-ATP (GE Healthcare). The labeling reaction was stopped through addition of 30 µl of annealing buffer (66 mM Tris-HCl pH 8.0, 66 mM NaCl, 1.33 mM EDTA). Complimentary reverse strand oligonucleotides were then added to the labeled strands in a 2:1 ratio, boiled for 5 minutes and gradually cooled to room temperature to generate labeled double-stranded DNA. Radiolabeled probes were purified through a G-25 column (GE healthcare). Cold competitor probes were generated following the same annealing conditions. Purified recombinant Pax7 protein (0.5 µg) was added to 3 pmol of probe in a binding reaction mixture containing 75 mM NaCl, 1mM EDTA, 1mM DTT, 10 mM Tris-HCl (pH 7.5), 6% glycerol, 2.5 µg of BSA, 1 µg of poly-dIdC (Sigma) and 0.2 µl of rabbit reticulocyte lysate (Promega). Competition experiments were carried out with varying amounts of either cold wildtype probe or probes whose entire 10 bp core sequence had been mutated with purine/pyrimidine base substitutions. *In vitro* translated MyoD protein was used as a negative control; Pax3 (Bajard et. al., 2006) was used as a positive control. Binding reactions proceeded for 20 minutes at room temperature, then were loaded onto a pre-run 0.5X TBE

acrylamide gel (49:1 bis-acrylamide). 180 V were applied to the gel for 1.5 hours and the gel was dried directly onto Whatman paper at 60°C for 1 hour. Gels were exposed to Biomax film (Kodak) overnight at -80°C.

2.7 Real-time and RT-PCR

Primers: PCR primers for RT- and real time PCR were designed online using Primer3 (http://Frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi; Whitehead Institute) (Rozen and Skaletsky, 2000). Primers used in real-time RT-PCR experiments were designed to span at least 1 intron. Primer sets used in RT-PCR as well as in ChIP assays amplified 100-150 bp fragments. Primer sequences are provided in Supplementary Table S1 and S2.

RT-PCR: Total RNA was isolated using RNeasy miniprep kits (Qiagen) and diluted 4-fold with RNase-free water to produce a common pool for analysis with multiple primer pairs. Reverse transcriptase reactions to generate cDNA were performed using Superscript II (Invitrogen).

Real-time PCR: SYBR green real-time PCR reactions were conducted in duplicate using MX4000 and MX3000 PCR machines (Stratagene). Fluorescence readings were based on FAM levels and ROX was included in the reactions as a loading normalization reference dye. Reactions were composed of 1X iQ SYBR green supermix (BioRad), 30 nM ROX passive dye, 50 nM of each forward and reverse primer and 2 µl of template in a final volume of 20 µl.

Fluorescence data was gathered for 40 cycles of {94°C - 30s; 58°C - 60 s; 72°C - 30s}. Primer specificity was validated by denaturation curve analysus (55-94°C) and sequencing of products. Amplification curve plotting and C_t value calculation was performed using MX4000 and MX3000 software (Stratagene) with further analysis performed using Excel (Microsoft).

2.8 Animal Models

To investigate the role of Zac1 in post-natal skeletal regeneration, we used *Zac1/Lacz* knock-in mice where β-galactosidase cDNA was expressed under the transcriptional control of the endogenous *Zac1* promoter through conventional gene targeting techniques (Fan et al., 2005), a kind gift from Serguei Kozlov (National Cancer Institute, Frederick, MD). Mice with the homozygous mutant allele were maintained on a C57Bl/6 background. Wildtype control mice (Charles River Laboratories) were also maintained in the C57Bl/6 genetic background.

2.9 Regeneration Assay

Adult mice (8-12 weeks old) were anaesthetized by halothane following the University of Ottawa Animal Care guidelines. Acute muscle injury to the tibialis anterior (TA) muscle was induced through an intramuscular injection of 25 µl of 10 µM cardiotoxin (CTX) in 0.9% sterile saline (Latoxan). Both the injured and contralateral TA muscles were dissected at 7 days, 12 days and 19 days post-injury. Muscles were weighed and fixed in 2% PFA at 4°C for 48 hrs prior to paraffin embedding and sectioning. Hematoxylin and eosin staining (H&E) techniques were used to delineate muscle fibre morphology as well as nuclei. Experimental groups consisted of two mice for each time point. Muscle hypertrophy was assessed as a measure of percent muscle weight increase in injured TA muscle relative to the contralateral resting muscle. Fibre number was quantified on H&E stained sections. Fibre numbers were quantified in through direct enumeration on H&E stained sections in a representative field for each sample through the use of ImageJ software (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://rsb.info.nih.gov/ij/, 1997-2008.).

CHAPTER 3

Results

3.1 Expression profile of *Zac1* in primary myoblasts.

Previous genome-wide comparative microarray studies identified a set of candidate Pax7 target genes in C2C12 myoblasts (McKinnell et al., 2008). *Zac1* was the most dramatically activated following Pax7 induction. We determined the expression profile of endogenous *Zac1* in primary myoblasts derived from wild-type adult muscle tissue through real-time PCR (Fig. 3). Fold-activation was measured relative to GAPDH expression and primers were designed to span the exon 7/8 boundary of *Zac1*. Under growth conditions, proliferating myoblasts express high levels of *Zac1* transcripts. Upon induction of myoblast differentiation through serum withdrawal, *Zac1* expression is dramatically down-regulated within 24 hours and its expression level remains low throughout differentiation (Fig. 3A). Expression of *Pax7* through differentiation of wildtype primary myoblasts reveals a strikingly similar profile (Fig. 3B). Therefore, these data are consistent with the hypothesis that Zac1 transcription is directly regulated by Pax7 during the myogenic developmental program.

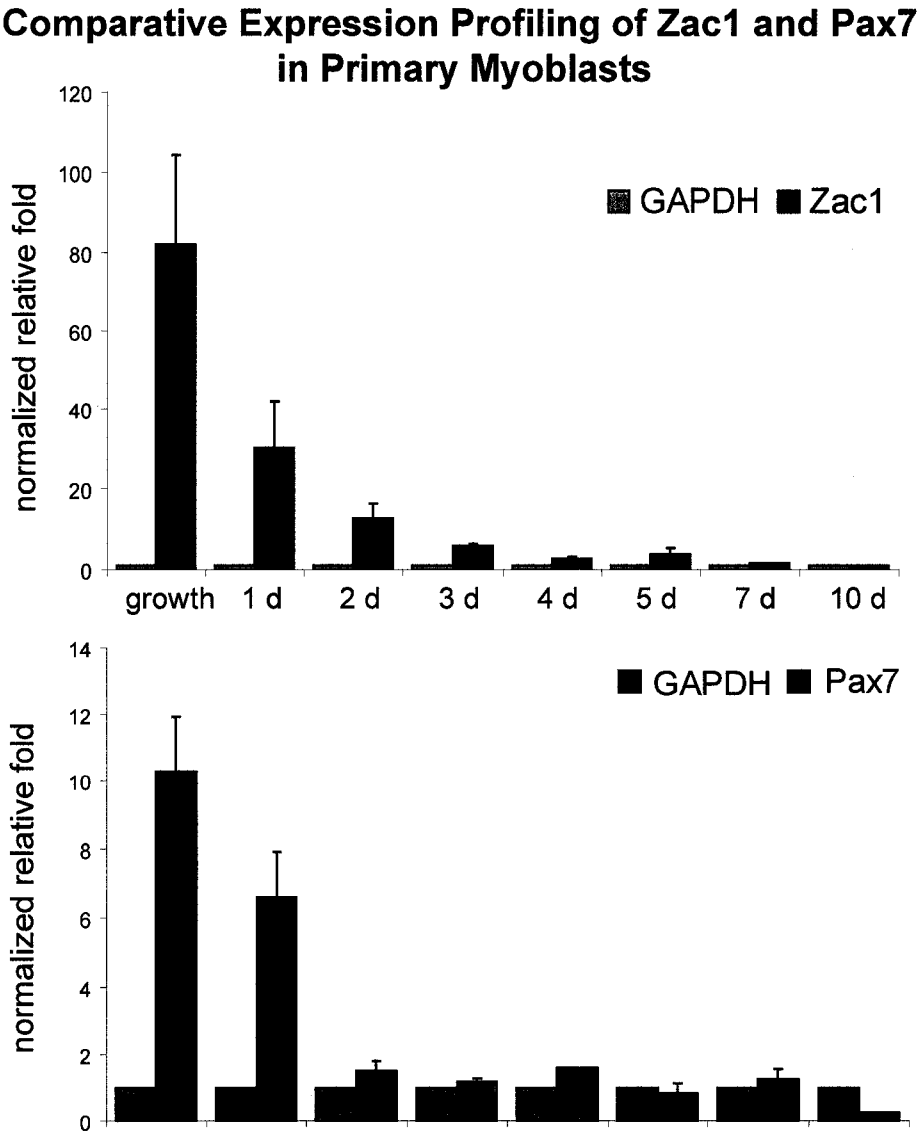


Figure 3. *Zac1* and *Pax7* share similar expression profiles throughout differentiation of primary myoblasts. Primary myoblasts were forced into differentiation through serum withdrawal over a period of 10 days. Expression levels were quantified by real-time RT-PCR and normalized to *GAPDH* levels. Fold change in expression is measured as a moving average of duplicate samples.(A) *Zac1* expression levels were high while myoblasts were maintained under growth conditions and dramatically down regulated 24 hours following the onset of differentiation. *Zac1* expression levels continued to fall throughout differentiation. (B) *Pax7* expression profile of differentiating myoblasts parallels that of *Zac1* where myoblasts express high levels of *Pax7* during growth and subsequently down-regulate its expression upon terminal differentiation.

3.2 Isolation and identification of *Zac1* splice variant

We isolated and identified a specific *Zac1* splice variant present in wild-type primary myoblasts (Fig. 4C,D). Identification of the isoform expressed in proliferating myoblasts was determined through RT-PCR using primers that mapped to both the previously identified 3' end as well as a recently described alternatively spliced 3' end. Amplicons reflecting the splice variant were uniquely present in proliferating myoblasts. RT-PCR products were then directly cloned into the pCR II vector and sequenced, revealing sequence identical to a previously described *Zac1* splice variant present in 17-day embryos (Huang and Stallcup, 2000). In accordance with the characterized Zac1 model molecular structure (Huang and Stallcup, 2000), *Zac1* mRNA from myoblasts encode seven Zn-finger domains and a proline-rich region characterized by PLE repeats and PE/E repeats. The structure of this splice variant differs from the *Zac1* sequence originally isolated from adult mouse pituitary cells (Spengler et al., 1997) in two ways: it contains a sequence for an additional 11-amino acid insertion following exon 8 and has an extended 3' end that encodes a distinctive 36 amino acids (Fig. 4C). The functional consequences of the splice events are unknown; however, the splice isoform originally identified in the adult mouse pituitary gland are absent from proliferating myoblasts as determined by RT-PCR using primers amplifying the 3' end.

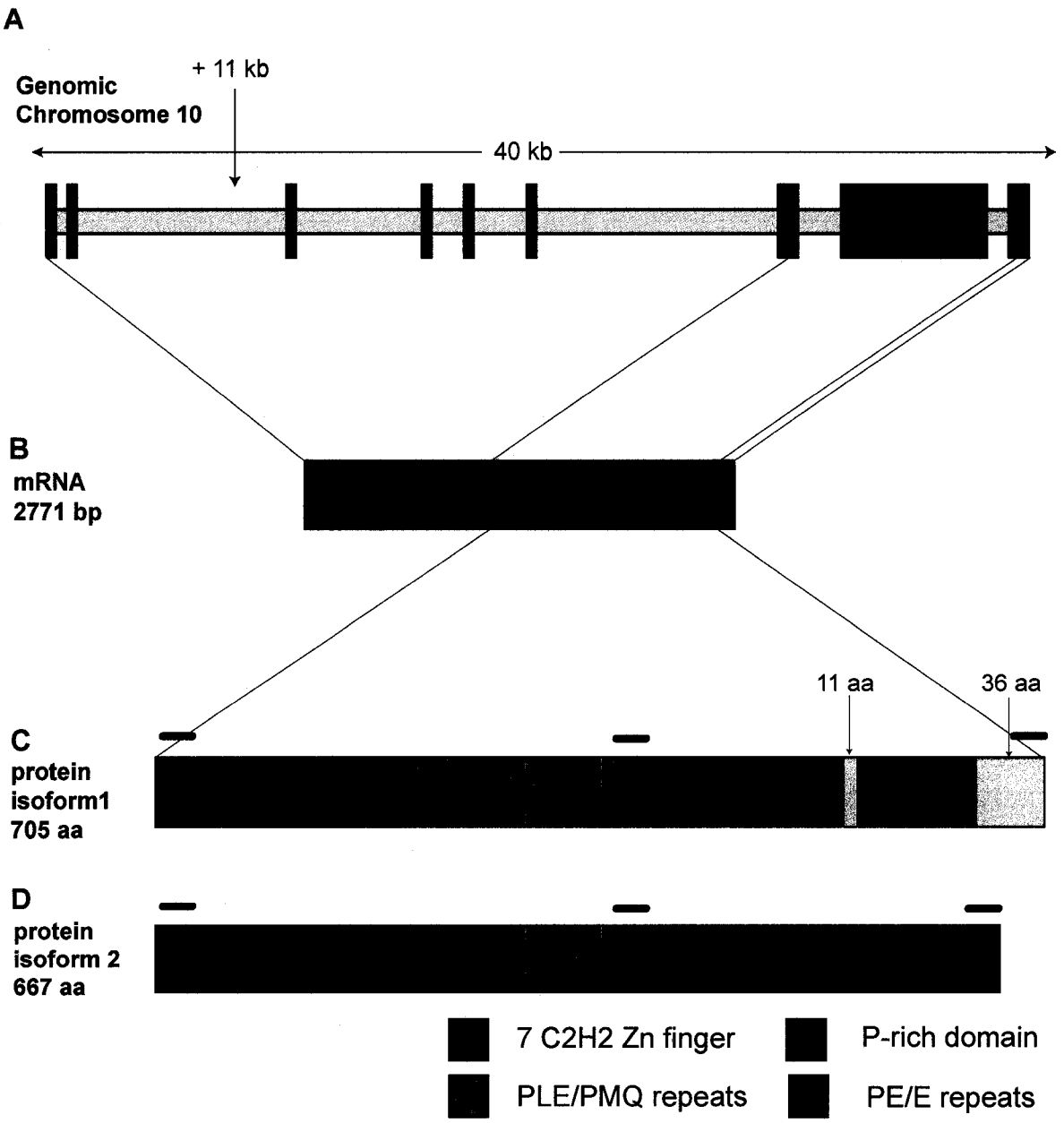


Figure 4. Primary myoblasts express a Zac1 splice variant. A schematic representation of mouse Zac1. (A) Representation of genomic *Zac1* and its mRNA transcript. Shown in blue is non-coding mRNA. (B) Zac1 protein isoforms. Zac1 contains 7 zinc finger motifs, two distinct regions containing PLE and PMQ repeats and a proline rich domain. (C) The Zac1 isoform expressed in proliferating primary myoblasts contains an 11 amino acid insert as well as a carboxy terminus that differs in 36 amino acids as determined by RT-PCR. This isoform was also identified in RNA derived from whole D 17 mouse embryos (Huang et al., 2000) (D) Schematic representation of the Zac1 isoform expressed in adult pituitary murine cells (Spengler et al., 1997). Black lines indicate primers used to identify splice variants by RT-PCR

3.3 Over-expression of Pax7 up-regulates expression of *Zac1*

Pax7 induces expression of *Zac1* in microarray experiments conducted on C2C12 myoblasts ectopically expressing Pax7 (McKinnell et al, 2008). To further characterize the relationship between Pax7 expression and *Zac1* induction in a more physiologically relevant *in vitro* model, we ectopically expressed Pax7 in primary myoblasts and quantified the induction of putative Pax7 target genes (*Myf5* and *Zac1*) through real-time PCR (Fig. 5). Pax7-Flag was expressed in primary myoblasts by infecting them with retrovirus containing a puromycin selectable marker. RNA was isolated following stable selection of infected myoblasts. Fold-induction was measured relative to the empty vector control and normalized to *β-actin* levels

Expression of *Zac1* increases 3.5-fold in myoblasts expressing *Pax7* transcripts at levels 18-fold higher than found in control wildtype primary myoblasts. Importantly, MyoD is down-regulated 80% as expected (McKinnell et al., 2008; Olguin and Olwin, 2004). *Myf5*, a novel Pax7 target gene similarly identified in the Pax7 microarray analysis in C2C12 myoblasts (McKinnell et al, 2008), also shows a 13-fold increase in expression in primary myoblasts ectopically expressing Pax7-Flag. Taken together, these data confirm that Pax7 is regulating expression of *Zac1.*

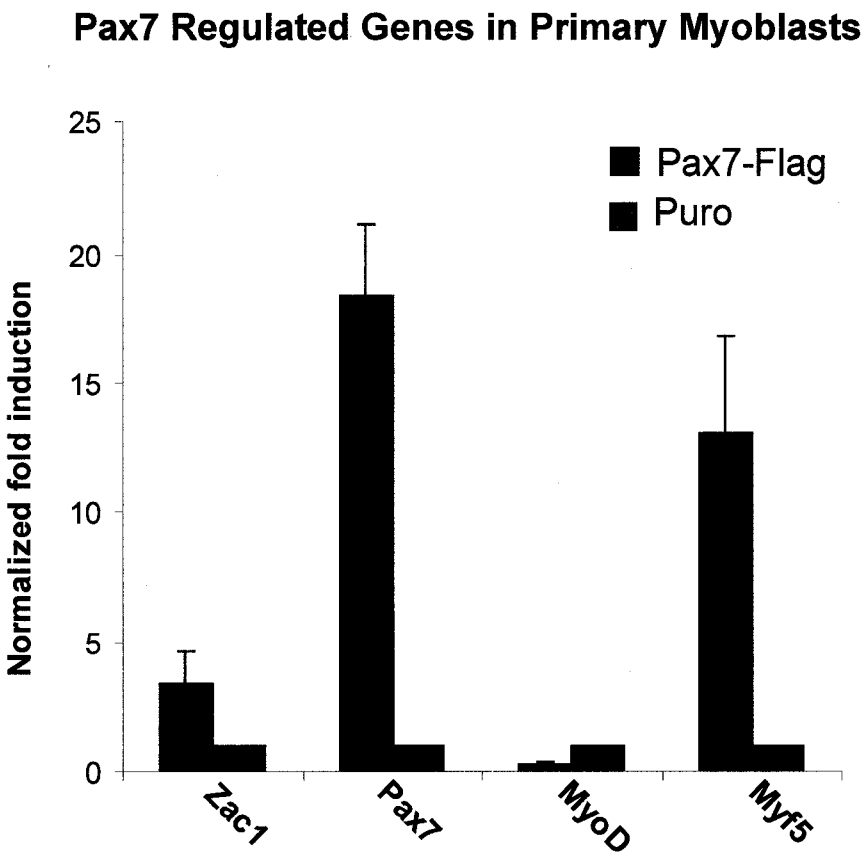


Figure 5. Pax7 regulates expression of *Zac1* and *Myf5* in primary myoblasts. Primary myoblasts were stably infected with a retrovirus expressing Pax7. Expression levels of the candidate target genes, *Zac1* and *Myf5* were examined by real-time PCR. Fold induction was measured relative to primary myoblasts expressing puro resistance alone and normalized to *β-actin* levels. Expression profile is representative of expression levels obtained from three independently infected pools of cells stably expressing Pax7. Ectopic expression of Pax7 increases both *Zac1* and *Myf5* expression levels. *MyoD* expression was not activated by Pax7.

3.4 Pax7 binds *Zac1* intronic sequences

To further define the regulation of *Zac1* by Pax7, we conducted Pax7 chromatin immunoprecipitation assays on C2C12 myoblasts ectopically expressing Pax7. Cells were cultured under growth conditions and expression was verified through Western blot analysis with both anti-Pax7 and anti-Flag antibodies (Fig. 6A). The ectopic expression level of Pax7 was found to be comparable to that of endogenous Pax7 expression in primary myoblasts (data not shown). Sonication of cross-linked DNA was optimized to randomly shear DNA into 300-500 bp fragments to allow for both efficient immunoprecipitation and high-resolution genomic mapping. The currently available Pax7 antibody is a hybridoma cell line supernatant, with a concentration and purity that do not allow for effective immunoprecipitation of Pax7. In light of this, we used a monoclonal anti-Flag antibody to immunoprecipitate Pax7-Flag protein-DNA complexes, followed by real time PCR to map Pax7 interactions with genomic elements at the *Zac1* locus. To ensure that Pax7 DNA complexes were efficiently immunopreciptated, we conducted Western blot analyses at various stages through the ChIP procedure (Fig.6B). The proportion of protein lysate loaded was kept constant to estimate relative levels of immunoprecipitated Pax7 protein versus input. At least 50% of complexed Pax7 DNA was precipitated using the Flag antibody. Therefore recovery was adequate for further immunoprecipitation studies.

Through these chromatin immunoprecipitation studies, we established binding of Pax7 to the previously described Pax3 binding site located in the -57.5 kb enhancer region upstream of *Myf5* (Fig. 7). This result confirmed previous

data that established Pax7 binding to this region through ChIP studies conducted in C2C12 expressing a Pax7-His-Flag protein (McKinnell et al., 2008). MyoD is known not to be directly activated by Pax7 (Olguin and Olwin, 2004), and therefore relative fold-enrichment was normalized to a *MyoD* genomic region. The 6-fold enrichment (relative to the puro control) of immunoprecipitated DNA from the -57.5 kb region indicates a physical interaction between Pax7 and this enhancer.

We identified a putative Pax3 binding site located 29 bp upstream of *Zac1* through sequence scanning based on a 9-of-10 base match to a hybrid sequence derived from a Pax3 consensus sequence (Epstein et al., 1996) and the Pax3 binding sequence described by Bajard et al., 2006 (Fig. 8B). Chromatin immunoprecipitation and real-time PCR were used to assess relative enrichment at distinct sites around the *Zac1* transcriptional start site (Fig.8A). We included a primer set amplifying the *MyoD* -22.9 kb region as a negative control. A 3.5-fold enrichment (comparable to that observed at the -57.5 kb site of *Myf5;* Fig.7*)* was observed in the proximal promoter region of *Zac1* (-100 bp). Lesser enrichment was also observed at +500 bp and at +1600 bp from *Zac1*. This pattern of fluctuating enrichment may reflect the presence of protein-protein interaction causing DNA bending (Kerppola and Curran, 1997). Also, it has been suggested that partial cognate binding sites may surround transcription factor binding sites to create a 'sink' that increases the likelihood of transcription factors binding to their specific sites of target genes (Zhang et al., 2006). It should be noted that the random shearing of DNA through sonication as well as subtle differences in

primer efficiency make high resolution DNA mapping through ChIP challenging. However, our results establish that the promoter region of *Zac1* is bound by Pax7 near a sequence with strong similarity to a Pax3 consensus binding site.

Genome wide ChIP sequencing (ChIPseq) studies conducted on primary myoblasts expressing ectopic Pax7 identified *Zac1* genomic regions at +11 kb and -42 kb with which Pax7 physically interacts (V. Punch, unpublished). To further investigate these potential novel non-canonical Pax7-bound sequences, we designed primers with identity to these two regions and conducted ChIP assays (Fig. 9). Template DNA was enriched by 3.5 fold at the +11 kb locus downstream of the *Zac1* transcriptional start site, whereas the -42 kb locus was only marginally enriched following immunoprecipitation. We therefore focused on further defining the interaction between Pax7 and the +11 kb *Zac1* region.

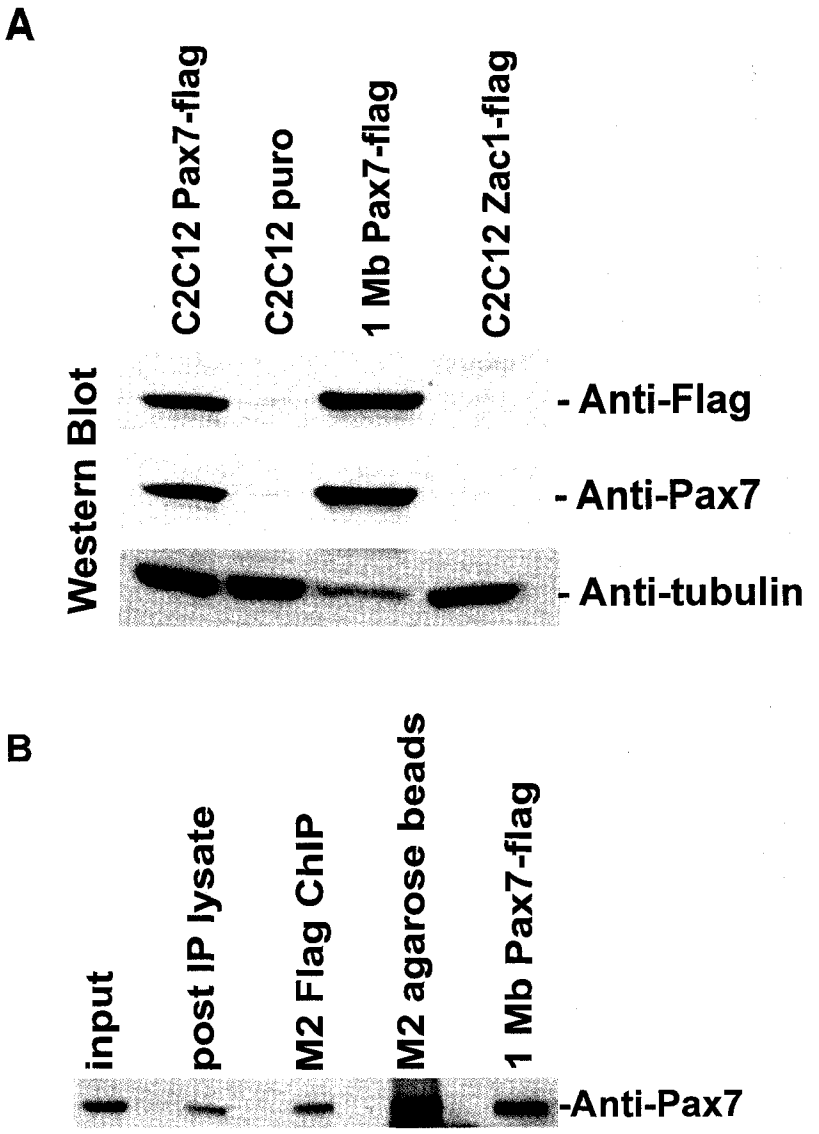


Figure 6. Pax7 is efficiently immunoprecipitated with the Flag M2 antibody.(A) Expression of Pax7-Flag epitope in C2C12 myoblasts was verified using both Flag polyconal and Pax7 monoclonal antibodies. Primary myoblasts (1 Mb) ectopically expressing Pax7 were used as a positive control for both Flag and Pax7 expression. C2C12 myoblasts expressing puro resistance alone show no Pax7 expression. (B) Western blot analysis of ChIP conducted on C2C12 Pax7-Flag cells (A) using a Flag M2 antibody conjugated to agarose beads. Equal proportions of lysate from input material, supernatant following immunoprecipitation and the eluant from the M2 agarose beads were loaded. Following elution, the total amount of M2 agarose beads was denatured and loaded. Based on densiometry analysis, 60% of Pax7 is recovered following immunoprecipitation.

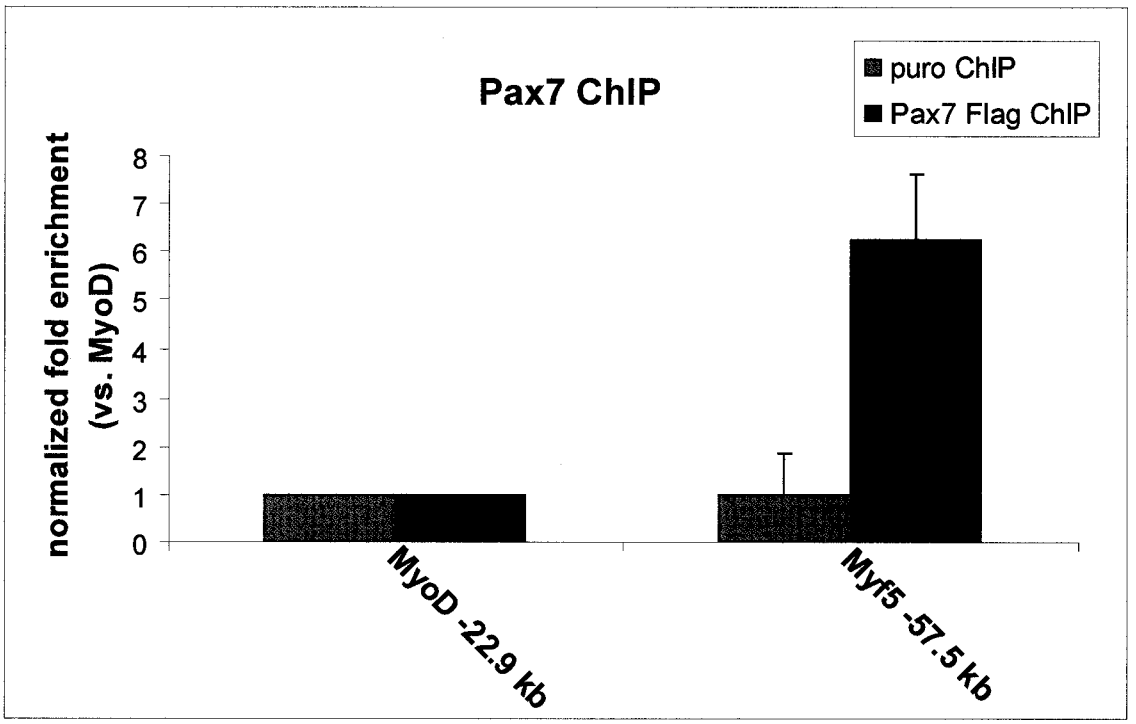
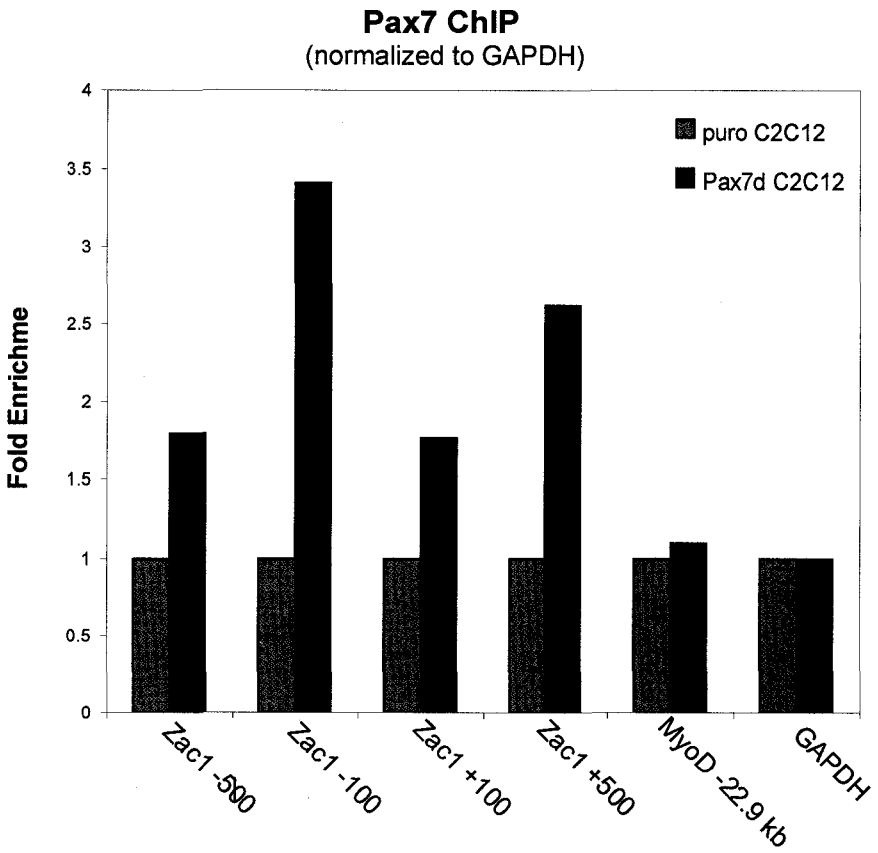


Figure 7. Pax7 binds to a region 57.5 kb upstream of *Myf5*. C2C12 myoblasts were stably infected with Pax7-Flag retrovirus. Triplicate chromatin immunoprecipitations were conducted using a Flag monoclonal antibody. Representative enrichment was measured by real-time PCR and compared to the empty vector control. Relative fold enrichment was normalized to levels of a region 22.9 kb upstream of *MyoD*. A 6-fold enrichment of the *Myf5* enhancer region containing the Pax3 consensus sequence located at -57.5 kb from *Myf5* was observed in C2C12 myoblasts ectopically expressing Pax7.

A



B

Pax3 motif (*Myf5* -57.5kb)	A G T C A T G C C T
Pax3 motif (*Zac1* +29 bp)	C G T C A T G G C T
Pax3 Consensus	C G T C A C G C T T

Figure 8. Pax7 binds to the *Zac1* proximal promoter. Chromatin immunoprecipitations (ChIPs) were conducted in triplicate on cells ectopically expressing Pax7-Flag epitope using a monoclonal flag antibody. Representative enrichment levels were quantified through real-time PCR and measured relative to ChIPs conducted on C2C12 myoblasts expressing the empty vector. Fold enrichment was normalized to *GAPDH* levels. Highest fold enrichment (3.5) was observed at the region 100 bp upstream of the *Zac1* transcriptional start site. The genomic region 22.9 kb upstream of *MyoD* was tested as a negative control and showed no enrichment following Pax7 immunoprecipitation. (B) A Pax3 motif is is located +29 bp upstream of *Zac1*; its sequence is a hybrid of the Pax3 consensus site described by Epstein (1995) and the Pax3 binding site 57.5 kb upstream of *Myf5* (Bajard et al., 2006).

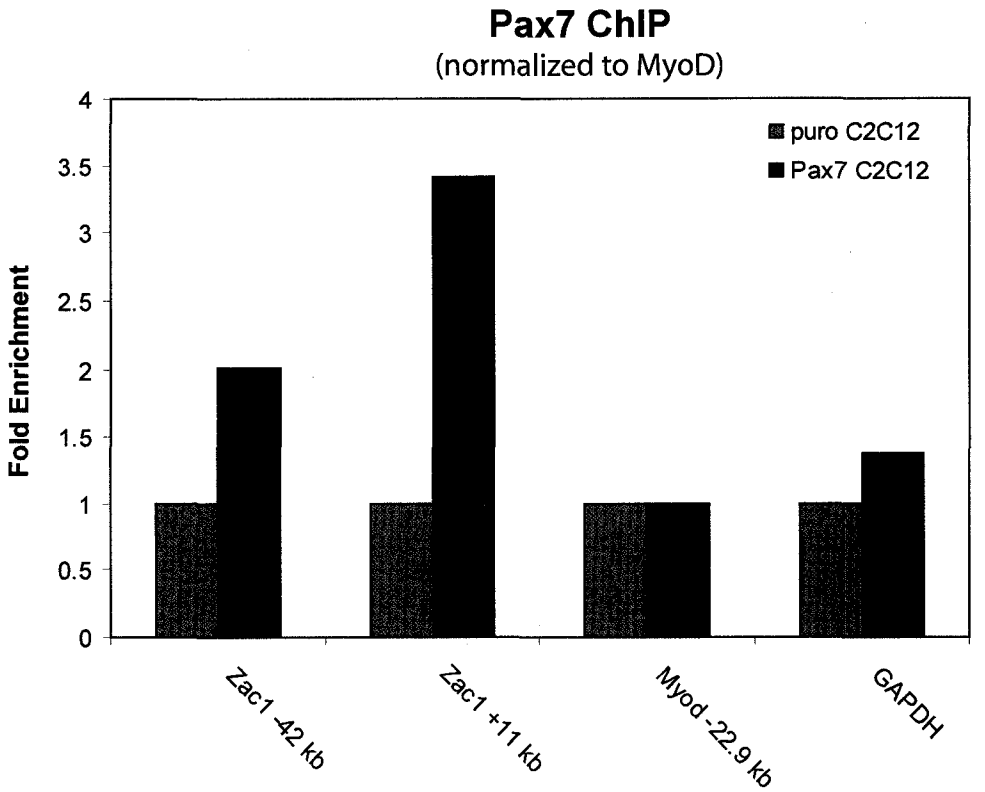


Figure 9. Pax7 binds to a region +11 kb upstream of *Zac1*. Chromatin immunoprecipitations (ChIP) were conducted in triplicate on cells ectopically expressing Pax7-Flag epitope using the Flag antibody. Shown is a representative ChIP experiment. Enrichment levels were quantified through real-time PCR and measured relative to ChIP conducted on C2C12 myoblasts expressing puro resistance alone. Fold enrichment was normalized to genomic *MyoD* levels. Highest enrichment was observed at 11 kb upstream of *Zac1*. This region does not contain a *Pax3* consensus sequence. A marginal enrichment was observed at the -42 kb region from *Zac1*. The Level of *GAPDH* enrichment was included as a negative control.

3.5 Pax7 directly binds a novel sequence in *Zac1*

To establish direct binding of Pax7 to a novel and distinct putative Pax7 binding sequence and to define its sequence, we conducted DNA fingerprint analysis on the +11 kb region in intron 2 of *Zac1* that was enriched following Pax7 ChIP (Fig. 10). We used an adapted footprinting technique which combined the use of a fluorescent automated capillary electrophoresis instrument to resolve DNase I-digested FAM-labeled probes following *in vitro* Pax7 binding reactions. Dideoxynucleotide sequencing reactions were resolved by capillary electrophoresis to define the sequence of *Zac1* +11 kb region bound by Pax7. To map out this region, three probes were generated through PCR amplification using mouse genomic template that spanned a 1 kb region located +11 kb downstream of *Zac1*. Forward primers contained a FAM fluorophore covalently linked to the 5' end. Following DNase I digestion, fragments were resolved using an automated fluorescent capillary sequencer to produce an electropherogram representing a classical footprint cleavage pattern. We replaced Pax7 with BSA in the binding reactions to generate a reference DNase I fragmentation pattern against which to compare those digestion sites protected by Pax7 binding. In addition to the BSA control, we conducted a footprint analysis on the promoter of *Myogenin*, a gene known not to interact with Pax7 during myogenesis (Olguin and Olwin, 2004) (Fig. S1).

Electropherograms from DNase I digestion of BSA and Pax7 *in vitro* binding reactions reveal a Pax7-protected region of 20 bp in the probe spanning 350 bp at +11 kb from *Zac1* (Fig. 10). This protected region is flanked by DNase I hypersensitive sites as characterized by the spiked electropherogram (Fig. 10).

These DNase I hypersensitive regions have been described as hallmarks of transcription factor binding to DNA and have been well characterized in the locus control regions (LCR) of erythrocytes (Cho et al., 2008). It is hypothesized that protein binding results in bending of the DNA, making these adjacent regions more accessible to digestion (Henikoff, 2008). Both *in vitro* translated Pax7 (Fig. 10A) and purified recombinant Pax7-Flag (Fig. 10B) were used in separate binding reactions and found to bind similar regions. While both of these binding reactions produced a footprinted region flanked by hypersensitive peaks, there are subtle differences in the electropherogram possibly reflecting minor changes in binding affinity resulting from the Flag epitope. Importantly, Pax7 footprint analysis conducted on the myogenin enhancers do not reveal any protected regions as expected (Fig. S1).

Dideoxynucleotide sequencing in conjunction with Genemapping software (ABI) was conducted on template using the FAM labeled primers to exactly align the genomic sequence with the footprint. The core 10 bp sequence, consisting of ACGCATTCA, was similarly identified in a region -115 kb upstream of *Myf5* through Pax7 footprint analysis conducted in parallel (V. Punch, unpublished).

Our findings identify a novel Pax7 binding site that is unrelated to the canonical Pax3 sequence. Therefore, these findings demonstrate that Pax7 binds to the *Zac1* gene.

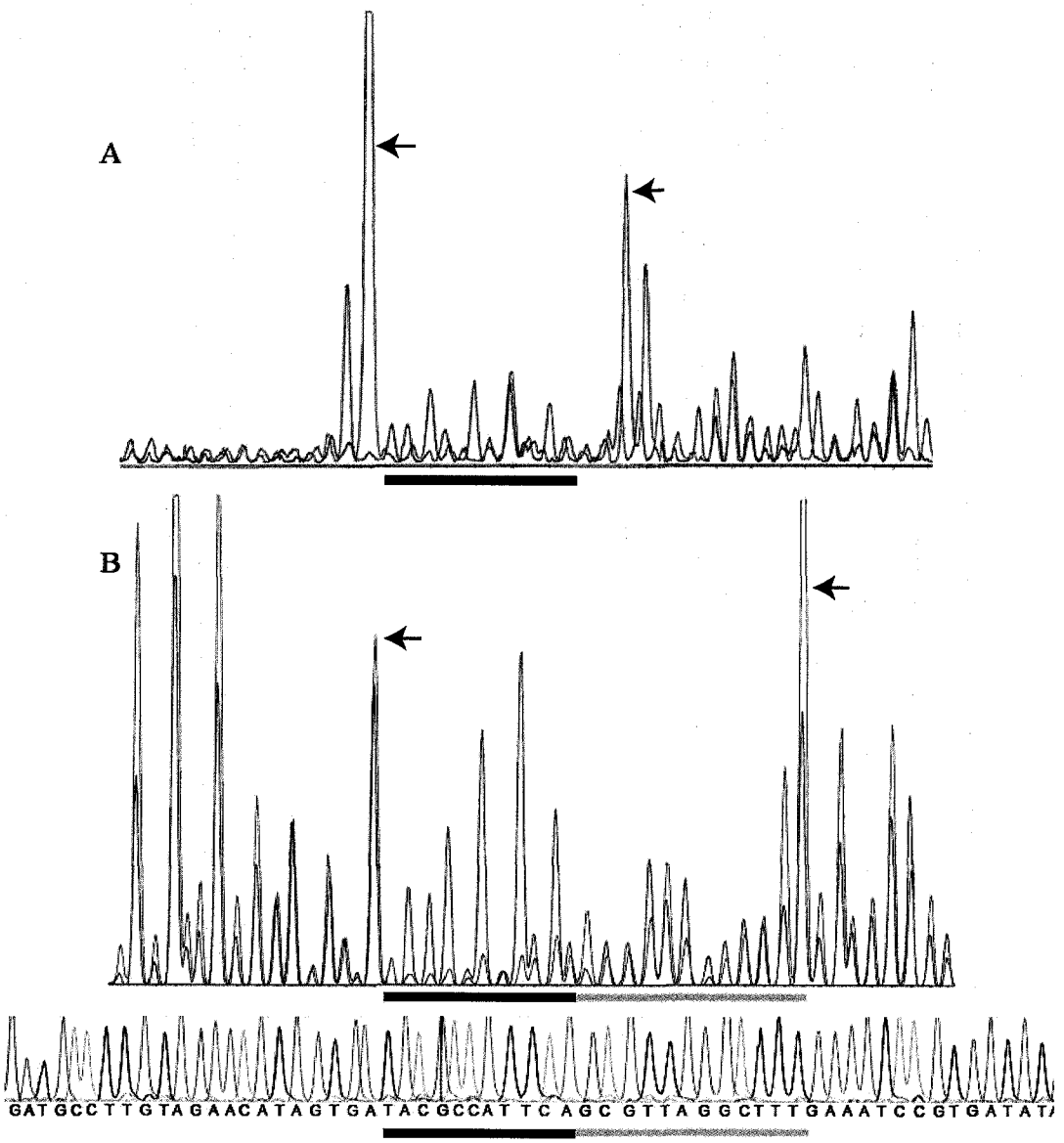


Figure 10. Pax7 directly binds to a novel DNA motif 11 kb upstream of *Zac1*. Electropherogram of DNAse I digested fragments following a binding reaction with Pax7 (red) and BSA (blue). Pax7 protein was generated from In vitro translated Pax7 using rabbit reticulocyte system to couple transcription and translation(A), and purified recombinant Pax7-Flag protein using the bacculovirus expression system (B). Probes were generated through PCR of mouse genomic template using FAM labeled primers that amplified a 1 kb region located 11 kb upstream of *Zac1*. Dideoxynucleotide DNA sequencing on template probe was aligned with the footprint through Genemapper software. Hypersensitive regions (arrows) flank either side of the footprinted region. The sequence underlined in black was found to bind Pax7 115 kb upstream of *Myf5*

3.6 Pax7 specifically binds a 10 bp sequence in *Zac1*

To further confirm that Pax7 binds to the 10 bp core nucleotide sequence identified in the DNA footprint analysis, we conducted electrophoretic mobility gel shift assays (EMSA) combined with competition experiments using *Zac1* ^{32}P-labeled probes (Fig. 11 and 12). Purified recombinant Pax7-Flag protein bound *in vitro* to 40 base pair oligos labeled with ^{32}P, consistent with the previously determined footprinted region of *Zac1*. Similarly, purified recombinant Pax3 protein bound at the Pax3 consensus sequence located -57.5 kb upstream from *Myf5* (Fig. 11). In contrast, binding reactions containing MyoD instead of Pax7 were incapable of binding to either the *Zac1* (Pax7 site) or *Myf5* (Pax3 site) probe (Fig. 11 and 12). Addition of cold wildtype competitor successfully competed with labeled probe with increasing concentration (Fig. 11B, lanes 6-8). Conversely, addition of cold mutant competitor (purine bases exchanged for pyrimidine bases and vice versa) in 100-fold excess did not affect binding to labeled probe, further confirming specificity (Fig. 11B, lane 9). To ensure that the interactions between the purified recombinant proteins and their binding sites were not artifacts of protein-to-DNA stoichiometry, we reduced the concentrations of Pax3 and Pax7 by 10-fold (Fig. 12). We also examined the binding of Pax3 to the non-canonical Pax7 binding site identified in *Zac1* (Fig. 12). Interestingly, Pax3 is also capable of binding to this sequence (Fig. 12, lanes 4-6) in our *in vitro* assay therefore suggestive of 2 unique DNA binding domains.

A

t g a t a c g c c a t t c a g c g

B

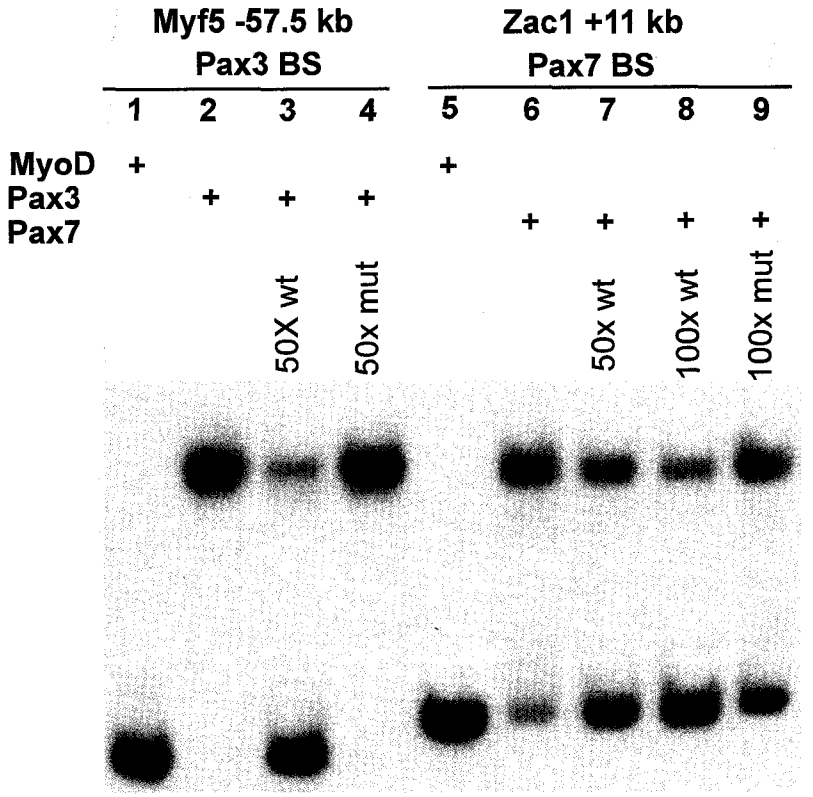


Figure 11. Pax7 specifically binds a novel DNA motif in *Zac1*. (A) Putative Pax7 binding sequence located + 11 kb from *Zac1*. Bases mutated in competition reactions are indicated in red type. (B) Autoradiography of an EMSA showing binding of purified recombinant Pax3-Flag protein to the Pax3 consensus sequence found -57.5 kb upstream of *Myf5* as a positive control (1-4), and binding of purified Pax7-Flag recombinant protein to a novel Pax7 binding sequence identified by footprinting experiments located +11 kb upstream of *Zac1* (5-9). Probes containing these binding sequences were radiolabeled with ^{32}P isotope. MyoD does not bind to the labeled probes (1,5). Pax3-Flag and Pax7-Flag causes a band shift (2,6) Addition of wild type cold competitor in increasing amounts out competes binding of labeled probes accordingly (3, 7, 8). Addition of cold mutant competitor in excess of 100-fold does not out compete binding to labeled probes (4,9)

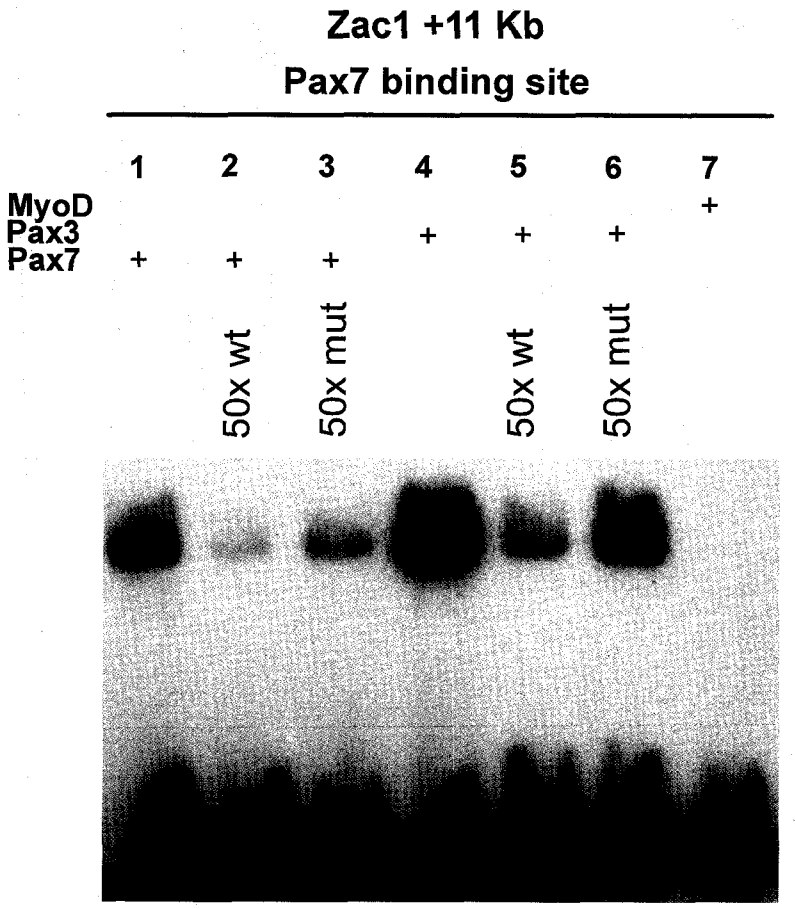


Figure 12. Both Pax3 and Pax7 bind a novel binding motif in *Zac1*
Autoradiography of an EMSA showing binding of purified recombinant Pax7-Flag (1-3) as well as Pax3-Flag (4-6) to the novel Pax7 binding sequence identified by footprinting experiments located +11 kb upstream of *Zac1*. Probes containing this binding sequence were radiolabeled with ^{32}P isotope. MyoD does not bind to the labeled probes (7). Pax3-Flag and Pax7-Flag cause a band shift (1,4) Addition of wild type cold competitor in excess of 50 fold competes with binding of radiolabeled probes (2,5). Addition of cold mutant oligos 50 fold in excess of labeled probe restores the band shift (3,6).

3.7 Functional regulation of *Zac1* by Pax7

Pax7 induces expression of *Zac1* in various cell types (Fig. 2 and 5) and directly binds to a DNA element located 11 kb downstream of the *Zac1* transcriptional start site (Fig. 9-12). To characterize the functional regulation of this element by Pax7, we tested the ability of Pax7 to transactivate expression of a firefly luciferase reporter plasmid consisting of a thymidine kinase promoter controlled by 300 bp of *Zac1* sequence containing the Pax7 binding site (Fig. 13). A constitutive Renilla luciferase reporter was use for internal normalization. Transactivation of the *Zac1* element was compared to a control vector containing 300 bp of *Zac1* genomic sequence lacking the Pax7 binding site. As Pax proteins are known to poorly transactivate target genes (Keller et al., 2004) we also co-transfected a Pax7/FKHR fusion construct with the *Zac1* reporter plasmid into C2C12 myoblasts. Pax7/FKHR fusion proteins contain the N-terminal DNA binding domain of Pax7 and the C-terminal transactivation domain of FKHR (FoxO1). This fusion has been identified as a strong transducer of myogenic gene activation in rhabdomyosarcomas through over-activation of genes bound through the Pax7 DNA binding domain, bypassing the inherently weak Pax7 transactivation properties (Keller et al., 2004).

Our results show a significant effect of Pax7/FKHR on the activity of the reporter containing the *Zac1* element but no effect on the control reporter lacking this sequence (Fig. 13). This data demonstrates functional activation through the Pax7 binding site.

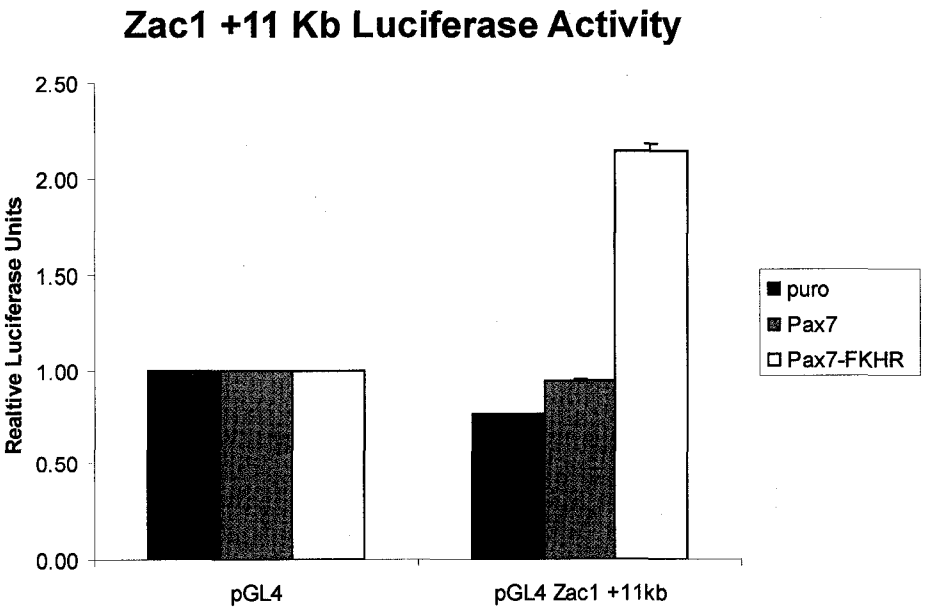


Figure 13. Pax7 activates the novel Pax7 binding site
Firefly luciferase reporter plasmids (pGL4.23) containing the element spanning the +11 kb region of *Zac1* placed upstream of the TK minimal promoter are upregulated 2-fold by Pax7/FKHR. Co-transfections were conducted in C2C12 myoblasts. Firefly activity was measured relative to empty vector puro control and compared to co-transfections with the empty pGL4.23 construct containing the TK minimal promoter. Co-transfections were conducted in triplicate.

3.8 Ash2L occupancy and methylation status of Pax7-induced *Zac1.*

Pax7 has recently been shown to directly interact with the MLL2 methyltransferase family member, Wdr5, and to mediate activation of its target genes through the recruitment of a chromatin remodeling complex that epigenetically modifies chromatin by methylation. H3K4 trimethylation has been established as a hallmark of active gene transcription (Margueron et al., 2005). To determine if the *Zac1* locus displays patterns of histone methylation associated with gene transcription, we conducted a series of ChIP assays on Pax7-infected C2C12 myoblasts using a trimethyl H3K4 antibody. Immunoprecipitated DNA enrichment was quantified by real-time PCR using primers that spanned both the proximal promoter region of *Zac1* (Fig. 14A), as well as the distant sites at -42 kb of *Zac1* and the novel non-canonical Pax7 binding site at +11 kb downstream of *Zac1* (Fig. 15A). Fold-enrichment was measured relative to the empty vector puro control and normalized to the *GAPDH* locus.

In addition to defining its methylation status, we also assessed Ash2L occupancy at the regions of *Zac1* bound by Pax7. Ash2L is a member of the methyltransferase complex, functioning as a structural platform on which to assemble other members of the complex and allowing for trimethylation of H3K4. We therefore conducted a series ChIP assays using the Ash2L antibody in tandem with the H3K4 ChIP experiments (Fig. 14B and 15B) to characterize the functional regulation of *Zac1* by Pax7.

Our data confirm that, with Pax7 expression in C2C12 myoblasts, the *Zac1* proximal promoter displays patterns of histone methylation and Ash2L

occupancy consistent with active gene transcription (Fig. 14A). The -500 bp region upstream of *Zac1* is highly enriched following both Ash2L and H3K4 immunoprecipitation when Pax7 is expressed. Whereas enrichment in the Ash2L ChIP is seen exclusively at the +500 site, H3K4 immunoprecipitation assays indicate trimethylation along the region proximal to the transcriptional start site of *Zac1*. The Pax7 ChIP assays suggest that Pax7 binds at a region -100 bp from *Zac1*, and its enrichment profile closely resembles that observed in the H3K4 and Ash2L ChIP assays (Fig. 8A and 14). These data parallel the model of Pax7 activation of *Myf5* (McKinnell, 2008) and suggest that Pax7 binds to the promoter of *Zac1* and activates its expression by recruiting an Ash2L-based methyltransferase complex.

Notably, there were only modest changes in H3K4 and Ash2L enrichment at the distant +11 kb region of *Zac1* bound by Pax7 (Fig. 15). Little change in Ash2L occupancy was observed (Fig. 15A). Reduced enrichment (2.5-fold at +11 kb vs. 10- to 13-fold at -100/+500) in the H3K4 ChIP assays was also observed (Fig. 15B). Therefore, while Pax7 binds to this region, it does not recruit Ash2L or instigate H3K4 modification in that immediate sequence.

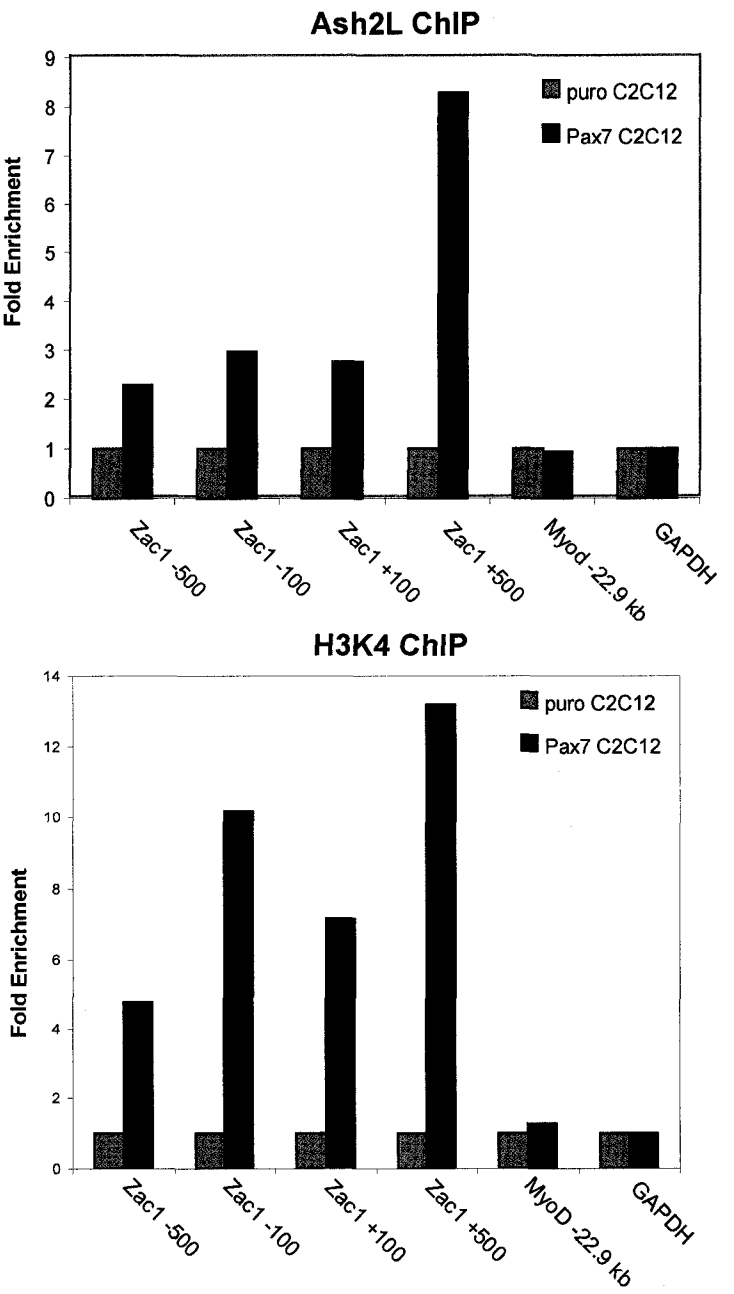


Figure 14. Functional regulation of *Zac1* promoter by Pax7. Chromatin immunoprecipitations (ChIP) were conducted in triplicate on cells ectopically expressing Pax7-Flag epitope using an Ash2L antibody (A) and an H3K4 trimethylation antibody (B). Shown are representative ChIP experiments. Enrichment levels were quantified through real-time PCR and measured relative to ChIPs conducted on C2C12 myoblasts expressing puro resistance alone. Fold enrichment was normalized to *GAPDH* levels. (A) The promoter of *Zac1* bound by Pax7 displays Ash2L occupancy at 500 bp upstream of *Zac1* (B) The promoter of *Zac1* bound by Pax7 shows patterns of H3K4 methylation that parallel Ash2L occupancy. (A, B) The region -22.9 kb from *MyoD* was tested as a negative control.

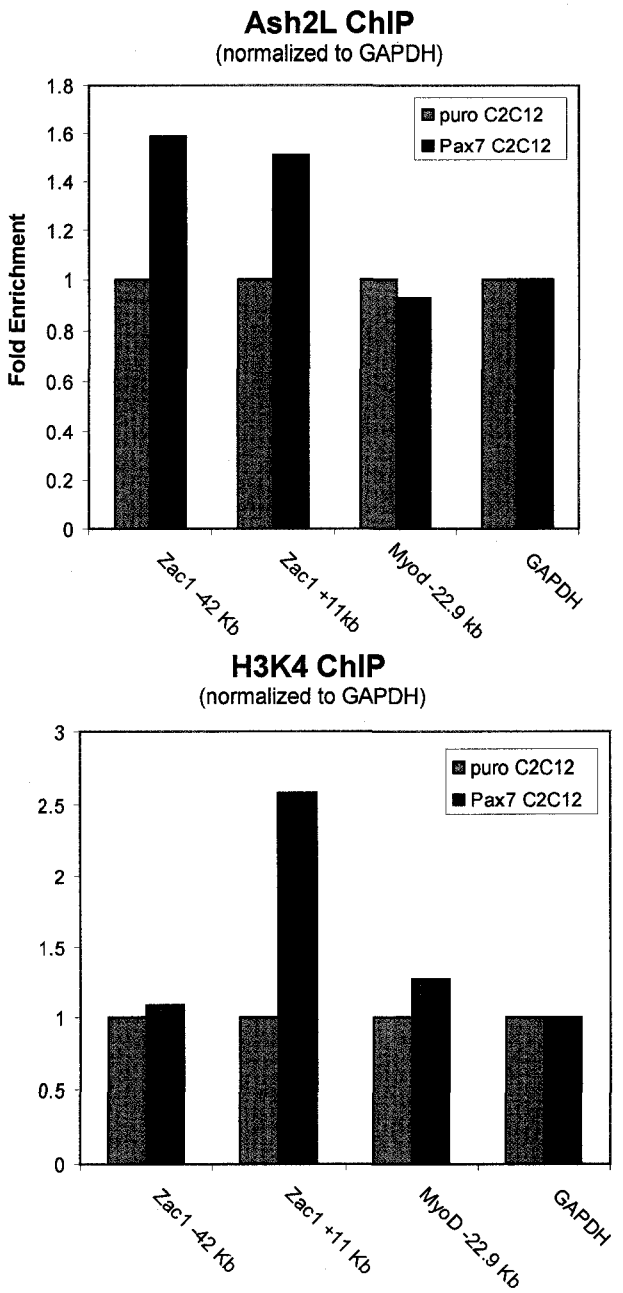


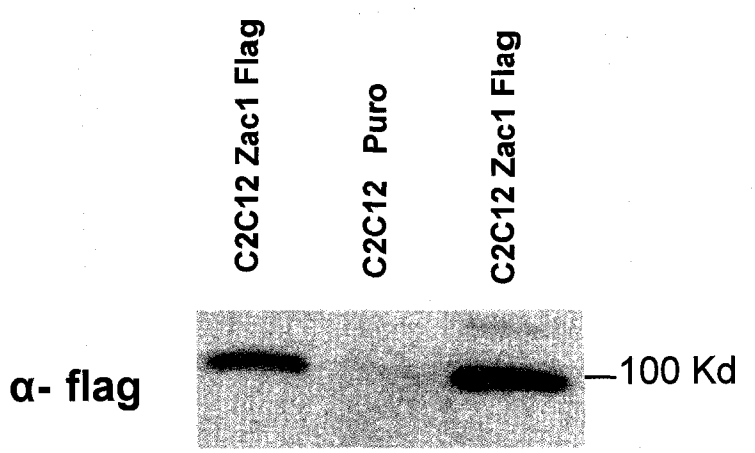
Figure 15. Regulation of +11 Kb from *Zac1* by Pax7 is not mediated by Ash2L. Chromatin immunoprecipitations (ChIP) were conducted in triplicate on cells ectopically expressing Pax7 using an Ash2L antibody (A) and an H3K4 trimethylation antibody (B). Representative enrichment levels were quantified through real-time PCR and measured relative to ChIPs conducted on C2C12 myoblasts expressing puro resistance alone. Fold enrichment was normalized to GAPDH levels. (A) No Ash2L occupancy was observed at the +11 Kb site from *Zac1* containing a novel Pax7 binding site sequence. (B) This region does not display strong patterns of H3K4 methylation showing 2.5 fold enrichment compared to the 13.5 fold H3K4 enrichment seen proximal to the *Zac1* start site.

3.9 Growth kinetics of C2C12 myoblasts expressing Zac1

The C2C12 myoblast cell line does not express any detectable *Zac1*, thereby providing a low-background environment in which to study the function of Zac1. To determine the functional consequences of *Zac1* expression in a myogenic context, we ectopically expressed Zac1 in C2C12 myoblasts and studied their growth kinetics *in vitro* (Fig. 16). The *Zac1* coding sequence was subcloned into the puromycin-selectable pBRIT retroviral expression vector and expressed with a C-terminal Flag epitope tag for tracking Zac1-Flag expression. Empty virus expressing puromycin resistance alone was used as a control. Cells were seeded at equal densities and maintained under growth conditions to construct a growth curve over the course of 4 days.

Interestingly, *Zac1* expression significantly increases the growth rate of C2C12 myoblasts. This finding contrasts with its published role as a tumor suppressor, where *in vitro* studies conducted on tumor cell lines showed that expression of Zac1 inhibits proliferation in colony forming assays and growth rate characterization (Spengler et al., 1997). *In vivo*, Zac1 expression prevented tumor formation in nude mice through G1 cell cycle arrest and induction of apoptosis (Spengler et al., 1997). Therefore, our data suggest that Zac1 has different functions in a skeletal myogenesis context.

A



B

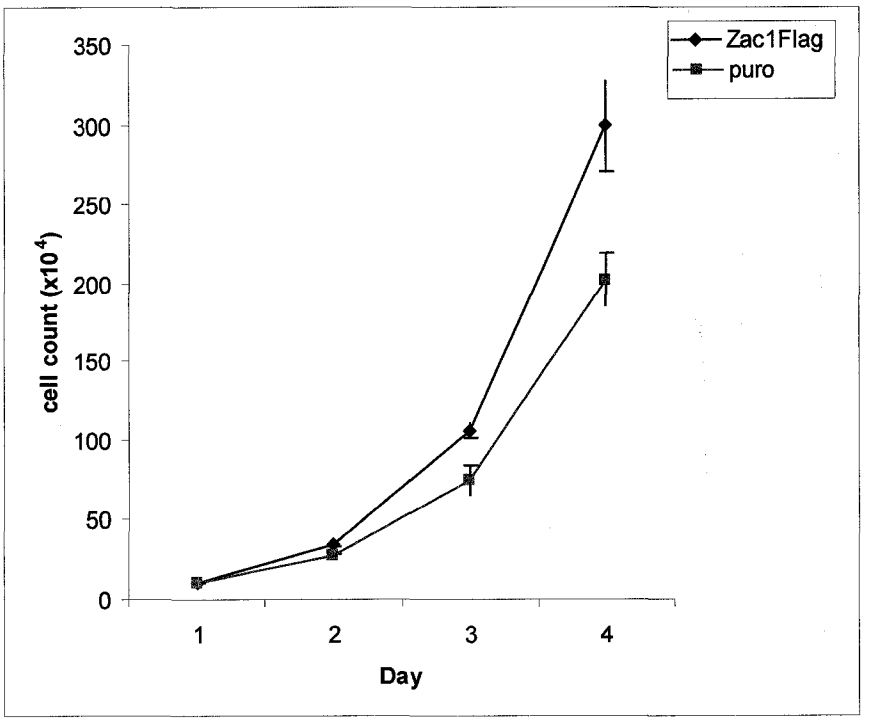


Figure 16. Ectopic expression of Zac1 increases the growth rate of C2C12 myoblasts. C2C12 myoblasts stably infected with a retrovirus expressing Zac1-flag epitope were plated at similar densities and grown under growth conditions for 4 days. Cells were trypsinized and counted every 24 hours in triplicate. (A) Expression of Zac1-Flag in C2C12 myoblasts following puromycin selection was confirmed by western blot analysis. (B) C2C12 myoblasts ectopically expressing Zac1 show a significantly higher growth rate (12.5%) compared to the empty control cells expressing puro resistance alone.

3.10 *Zac1*-null mice display compromised regenerative myogenesis

To further characterize the role played by Zac1 in myogenesis, we examined the regenerative capacity of skeletal muscle in *Zac1*-null mice. We used an acute injury model whereby cardiotoxin (CTX) was injected directly into the *tibialis anterior* (TA) leg muscle. Cardiotoxin-induced injury is advantageous in that it is highly reproducible in stimulating the degeneration (and subsequent regeneration) of the injected muscle. CTX acts by inhibiting protein kinase C activity, resulting in depolarization and contraction of muscle cells, disruption of membrane organization and cell lysis (Charge and Rudnicki, 2004). The morphological profile of CTX-induced regeneration has been well established, with myogenic differentiation and new myotube formation occuring at 5-6 days post injury and with complete restoration of the muscle observed at 3-4 weeks post-injury. Centrally located nuclei identify the newly formed fibres.

The regeneration assay was conducted on adult *Zac1-LacZ* knock-in mice expressing β-galactosidase under the control of the endogenous Zac1 promoter. *Zac1*$^{LacZ/LacZ}$ mice (homozygous *Zac1*-null) were compared to wildtype mice in the same C57Bl/6 genetic background. Myogenic progression was assessed during regeneration through histological analysis of muscle tissue using standard hematoxylin and eosin staining techniques. TA muscles were harvested at the 7-day, 12-day and 19-day times after CTX-injection to include the complete regenerative progression.

Strikingly, we observed a distinct deficit in the regenerative capacity of *Zac1*-null mice. While both *Zac1*-null and control mice had large numbers of newly regenerated fibres (distinguished by their centrally located nuclei), *Zac1*-

null mice revealed a dramatic accumulation of adipocytes (Fig. 17). Oil red O staining for lipids in frozen sections confirmed the identity of these adipocytes. This phenotype was detected at each assayed time point, with the most notable fat deposition occurring at the latest 19-day time point.

Interestingly, this compromised myogenesis is further reflected in the overall muscle mass of the TA muscle throughout the regenenerative process. Whereas injured wildtype muscle tissue continues to increase in mass relative to the contralateral resting muscle, *Zac1*-null muscle decreases in mass following injury (Fig. 18). This marked reduction in injury-induced hypertrophy is most prominent at 19 days post-CTX injection, strongly suggesting a deficiency in restoring the muscle architecture. Preliminary quantification of fibre size suggests that *Zac1*-null fibres are smaller than their wildtype controls at 19 days post-injury. Based on fibres counted in representative fields of identical dimensions and magnification across histological samples, *Zac1*-null mice have a 46.3 % decrease in fibre cross-section relative to the wildtype control at 19 days post-CTX injury (Fig. 19).

These data demonstrate that Zac1 is functionally important in the regeneration of adult skeletal muscle. Interestingly, our results parallel those seen in regenerative assays conducted on *Myf5*-null mice (Gayraud-Morel et al., 2007), where regeneration deficits have been attributed to a reduced proliferative capacity of satellite cells.

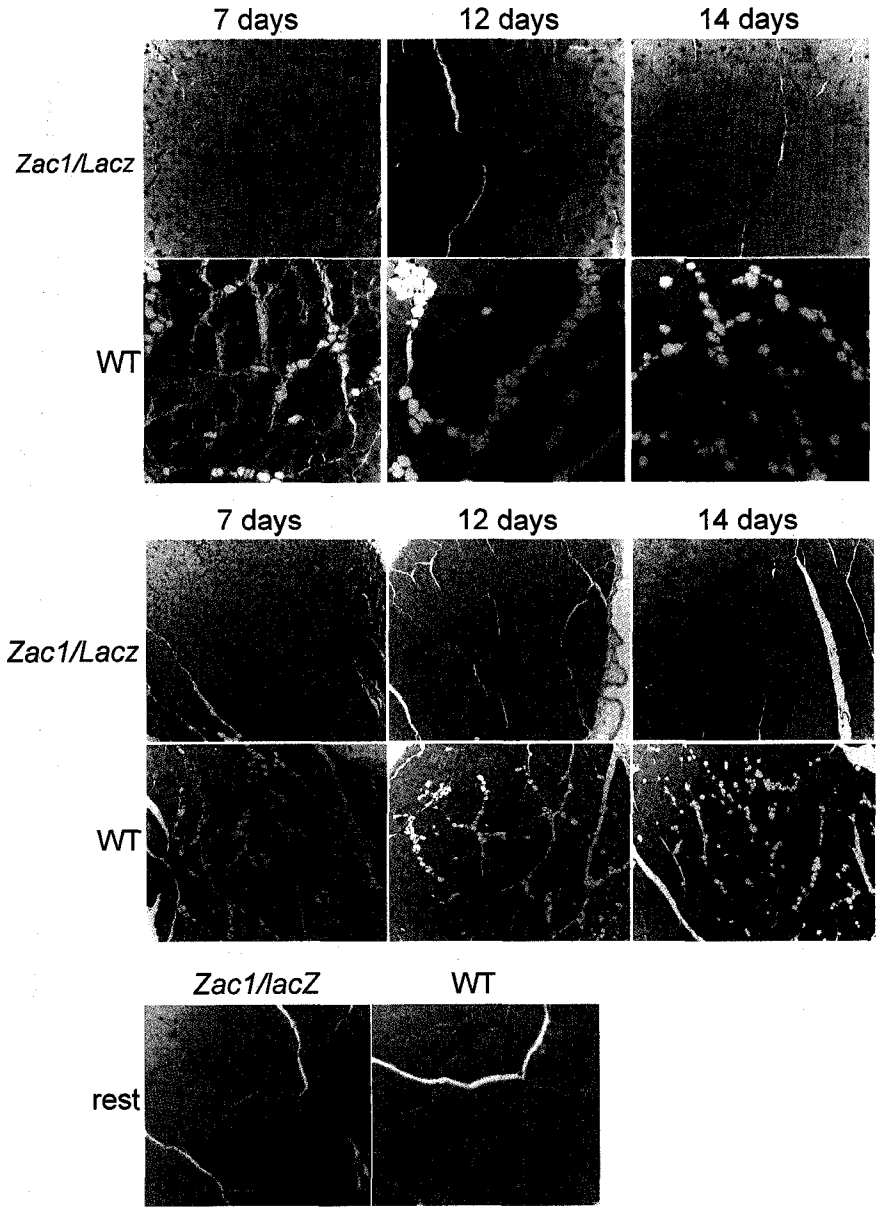


Figure 17. *Zac1*-null mice have a regeneration deficit following acute injury. Tibialis anterior of adult mice (8-12 weeks) were injured through intramuscular injection of cardiotoxin. Morphological features of regenerative progression was examined through histological analysis using H&E stain on muscle harvested at 7 days, 12 days and 19 days post injury. Control wild type mice (WT) had same genetic background (C57Bl/6). (A) Representative histological analysis at 20 X magnification shows the presence of fat cells (arrows) exclusively in *Zac1*$^{-/-}$ regenerating fibres. (B) Lower magnification reveals widespread fat accumulation throughout the regenerating TA muscle of *Zac1*$^{-/-}$ mice. (C) Resting muscle of *Zac1*$^{-/-}$ and wildtype mice have similar morphologies. Visual observation revealed a subtle decrease in *Zac1*$^{-/-}$ fibre circumference.

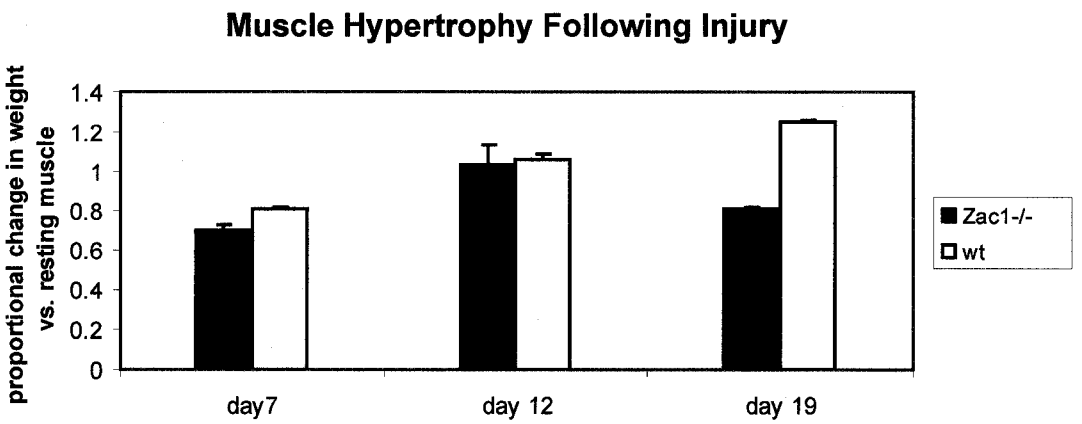


Figure 18. *Zac1*-null mice display a reduction in muscle hypertrophy after acute injury. Hypertrophy is measured as the proportion of tibialis anterior weight increase following CTX induced injury relative to the contra-lateral resting muscle. A dramatic reduction in total muscle growth is observed in the *Zac1*-null regenerating muscle 19 days post injury compared to the wild type control mice of the same genetic background. N=2

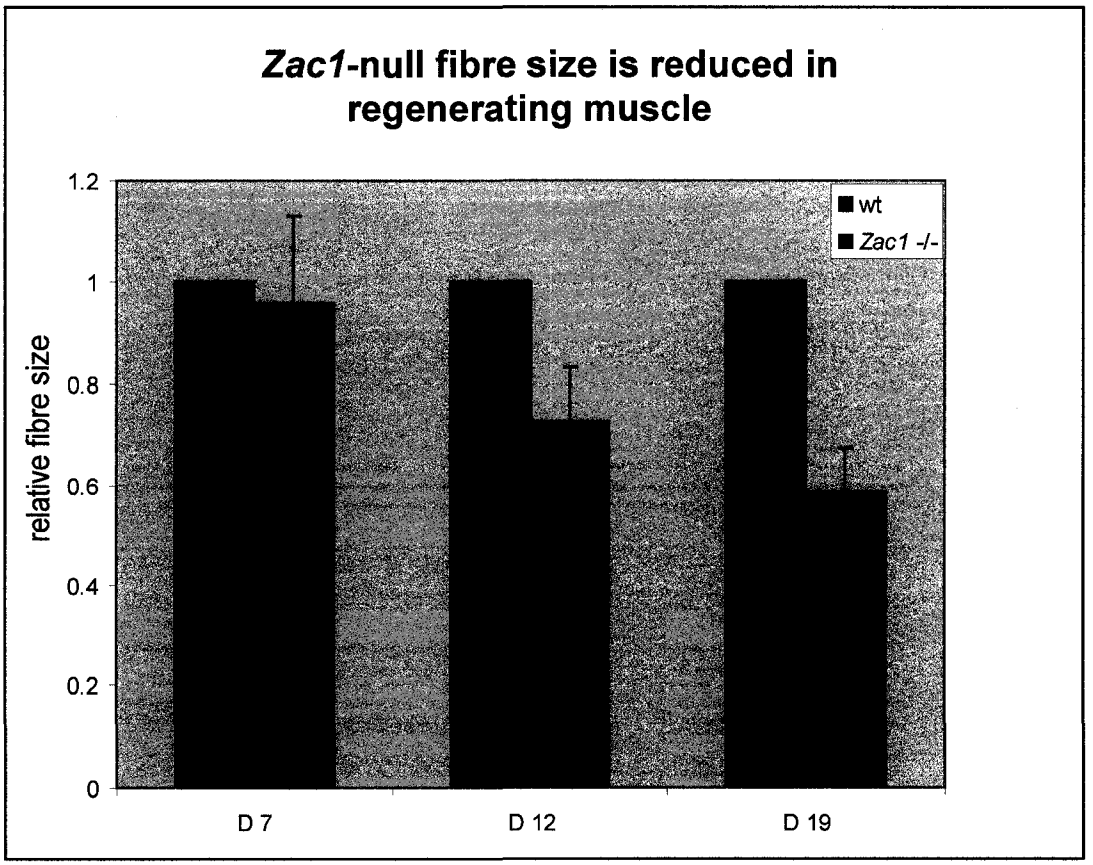


Figure 19. *Zac1*-null fibre size is decreased during regeneration. Injury was induced following CTX injection into the TA muscle of adult wildtype C57Bl/6 and *Zac1*$^{-/-}$ mice. Muscle was harvested at 7, 12 and 19 days following CTX injection. Fibre numbers were quantified in through direct enumeration on H & E stained sections in a representative field for each sample. Field size was constant between samples. Increase in fibre number corresponds to a 27.3% and a 41.3 % relative decrease in fibre size in Zac1-null mice for 12 and 19-days post injury respectively.

CHAPTER 4

Discussion

4.1 Discussion

Satellite cells are the key mediators of post-natal skeletal muscle growth and adult muscle regeneration. These unique cells have been the focus of intensive study due to their remarkable regenerative potential owing to their stem cell like properties. Indeed, it is now established that satellite cells are a heterogeneous population composed of cells that, through asymmetric divisions, give rise to both committed progenitors that undergo terminal differentiation as well as self-renewing cells that occupy and maintain the stem cell niche (Collins et al., 2005; Kuang et al., 2007). Therefore, elucidating the molecular mechanisms of satellite cell-mediated growth and repair has wide reaching implications in developing ways to treat diseased and injured muscle, as well as shedding light on the regenerative field as a whole.

Satellite cells are defined by their expression of the paired box transcription factor, Pax7 (Seale et al., 2000). The unique role played by Pax7 in the specification, expansion and survival of myogenic progenitors during post-natal growth and repair has been well documented and highlights its vital function in adult myogenesis (Zammit et al., 2006a). However, due to their heterogeneity, characterizing their cellular phenotype beyond Pax7 expression has remained challenging. The recent identification of putative Pax7 target genes through comparative genome wide microarray studies (McKinnell et al., 2008) has provided a foundation from which to investigate the biological pathways that are regulated by Pax7 and, in so doing, gaining insight into how satellite cells orchestrate muscle tissue growth and regeneration.

Zac1 (PlagL1) was found to be a strongly affected candidate Pax7 target gene. In microarray analyses, *Zac1* displayed the most striking level of activation upon ectopic expression of Pax7 in C2C12 myoblasts (McKinnell et al., 2008). Activation of *Zac1* by Pax7 was also seen in the non-myogenic 10T1/2 fibroblast cell line, suggesting that it is a *bona fide* Pax7 target (J. Ishibashi, unpublished). Moreover, expression profiling of *Zac1* in primary myoblasts (Fig. 3) provided further evidence that a direct functional relationship between Pax7 activation and *Zac1* expression exists. To further explore the hypothesis that *Zac1* is a direct Pax7 target gene, we defined the genomic regions of the *Zac1* gene bound by Pax7 and sought to characterize its functional regulation by Pax7.

To further investigate the relationship between Pax7 induction and *Zac1* activity, we studied their expression profiles in primary myoblasts. Cultured primary myoblasts are derived from isolated adult mouse satellite cells and represent committed myogenic precursor cells that are proliferative and express high levels of *Pax7*. As these cells undergo terminal differentiation, Pax7 expression is turned off, cells cease to cycle, then elongate and fuse to form multinucleated myofibres that express the differentiation markers Myogenin and MRF4 (Beauchamp et al., 2000; Rosenblatt et al., 1995; Yablonka-Reuveni and Rivera, 1994). Strikingly, we discovered that *Zac1* is also highly expressed in proliferating myoblasts, and is subsequently and dramatically down-regulated as cells terminally differentiate. *Pax7* displays a markedly similar expression profile in differentiating myoblasts (Fig. 3) consistent with a direct relationship between Pax7 expression and *Zac1* activation.

In order to further illustrate the functional relationship between Pax7 activation and *Zac1* expression, we showed activation of *Zac1* by Pax7 in primary myoblasts. Primary myoblasts ectopically expressing Pax7 exhibit a nearly 4-fold increase in *Zac1* transcripts compared to control myoblasts (Fig. 5). As both *Pax7* and *Zac1* are highly expressed in primary myoblasts under growth conditions, this relative fold-activation of *Zac1* is substantial, and provides further compelling evidence of Pax7 regulation of *Zac1*. In addition, *Myf5* expression was highly activated following Pax7 induction, as was observed in C2C12 myoblasts (McKinnell et al., 2008). *MyoD* levels fell upon ectopic expression of Pax7 in primary myoblasts, supporting previous findings that over-expression of Pax7 down-regulates MyoD (Olguin et al., 2007). The observations that Pax7 induces expression of *Zac1* in both myogenic and non-myogenic cells and both *Pax7* and *Zac1* share parallel expression profiles through primary myoblasts differentiation, strongly suggests that Pax7 activates *Zac1*.

The discovery that Zac1 is highly expressed in primary myoblasts under growth conditions enabled us to isolate and identify the presence of a unique *Zac1* splice variant (Fig. 4). To date, only two expressed isoforms have been described in the mouse: one originally isolated from adult mouse pituitary cells containing a unique 9 amino acid C-terminus (Fig.3B; Spengler et al., 1997) and the other containing an 11 amino acid insert and an alternative 36 amino acid C-terminus isolated from 17-day embryos (Fig.4A; Stallcup et al., 1999). Through RT-PCR, we identified *Zac1* transcripts in primary myoblasts that had been previously identified in embryonic tissues. While the functional significance of this

particular isoform remains undetermined, it does highlight the fact that *Zac1* contains various splice acceptor sites and has the capacity to undergo complex alternative splicing.

Given the ability of Pax7 to induce expression of *Zac1* in various cell culture models and their similar expression profiles in differentiating primary myoblasts, we hypothesized that Pax7 could act directly upstream of *Zac1* by binding to a genomic regulatory element. As no clearly defined consensus sequence for Pax7 has been established (reflecting its poorly understood mechanism of action), the possibility of identifying a direct *bona fide* Pax7 target gene presented us with the intriguing opportunity to both map a Pax7 binding site and gain insight into the molecular mechanisms through which Pax7 mediates post-natal myogenesis.

Pax3, the paralogue of Pax7, has recently been shown to bind to a 10 bp motif 57.5 kb upstream of *Myf5* during embryonic myogenesis (Bajard et al., 2006). The similar genomic and structural organization of both Pax proteins raised the possibility that Pax7 could bind this element located in the *Myf5* enhancer. Indeed, our lab recently demonstrated, through chromatin immunoprecipitation experiments conducted on C2C12 myoblasts ectopically expressing the Pax7-HIS-Flag epitope, that the enhancer region -57.5 kb upstream of *Myf5* was bound by Pax7, implying that Pax7 and Pax3 could activate target gene expression through binding of a similar consensus motif.

To establish a direct physical interaction between Pax7 and *Zac1* and to subsequently define the region of *Zac1* bound by Pax7, we conducted ChIP

experiments on C2C12 myoblasts ectopically expressing Pax7-Flag using an antibody recognizing the Flag epitope. Use of the Flag-ChIP strategy was critical because the existing monoclonal Pax7 antibody was ineffective for immunoprecipitation of native protein. We first focused on mapping out the region proximal to the transcriptional start site of *Zac1*. This region has been classically defined as the promoter for *Zac1* and it contains a DNA motif that represents a hybrid of the previously characterized Pax3 consensus sequence (Epstein et al., 1996) and the Pax3 binding site recently identified in *Myf5* (Bajard et al., 2006). Enrichment was observed at a site proximal to the transcriptional start site of *Zac1* upon induction of Pax7, suggestive of a direct interaction of Pax7 with the proximal promoter of *Zac1*. Importantly, enrichment was also established at the positive control locus -57.5 kb upstream of *Myf5* as has been previously described (McKinnell et al., 2008).

Interestingly, putative Pax7 binding regions have been identified -42 kb and +11 kb from the *Zac1* locus using a genome-wide binding site mapping approach (Punch, unpublished). A novel region 115 kb upstream *Myf5*, was similarly identified in that study as a putative Pax7 binding region. Large scale immunoprecipitation assays combined with high-throughput sequencing on the immunoenriched DNA were conducted on primary myoblasts ectopically expressing Pax7. The intriguing possibility of a unique Pax7 binding site distinct from that of Pax3 reinforced the notion that Pax3 and Pax7 have distinct roles during myogenesis. Pax3 plays a pivotal role in the patterning and development of embryonic skeletal muscle (Buckingham, 2001) whereas Pax7 expression is

critical for the growth and regenerative capacity of muscle after birth (Kuang et al., 2006; Seale et al., 2000). While the significance of Pax3 expression in a subset of satellite cells remains to be determined, it is clear that its expression is not required for satellite cell-mediated regeneration. We validated a physical interaction between Pax7 and a region located in intron 2 of *Zac1* through chromatin immunoprecipitation studies conducted in the myogenic C2C12 system, suggestive of a novel non-canonical Pax7 binding site. The -42 kb region from *Zac1* showed marginal enrichment following immunoprecipitation of Pax7-bound DNA complexes; we therefore focused on further defining the +11 kb region bound by Pax7.

Using an adapted DNA footprinting technique (Zianni et al., 2006), we identified a Pax7 binding site located in the region +11 kb downstream of *Zac1* previously identified in both ChIP and ChIPseq experiments. A region protected from DNase I digestion through in vitro binding of Pax7 to a fluorescently labeled probe spanning the +11 kb region was resolved through fluorescent automated capillary electrophoresis. The resulting electropherogram, which simulates the classical 'footprint' based on the comparative frequency and size of digested DNA, also displayed DNase I hypersensitive regions flanking the protected region when compared to the BSA control.

These hypersensitive regions are considered hallmarks of transcription factor binding and have been previously illustrated in erythrocyte locus control regions, where structural changes in chromatin as a result of protein binding render DNA more vulnerable to enzymatic digestion and create a signature

pattern of DNAse I hypersensitivity (Cho et al., 2008). Importantly, binding reactions between Pax7 and the *Myogenin* promoter, a region known not to interact with Pax7 (Zammit et al., 2006b), did not reveal a protected region (Fig. S1). Using fluorescently labeled probes in conjunction with the automated system allowed us to conduct sequencing reactions in parallel to fully define the sequence protected by Pax7 binding. Remarkably, the footprinting experiments conducted in parallel on the *Myf5* -115 kb locus revealed a footprint containing an identical 10 bp DNA element (data not shown). Moreover, this element was conserved across species, providing strong evidence that Pax7 could bind a novel DNA element that was unrelated to previously described Pax consensus sites.

We determined through electrophoretic mobility shift assays (EMSA), that Pax7 specifically binds this core motif. Addition of unlabeled oligos out-competed Pax7 binding with labeled probes, and oligos containing a mutation of the 10 base pair motif (direct purine to pyrimidine exchange) had no competitive effect on the binding reactions. Interestingly, Pax3 was also capable of binding to this *Zac1* element *in vitro*. Pax7 was shown to bind the Pax3 consensus site located -57.5 kb upstream of *Myf5* along with Pax3 in EMSA experiments (V. Punch, unpublished). Together, these data, along with the ChIP studies, raise the possibility that both Pax3 and Pax7 have both similar and distinct DNA binding domains.

To ascertain that *Zac1* is indeed a *bona fide* Pax7 target gene that plays an essential role in myogenesis, we sought to characterize its functional

regulation by Pax7. We tested the response of the novel non-canonical Pax7 binding site to Pax7/FKHR transactivation through a luciferase reporter assay system. Pax7 is known to have weak transactivation properties due to *cis* repression (Keller et al., 2004); however, naturally occurring translocations of the Pax7 DNA binding domain to the transactivation domain of FKHR result in higher levels of myogenic gene activation as reported in highly aggressive rhabdomyosarcoma tumours (Bennicelli and Barr, 1999). Therefore, we co-transfected a Pax7/FKHR fusion construct along with a luciferase expression plasmid driven by a thymidinine kinase minimal promoter that was preceded by the Pax7 binding site response element we identified in intron 2 of *Zac1*. Pax7/FKHR activated reporter expression was significantly increased by 2-fold in the presence of the 10 bp Pax7 binding site. As expected, reporter activity was increased by Pax7/FKHR transactivation relative to the empty vector control. Therefore, evidence of Pax7 binding to a specific DNA element in an intronic region of *Zac1*, combined with its functional activation, support the assertion that Zac1 is a direct Pax7 target gene during skeletal myogenesis.

Recently, our lab demonstrated that Pax7 activates transcription of target genes through interaction with a chromatin remodeling complex that modifies chromatin structure through H3K4 trimethylation (McKinnell et al., 2008). It has been widely accepted that H3K4 trimethylation is a signature epigenetic modification that indicates active gene transcription (Margueron et al., 2005). Mass spectrometry analysis of Pax7-interacting proteins identified the winged helix protein, Wdr5 (McKinnell et al., 2008). This interaction is significant in that

Wdr5, together with Ash2L and RbBP5, form a core subunit complex that interacts with the trithorax group MLL methyltransferase proteins to trimethylate H3K4 (Ruthenburg et al., 2007). Furthermore, the *Myf5* enhancer region that is bound by Pax7 was shown to display patterns of H3K4 methylation as revealed by ChIP studies (McKinnell et al., 2008). Thus, Pax7 can activate target gene transcription through chromatin remodeling via an association with a methyltransferase complex.

Based on those findings, we characterized the methylation status of the proximal promoter as well as the more distal sites identified in the genome wide ChIPseq results of *Zac1*. ChIP experiments were conducted on C2C12 myoblasts ectopically expressing Pax7 through the use of an antibody specifically recognizing tri-methylation of H3K4. The *Zac1* proximal promoter showed the hallmark trimethylation of H3K4, with the +500 bp site showing the highest enrichment, indicative of an actively transcribed gene in the presence of Pax7.

To further investigate the functional regulation of *Zac1* by Pax7, we looked at Ash2L occupancy of the various *Zac1* loci using a similar *in vitro* system. Surprisingly, enrichment of Ash2L immunoprecipitated DNA was observed exclusively at +500 bp downstream of *Zac1* transcriptional start site. Ash2L association with the *Zac1* promoter coincides with its H3K4 trimethylation status following Pax7 expression and suggests that Ash2L-mediated chromatin remodeling occurs through interaction with Pax7 in the *Zac1* promoter. Notably, there was no similar enrichment of Ash2L immunoprecipitates mapping to the

region containing the novel Pax7 binding site, +11 kb downstream of *Zac1*. Lack of Ash2L occupancy at this Pax7 binding site suggests a provocative mechanism for Pax7 action over long distance interactions.

Chromosomal 3-D structure is an important consideration when assessing the regulatory mechanisms involving protein-protein and protein-DNA interactions of distant loci. The higher structure of chromosomes is in a constant state of flux and has emerged as a key factor in regulating gene activity. The structural features of chromosomes allow them to loop, bend and connect with neighboring chromosomes to influence gene expression (Dekker et al., 2002; Simonis et al., 2007). In addition, protein-protein interactions add a level of complexity to this highly dynamic molecular setting, making it difficult to clearly dissect exact molecular mechanisms of target gene activation. These complex structural changes occurring at various chromosomal levels may in fact not be captured by straight ChIP protocols.

Such long-range interactions may reconcile our ChIP results if Pax7, while binding to the +11 kb element (Fig. 9), recruits Ash2L and allows it to bind at the *Zac1* -500 to +500 region (Fig. 14A), leading to H3K4 trimethylation of the proximal promoter (Fig. 14B). Consequent to the formaldehyde cross-linking in the ChIP procedure, the proximal promoter region would therefore be bridged through protein-protein interactions with the +11 kb region. This would manifest in enrichment, albeit indirect and decreased, of the +11 kb during Ash2L and 3Me-H3K4 ChIP (Fig. 15); as well as the presence of Pax7 at the proximal promoter region (Fig. 8A).

Recently, chromosome conformation capture (3C) techniques have been developed to better reflect temporal and spatial interactions occurring at the chromosomal level to influence gene transcription (Dekker et al., 2002; Simonis et al., 2007). Therefore, while it is clear that Pax7 binds to and activates transcription of *Zac1*, conducting 3C studies to directly test the existence of interactions between the promoter and the +11 kb region will provide conclusive information about the exact mechanism through which Pax7 interacts with protein complexes at its novel and distant binding site of *Zac1*.

Understanding the functional role played by Zac1 in post-natal myogenesis will provide further insight into how Pax7 mediates post-natal and regenerative skeletal muscle growth. Previous studies have described Zac1 as tumour suppressor due to its role in inducing cell cycle arrest and apoptosis in tumour cell lines as well as its expression profile in various carcinomas. More recently, Zac1 has been identified as a member of an imprinted gene network that controls embryonic growth and differentiation (Varrault et al., 2006). Interestingly, in this study, Zac1 was found to alter expression of *Peg3* and *Mest*, imprinted genes which were also amongst the top ten targets from our Pax7 microarray analysis (McKinnell et al., 2008). In addition to these activated imprinted genes, Zac1 was also shown to directly bind to *H19* 3' enhancers, resulting in the transactivation of the *Igf2* and *H19* promoters. Igf2 is a mitogenic growth factor that is involved in modulating muscle growth and differentiation (Pedone et al, 1994). Zac1 has also been reported to negatively regulate retinal size during development through a negative feedback loop involving TGFβII (Ma

et al., 2007). Therefore, based on previous studies, we can speculate that Zac1 may be involved in regulating proliferation and survival during myogenesis.

Our studies of Zac1 suggest a role in regulating proliferation in a myogenic context. Growth studies conducted with C2C12 myoblasts ectopically expressing Zac1 revealed enhanced proliferative activity *in vitro* (Fig.16B). While Zac1 has previously been reported to negatively affect proliferation *in vitro* in other cellular environments, its effect in C2C12 cells does reveal a context-dependent role and further studies are required to fully characterize its biological function in satellite cell-mediated myogenesis. Identification of its downstream targets in the myogenic pathway through microarray studies will provide valuable insight into how Zac1 is involved in adult muscle regeneration.

Normal muscle hypertrophy occurring as a result of injury-induced regeneration is severely compromised in *Zac1*-null mice. This regenerative deficit is reflected by a reduction in muscle mass relative to the contralateral resting muscle. Defects in muscle growth were particularly observed in the latter stages of the regenerative process. Similarly, a reduction in fibre size was detected in *Zac1*-null regenerating muscle as measured by fibre numbers across a defined field size between samples (Fig. 19). Quantification of fibre diameter across a larger sample size is required to fully document this potential myopathy; however, preliminary data indicate that there is a reduction in fibre size as regeneration proceeds in *Zac1*-null mice.

Intriguingly, *in vivo* regeneration assays conducted on *Zac1*-null mice suggest a critical role for Zac1 in regulating adult muscle regenerative capacity.

Histological analysis of regenerating *Zac1*-null muscle tissue reveal a striking accumulation of adipocytes when compared to wildtype control muscle in the same genetic background. This phenotype was most pronounced at three weeks post-injury (Fig. 17). Resting muscle of both *Zac1*-null and wildtype mice had an absence of these adipocytes (Fig. 17).

Fat deposition has been previously noted in regeneration studies conducted on *Myf5*-null muscle tissue, in *Pax7*-null mice and in Duchenne muscular dystrophy patients where delayed myogenic differentiation was also identified (Gayraud-Morel et al., 2007). Whether fat accumulation due to a regeneration deficit is a result of adipocyte invasion or due to a conversion of myogenic to adipogenic cell fates during compromised myogenic regeneration is unclear. Recent studies conducted by the Spiegelman group have shown that adipocytes originate from Myf5 expressing precursor cells and that these precursors can enter into either a myogenic lineage or an adipogenic lineage through the expression of the transcription factor, PRDM16 (Seale et al., 2008). This finding raises the possibility that adipocytes have arisen from myogenic precursor cells in the absence of *Zac1* expression.

Given the role of Pax7 in the expansion and survival of satellite cells, and the recently established function of Myf5 in the generation of transit amplifying cells, it is possible that the increased adiposity in muscle tissue is correlated to proliferation defects and hence a reduced capacity to regenerate. Further studies to characterize the underlying mechanism of adipocyte accumulation are

required to gain a clearer understanding of exactly how Zac1, and therefore Pax7, influences satellite cell behavior during myogenesis.

Elucidating the molecular mechanisms that underscore satellite cell activity is critical in our understanding of adult skeletal muscle regeneration. Furthermore, satellite cells not only represent myogenic progenitors with a remarkable capacity for skeletal muscle regeneration, but also provide a model system in which to study stem cell biology. Pax7 is a defining feature of satellite cell function as its expression is essential in the commitment, expansion and survival of myogenic progenitors. Its role in myogenesis is unique and vital for post-natal growth and regeneration of these cells. In identifying its downstream target genes, we are acquiring a clearer picture of the exact mechanisms through which Pax7 regulates gene activation. This research has identified a direct Pax7 target gene, *Zac1*, through expression analysis, *in vivo* immunoprecipitation studies and *in vitro* binding assays. Surprisingly, we have characterized a novel non-canonical Pax7 binding site whose sequence is unrelated to any other Pax consensus sites, including its paralogue, Pax3. This finding is of particular interest since it suggests a unique mechanism for Pax7 activity as distinct from Pax3. This notion supports previous data describing the distinct roles played by Pax3 and Pax7 in myogenesis (Keller et al., 2004; Kuang et al., 2006; Relaix et al., 2004). We have also established the functional regulation of *Zac1* by Pax7 through characterization of its methylation status and Ash2L occupancy at a region physically bound by Pax7. Finally, we have shed some light on a possible function of Zac1 in regulating proliferation of myogenic cells, thus setting the

stage for future in-depth studies for Zac1 involvement in myogenesis. These findings are important to defining the cellular phenotype of satellite cells and advance our knowledge of the regenerative process in adult tissues.

References

Abdollahi, A. (2007). LOT1 (ZAC1/PLAGL1) and its family members: Mechanisms and functions. J Cell Physiol *210*, 16-25.

Anastasi, S., Giordano, S., Sthandier, O., Gambarotta, G., Maione, R., Comoglio, P., and Amati, P. (1997). A natural hepatocyte growth factor/scatter factor autocrine loop in myoblast cells and the effect of the constitutive Met kinase activation on myogenic differentiation. J Cell Biol *137*, 1057-1068.

Armand, O., Boutineau, A. M., Mauger, A., Pautou, M. P., and Kieny, M. (1983). Origin of satellite cells in avian skeletal muscles. Arch Anat Microsc Morphol Exp *72*, 163-181.

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P., and Buckingham, M. E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. Genes Dev *20*, 2450-2464.

Baroffio, A., Hamann, M., Bernheim, L., Bochaton-Piallat, M. L., Gabbiani, G., and Bader, C. R. (1996). Identification of self-renewing myoblasts in the progeny of single human muscle satellite cells. Differentiation *60*, 47-57.

Beauchamp, J. R., Heslop, L., Yu, D. S., Tajbakhsh, S., Kelly, R. G., Wernig, A., Buckingham, M. E., Partridge, T. A., and Zammit, P. S. (2000). Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. J Cell Biol *151*, 1221-1234.

Beauchamp, J. R., Morgan, J. E., Pagel, C. N., and Partridge, T. A. (1999). Dynamics of myoblast transplantation reveal a discrete minority of precursors with stem cell-like properties as the myogenic source. J Cell Biol *144*, 1113-1122.

Ben-Yair, R., and Kalcheim, C. (2005). Lineage analysis of the avian dermomyotome sheet reveals the existence of single cells with both dermal and muscle progenitor fates. Development *132*, 689-701.

Bennicelli, J. L., and Barr, F. G. (1999). Genetics and the biologic basis of sarcomas. Curr Opin Oncol *11*, 267-274.

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

Bilanges, B., Varrault, A., Mazumdar, A., Pantaloni, C., Hoffmann, A., Bockaert, J., Spengler, D., and Journot, L. (2001). Alternative splicing of the imprinted candidate tumor suppressor gene ZAC regulates its antiproliferative and DNA binding activities. Oncogene *20*, 1246-1253.

Bischoff, R. (1994). The satellite cell and muscle regeneration., In Myogenesis (New York: McGraw-Hill), pp. 97-118.

Blanco-Bose, W. E., Yao, C. C., Kramer, R. H., and Blau, H. M. (2001). Purification of mouse primary myoblasts based on alpha 7 integrin expression. Exp Cell Res *265*, 212-220.

Bober, E., Franz, T., Arnold, H. H., Gruss, P., and Tremblay, P. (1994). Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. Development *120*, 603-612.

Braun, T., Rudnicki, M. A., Arnold, H. H., and Jaenisch, R. (1992). Targeted inactivation of the muscle regulatory gene Myf-5 results in abnormal rib development and perinatal death. Cell *71*, 369-382.

Buckingham, M. (2001). Skeletal muscle formation in vertebrates. Curr Opin Genet Dev *11*, 440-448.

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003). The formation of skeletal muscle: from somite to limb. J Anat *202*, 59-68.

Campbell, J. S., Wenderoth, M. P., Hauschka, S. D., and Krebs, E. G. (1995). Differential activation of mitogen-activated protein kinase in response to basic fibroblast growth factor in skeletal muscle cells. Proc Natl Acad Sci U S A *92*, 870-874.

Chalepakis, G., and Gruss, P. (1995). Identification of DNA recognition sequences for the Pax3 paired domain. Gene *162*, 267-270.

Charge, S. B., and Rudnicki, M. A. (2004). Cellular and molecular regulation of muscle regeneration. Physiol Rev *84*, 209-238.

Cho, Y., Song, S. H., Lee, J. J., Choi, N., Kim, C. G., Dean, A., and Kim, A. (2008). The role of transcriptional activator GATA-1 at human {beta}-globin HS2. Nucleic Acids Res.

Christ, B., and Ordahl, C. P. (1995). Early stages of chick somite development. Anat Embryol (Berl) *191*, 381-396.

Collins, C. A., Olsen, I., Zammit, P. S., Heslop, L., Petrie, A., Partridge, T. A., and Morgan, J. E. (2005). Stem cell function, self-renewal, and behavioral heterogeneity of cells from the adult muscle satellite cell niche. Cell *122*, 289-301.

Cornelison, D. D., and Wold, B. J. (1997). Single-cell analysis of regulatory gene expression in quiescent and activated mouse skeletal muscle satellite cells. Dev Biol *191*, 270-283.

Cossu, G., Tajbakhsh, S., and Buckingham, M. (1996). How is myogenesis initiated in the embryo? Trends Genet *12*, 218-223.

De Angelis, L., Berghella, L., Coletta, M., Lattanzi, L., Zanchi, M., Cusella-De Angelis, M. G., Ponzetto, C., and Cossu, G. (1999). Skeletal myogenic progenitors originating from embryonic dorsal aorta coexpress endothelial and myogenic markers and contribute to postnatal muscle growth and regeneration. J Cell Biol *147*, 869-878.

Decary, S., Mouly, V., Hamida, C. B., Sautet, A., Barbet, J. P., and Butler-Browne, G. S. (1997). Replicative potential and telomere length in human skeletal muscle: implications for satellite cell-mediated gene therapy. Hum Gene Ther *8*, 1429-1438.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Dou, Y., Milne, T. A., Ruthenburg, A. J., Lee, S., Lee, J. W., Verdine, G. L., Allis, C. D., and Roeder, R. G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol *13*, 713-719.

Epstein, J., Cai, J., Glaser, T., Jepeal, L., and Maas, R. (1994). Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. J Biol Chem *269*, 8355-8361.

Epstein, J. A., Shapiro, D. N., Cheng, J., Lam, P. Y., and Maas, R. L. (1996). Pax3 modulates expression of the c-Met receptor during limb muscle development. Proc Natl Acad Sci U S A *93*, 4213-4218.

Fan, T., Hagan, J. P., Kozlov, S. V., Stewart, C. L., and Muegge, K. (2005). Lsh controls silencing of the imprinted Cdkn1c gene. Development *132*, 635-644.

Fielding, R. A., Manfredi, T. J., Ding, W., Fiatarone, M. A., Evans, W. J., and Cannon, J. G. (1993). Acute phase response in exercise. III. Neutrophil and IL-1 beta accumulation in skeletal muscle. Am J Physiol *265*, R166-172.

Fuchtbauer, E. M., and Westphal, H. (1992). MyoD and myogenin are coexpressed in regenerating skeletal muscle of the mouse. Dev Dyn *193*, 34-39.

Garry, D. J., Yang, Q., Bassel-Duby, R., and Williams, R. S. (1997). Persistent expression of MNF identifies myogenic stem cells in postnatal muscles. Dev Biol *188*, 280-294.

Gayraud-Morel, B., Chretien, F., Flamant, P., Gomes, D., Zammit, P. S., and Tajbakhsh, S. (2007). A role for the myogenic determination gene Myf5 in adult regenerative myogenesis. Dev Biol *312*, 13-28.

Goulding, M., Lumsden, A., and Paquette, A. J. (1994). Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. Development *120*, 957-971.

Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R., and Gruss, P. (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. Embo J *10*, 1135-1147.

Gros, J., Manceau, M., Thome, V., and Marcelle, C. (2005). A common somitic origin for embryonic muscle progenitors and satellite cells. Nature.

Gros, J., Scaal, M., and Marcelle, C. (2004). A two-step mechanism for myotome formation in chick. Dev Cell *6*, 875-882.

Grounds, M. D., Garrett, K. L., Lai, M. C., Wright, W. E., and Beilharz, M. W. (1992). Identification of skeletal muscle precursor cells in vivo by use of MyoD1 and myogenin probes. Cell Tissue Res *267*, 99-104.

Grounds, M. D., and Yablonka-Reuveni, Z. (1993). Molecular and cell biology of skeletal muscle regeneration. Mol Cell Biol Hum Dis Ser *3*, 210-256.

Hawke, T. J., and Garry, D. J. (2001). Myogenic satellite cells: physiology to molecular biology. J Appl Physiol *91*, 534-551.

Henikoff, S. (2008). Nucleosome destabilization in the epigenetic regulation of gene expression. Nat Rev Genet *9*, 15-26.

Hensen, K., Van Valckenborgh, I. C., Kas, K., Van de Ven, W. J., and Voz, M. L. (2002). The tumorigenic diversity of the three PLAG family members is associated with different DNA binding capacities. Cancer Res *62*, 1510-1517.

Hoffmann, A., Ciani, E., Boeckardt, J., Holsboer, F., Journot, L., and Spengler, D. (2003). Transcriptional activities of the zinc finger protein Zac are differentially controlled by DNA binding. Mol Cell Biol *23*, 988-1003.

Huang, S. M., Schonthal, A. H., and Stallcup, M. R. (2001). Enhancement of p53-dependent gene activation by the transcriptional coactivator Zac1. Oncogene *20*, 2134-2143.

Huang, S. M., and Stallcup, M. R. (2000). Mouse Zac1, a transcriptional coactivator and repressor for nuclear receptors. Mol Cell Biol *20*, 1855-1867.

Huh, M. S., Parker, M. H., Scime, A., Parks, R., and Rudnicki, M. A. (2004). Rb is required for progression through myogenic differentiation but not maintenance of terminal differentiation. J Cell Biol *166*, 865-876.

Huxley, H. E. (2002). The mechanism of muscular contraction. Science 164:1356-1366, 1969. Clin Orthop Relat Res, S6-17.

Iberg, A. N., Espejo, A., Cheng, D., Kim, D., Michaud-Levesque, J., Richard, S., and Bedford, M. T. (2008). Arginine methylation of the histone H3 tail impedes effector binding. J Biol Chem *283*, 3006-3010.

Ishibashi, J., Perry, R. L., Asakura, A., and Rudnicki, M. A. (2005). MyoD induces myogenic differentiation through cooperation of its NH2- and COOH-terminal regions. J Cell Biol *171*, 471-482.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomes, D., Rocancourt, D., Buckingham, M., Shinin, V., and Tajbakhsh, S. (2004). Mrf4 determines skeletal muscle identity in Myf5:Myod double-mutant mice. Nature *431*, 466-471.

Kassar-Duchossoy, L., Giacone, E., Gayraud-Morel, B., Jory, A., Gomes, D., and Tajbakhsh, S. (2005). Pax3/Pax7 mark a novel population of primitive myogenic cells during development. Genes Dev *19*, 1426-1431.

Keller, C., Hansen, M. S., Coffin, C. M., and Capecchi, M. R. (2004). Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. Genes Dev *18*, 2608-2613.

Kerppola, T. K., and Curran, T. (1997). The transcription activation domains of Fos and Jun induce DNA bending through electrostatic interactions. Embo J *16*, 2907-2916.

Kiefer, J. C. (2007). Epigenetics in development. Dev Dyn *236*, 1144-1156.

Kuang, S., Charge, S. B., Seale, P., Huh, M., and Rudnicki, M. A. (2006). Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. J Cell Biol *172*, 103-113.

Kuang, S., Gillespie, M. A., and Rudnicki, M. A. (2008). Niche regulation of muscle satellite cell self-renewal and differentiation. Cell Stem Cell *2*, 22-31.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M. A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell *129*, 999-1010.

Le Grand, F., and Rudnicki, M. A. (2007). Skeletal muscle satellite cells and adult myogenesis. Curr Opin Cell Biol *19*, 628-633.

Ma, L., Cantrup, R., Varrault, A., Colak, D., Klenin, N., Gotz, M., McFarlane, S., Journot, L., and Schuurmans, C. (2007). Zac1 functions through TGFbetaII to negatively regulate cell number in the developing retina. Neural Develop *2*, 11.

Maniatis, T., Sambrook J., Fritsch, E.F. (1989). Molecular Cloning: A Laboratory Manual
Vol I, 2nd ed. edn: Cold Spring Harbor Laboratory Press
).

Mansouri, A., Stoykova, A., Torres, M., and Gruss, P. (1996). Dysgenesis of cephalic neural crest derivatives in Pax7-/- mutant mice. Development *122*, 831-838.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? Curr Opin Genet Dev *15*, 163-176.

Mauro, A. (1961). Satellite cell of skeletal muscle fibers. J Biophys Biochem Cytol *9*, 493-495.

McGeachie, J. K. (1985). The fate of proliferating cells in skeletal muscle after denervation or tenotomy: an autoradiographic study. Neuroscience *15*, 499-506.

McGeachie, J. K., and Grounds, M. D. (1989). The onset of myogenesis in denervated mouse skeletal muscle regenerating after injury. Neuroscience *28*, 509-514.

McKinnell, I. W., Ishibashi, J., Le Grand, F., Punch, V. G., Addicks, G. C., Greenblatt, J. F., Dilworth, F. J., and Rudnicki, M. A. (2008). Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. Nat Cell Biol *10*, 77-84.

McKinsey, T. A., Zhang, C. L., and Olson, E. N. (2002). Signaling chromatin to make muscle. Curr Opin Cell Biol *14*, 763-772.

Megeney, L. A., Kablar, B., Garrett, K., Anderson, J. E., and Rudnicki, M. A. (1996). MyoD is required for myogenic stem cell function in adult skeletal muscle. Genes Dev *10*, 1173-1183.

Miller, K. J., Thaloor, D., Matteson, S., and Pavlath, G. K. (2000). Hepatocyte growth factor affects satellite cell activation and differentiation in regenerating skeletal muscle. Am J Physiol Cell Physiol *278*, C174-181.

Milne, T. A., Dou, Y., Martin, M. E., Brock, H. W., Roeder, R. G., and Hess, J. L. (2005). MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A *102*, 14765-14770.

Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. (2005). Direct isolation of satellite cells for skeletal muscle regeneration. Science *309*, 2064-2067.

Moss, F. P., and Leblond, C. P. (1971). Satellite cells as the source of nuclei in muscles of growing rats. Anat Rec *170*, 421-435.

Nagata, Y., Partridge, T. A., Matsuda, R., and Zammit, P. S. (2006). Entry of muscle satellite cells into the cell cycle requires sphingolipid signaling. J Cell Biol *174*, 245-253.

Olguin, H. C., and Olwin, B. B. (2004). Pax-7 up-regulation inhibits myogenesis and cell cycle progression in satellite cells: a potential mechanism for self-renewal. Dev Biol *275*, 375-388.

Olguin, H. C., Yang, Z., Tapscott, S. J., and Olwin, B. B. (2007). Reciprocal inhibition between Pax7 and muscle regulatory factors modulates myogenic cell fate determination. J Cell Biol *177*, 769-779.

Orimo, S., Hiyamuta, E., Arahata, K., and Sugita, H. (1991). Analysis of inflammatory cells and complement C3 in bupivacaine-induced myonecrosis. Muscle Nerve *14*, 515-520.

Pagotto, U., Arzberger, T., Ciani, E., Lezoualc'h, F., Pilon, C., Journot, L., Spengler, D., and Stalla, G. K. (1999). Inhibition of Zac1, a new gene differentially expressed in the anterior pituitary, increases cell proliferation. Endocrinology *140*, 987-996.

Park, K. Y., and Pfeifer, K. (2003). Epigenetic interplay. Nat Genet *34*, 126-128.
Parker, M. H., Seale, P., and Rudnicki, M. A. (2003). Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat Rev Genet *4*, 497-507.

Pedone, P. V., Tirabosco, R., Cavazzana, A. O., Ungaro, P., Basso, G., Luksch, R., Carli, M., Bruni, C. B., Frunzio, R., and Riccio, A. (1994). Mono- and bi-allelic expression of insulin-like growth factor II gene in human muscle tumors. Hum Mol Genet *3*, 1117-1121.

Piras, G., El Kharroubi, A., Kozlov, S., Escalante-Alcalde, D., Hernandez, L., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Stewart, C. L. (2000). Zac1 (Lot1), a potential tumor suppressor gene, and the gene for epsilon-sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. Mol Cell Biol *20*, 3308-3315.

Puri, P. L., and Sartorelli, V. (2000). Regulation of muscle regulatory factors by DNA-binding, interacting proteins, and post-transcriptional modifications. J Cell Physiol *185*, 155-173.

Rando, T. A., and Blau, H. M. (1994). Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. J Cell Biol *125*, 1275-1287.

Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A., and Buckingham, M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. J Cell Biol *172*, 91-102.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. Genes Dev *18*, 1088-1105.

Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. Nature.

Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. Annu Rev Genet *38*, 413-443.

Rosenblatt, J. D., Lunt, A. I., Parry, D. J., and Partridge, T. A. (1995). Culturing satellite cells from living single muscle fiber explants. In Vitro Cell Dev Biol Anim *31*, 773-779.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol *132*, 365-386.

Rudnicki, M. A., Braun, T., Hinuma, S., and Jaenisch, R. (1992). Inactivation of MyoD in mice leads to up-regulation of the myogenic HLH gene Myf-5 and results in apparently normal muscle development. Cell *71*, 383-390.

Ruthenburg, A. J., Allis, C. D., and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell *25*, 15-30.

Sabourin, L. A., Girgis-Gabardo, A., Seale, P., Asakura, A., and Rudnicki, M. A. (1999). Reduced differentiation potential of primary MyoD-/- myogenic cells derived from adult skeletal muscle. J Cell Biol *144*, 631-643.

Schienda, J., Engleka, K. A., Jun, S., Hansen, M. S., Epstein, J. A., Tabin, C. J., Kunkel, L. M., and Kardon, G. (2006). Somitic origin of limb muscle satellite and side population cells. Proc Natl Acad Sci U S A *103*, 945-950.

Schmalbruch, H., and Lewis, D. M. (2000). Dynamics of nuclei of muscle fibers and connective tissue cells in normal and denervated rat muscles. Muscle Nerve *23*, 617-626.

Schultz, E. (1976). Fine structure of satellite cells in growing skeletal muscle. Am J Anat *147*, 49-70.

Schultz, E. (1989). Satellite cell behavior during skeletal muscle growth and regeneration. Med Sci Sports Exerc *21*, S181-186.

Schultz, E. (1996). Satellite cell proliferative compartments in growing skeletal muscles. Dev Biol *175*, 84-94.

Schultz, E., and Jaryszak, D. L. (1985). Effects of skeletal muscle regeneration on the proliferation potential of satellite cells. Mech Ageing Dev *30*, 63-72.

Schultz, E., and Lipton, B. H. (1982). Skeletal muscle satellite cells: changes in proliferation potential as a function of age. Mech Ageing Dev *20*, 377-383.

Seale, P., Bjork, B., Yang, W., Kajimura, S., Chin, S., Kuang, S., Scime, A., Devarakonda, S., Conroe, H. M., Erdjument-Bromage, H., *et al.* (2008). PRDM16 controls a brown fat/skeletal muscle switch. Nature *454*, 961-967.

Seale, P., and Rudnicki, M. A. (2000). A new look at the origin, function, and "stem-cell" status of muscle satellite cells. Dev Biol *218*, 115-124.

Seale, P., Sabourin, L. A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M. A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell *102*, 777-786.

Simonis, M., Kooren, J., and de Laat, W. (2007). An evaluation of 3C-based methods to capture DNA interactions. Nat Methods *4*, 895-901.

Smith, C. K., 2nd, Janney, M. J., and Allen, R. E. (1994). Temporal expression of myogenic regulatory genes during activation, proliferation, and differentiation of rat skeletal muscle satellite cells. J Cell Physiol *159*, 379-385.

Smith, R. J., Arnaud, P., Konfortova, G., Dean, W. L., Beechey, C. V., and Kelsey, G. (2002). The mouse Zac1 locus: basis for imprinting and comparison with human ZAC. Gene *292*, 101-112.

Spengler, D., Villalba, M., Hoffmann, A., Pantaloni, C., Houssami, S., Bockaert, J., and Journot, L. (1997). Regulation of apoptosis and cell cycle arrest by Zac1, a novel zinc finger protein expressed in the pituitary gland and the brain. Embo J *16*, 2814-2825.

Steward, M. M., Lee, J. S., O'Donovan, A., Wyatt, M., Bernstein, B. E., and Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat Struct Mol Biol *13*, 852-854.

Stuart, E. T., Kioussi, C., and Gruss, P. (1994). Mammalian Pax genes. Annu Rev Genet *28*, 219-236.

Tajbakhsh, S., Rocancourt, D., Cossu, G., and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. Cell *89*, 127-138.

Tatsumi, R., Anderson, J. E., Nevoret, C. J., Halevy, O., and Allen, R. E. (1998). HGF/SF is present in normal adult skeletal muscle and is capable of activating satellite cells. Dev Biol *194*, 114-128.

Thomas, M., Langley, B., Berry, C., Sharma, M., Kirk, S., Bass, J., and Kambadur, R. (2000). Myostatin, a negative regulator of muscle growth, functions by inhibiting myoblast proliferation. J Biol Chem *275*, 40235-40243.

Tidball, J. G. (2005). Inflammatory processes in muscle injury and repair. Am J Physiol Regul Integr Comp Physiol *288*, R345-353.

Tremblay, P., and Gruss, P. (1994). Pax: genes for mice and men. Pharmacol Ther *61*, 205-226.

Ustanina, S., Carvajal, J., Rigby, P., and Braun, T. (2007). The myogenic factor Myf5 supports efficient skeletal muscle regeneration by enabling transient myoblast amplification. Stem Cells *25*, 2006-2016.

Valente, T., Junyent, F., and Auladell, C. (2005). Zac1 is expressed in progenitor/stem cells of the neuroectoderm and mesoderm during embryogenesis: differential phenotype of the Zac1-expressing cells during development. Dev Dyn *233*, 667-679.

Varrault, A., Ciani, E., Apiou, F., Bilanges, B., Hoffmann, A., Pantaloni, C., Bockaert, J., Spengler, D., and Journot, L. (1998). hZAC encodes a zinc finger protein with antiproliferative properties and maps to a chromosomal region frequently lost in cancer. Proc Natl Acad Sci U S A *95*, 8835-8840.

Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Aknin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., *et al.* (2006). Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. Dev Cell *11*, 711-722.

Vogan, K. J., Underhill, D. A., and Gros, P. (1996). An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. Mol Cell Biol *16*, 6677-6686.

Wozniak, A. C., and Anderson, J. E. (2007). Nitric oxide-dependence of satellite stem cell activation and quiescence on normal skeletal muscle fibers. Dev Dyn *236*, 240-250.

Yablonka-Reuveni, Z., and Rivera, A. J. (1994). Temporal expression of regulatory and structural muscle proteins during myogenesis of satellite cells on isolated adult rat fibers. Dev Biol *164*, 588-603.

Yoshida, N., Yoshida, S., Koishi, K., Masuda, K., and Nabeshima, Y. (1998). Cell heterogeneity upon myogenic differentiation: down-regulation of MyoD and Myf-5 generates 'reserve cells'. J Cell Sci *111 (Pt 6)*, 769-779.

Zammit, P. S., Partridge, T. A., and Yablonka-Reuveni, Z. (2006a). The skeletal muscle satellite cell: the stem cell that came in from the cold. J Histochem Cytochem *54*, 1177-1191.

Zammit, P. S., Relaix, F., Nagata, Y., Ruiz, A. P., Collins, C. A., Partridge, T. A., and Beauchamp, J. R. (2006b). Pax7 and myogenic progression in skeletal muscle satellite cells. J Cell Sci *119*, 1824-1832.

Zhang, C., Xuan, Z., Otto, S., Hover, J. R., McCorkle, S. R., Mandel, G., and Zhang, M. Q. (2006). A clustering property of highly-degenerate transcription factor binding sites in the mammalian genome. Nucleic Acids Res *34*, 2238-2246.

Zianni, M., Tessanne, K., Merighi, M., Laguna, R., and Tabita, F. R. (2006). Identification of the DNA bases of a DNase I footprint by the use of dye primer sequencing on an automated capillary DNA analysis instrument. J Biomol Tech *17*, 103-113.

Ziman, M. R., Rodger, J., Chen, P., Papadimitriou, J. M., Dunlop, S. A., and Beazley, L. D. (2001). Pax genes in development and maturation of the vertebrate visual system: implications for optic nerve regeneration. Histol Histopathol *16*, 239-249.

Appendix A

Table S1. List of primers for cloning and expression profile studies

name	sequence (5'-3')	location
Cloning of Zac1		
Zac1 5' F EcoRI	GAATTCGCCGCCACCATGGCTCCATTCCGCT	exon 7
Zac1 R mid 3' R	CATGGGTTGAACTTGAGAGG	exon 8
Zac1 mid 5' F	TCCTCCCCCAATTATTCTTC	exon 8
Zac1 R no stop xho I R	CTCGAGAACTGTCCATTTCTTATAGACGAGT	exon 9
isoformR2 no stop (xhol)	CTCGAGTCTAAATGCGTGATGGAAATG	exon 9
Real-time PCR		
Zac1 F	GTGGCAAAGCCTTCGTCT	exon 7
Zac1 R	CACACCCCACAGGTGAGA	exon 8
GAPDH F	ATGCCAGTGAGCTTCCCGTC	exon 1
GAPDH R	CATCACCATCTTCCAGGAGC	exon 1
Myf5 F	TGAAGGATGGACATGACGGACG	exon 1
Myf5 R	TTGTGTGCTCCGAAGGCTGCTA	exon 1
MyoD F	TACCCAAGGTGGAGATCCTG	exon 1
MyoD R	CATCATGCCATCAGAGCAGT	exon 1
Pax7 F	CTGGATGAGGGCTCAGATGT	exon 3
Pax7 R	GGTTAGCTCCTGCCTGCTTA	exon 4
Beta-actin F	AGAAGATCTGGCACCACACC	exon 3
Beta-actin R	CCATCTCCTGCTCGAAGTCT	exon 4

Table S2. List of primers for chromosome mapping and functional regulation studies

name	sequence (5'-3')	location
EMSA		
Zac1 EMSA F	TAGAACATAGTGATACGCCATTCAGCGTTAGGCTTTGAA	
Zac1 EMSA R	TTCAAAGCCTAACGCTGAATGGCGTATCACTATGTTCTA	
Zac1 EMS mut F	TAGAACATAGTGATCATAACGGACGCGTTAGGCTTTGAA	
Zac1 EMSA mut R	TTCAAAGCCTAACGCGTCCGTTATGATCACTATGTTCTA	
Pax3 EMSA	AAT CATAAAGGCATGACTAATTGCATGG	
Footprinting probes at +11 kb from *Zac1*		
Footprint +11 kb F1	CCT TGC TCG CAT GCT TTT AG	intron 2
Footprint +11 kb R1	CCC AGG AAG ACA CAC TTT GC	intron 2
Footprint +11 kb F2	GCA GCA AGT TTG AGA CAG GA	intron 2
Footprint +11 kb R2	CAC ATG ACT GCC TGA AAA CG	intron 2
Footprint myoG F(-426 kb)	CCGTCCGTCCAAGACAACCC	
Footprint myoG R(-426 kb)	CCCCCCCTCTAAGCTGTTGC	
Cloning of luciferase constructs		
fp1+11 luc F	AGA TCT CCT TGC TCG CAT GCT TTT	intron 2
fp1 +11 luc R	GGA TCC CCC AGG AAG ACA CAC TTT GC	intron 2
fp2 +11 luc F	AGA TCT GCA GCA AGT TTG AGA CAG GA	intron 2
fp2 +11 luc R	GGA TCC CAC ATG ACT GCC TGA AAA CG	intron 2
ChIP assay		
Zac1 -500 kb F	TCTCTGCCCAGTTTCTGGTT,	
Zac1 -500 kb R	GTGACTCATCGGCTGTATGC	
Zac1 -100 kb F	GGCGTTCTCCAACCTCACT	
Zac1 -100 kb R	CCGTCAGATCTCCGATCC	
Zac1 +100 kb F	TCCGGCTAGGGTAGGTAAGT	exon 1
Zac1 +100 kb R	GAGCCTTTTGCTGCATCTCT	exon 1
Zac1 +500 kb F	GGAGAATTCAGAGCGCAACT	intron 1
Zac1 +500 kb R	CTCGATTCCTAAAACACCAAAG	intron 1
GAPDH -3kb F	TCCAGTGAGGACGG	
GAPDH -3kb R	CATAAAGATGGGGC	
MyoD -22.9 kb F	CCATAAGGCACTGG	
MyoD -22.9 kb R	TCTAAGTCTGTGCGGGTGAG	
Myf5 -57.5 kb F	TGTGGCTCTCTCTCCGTATG	
Myf5 -57.5 kb R	AATACAGACATGCAGGCTTCAC	

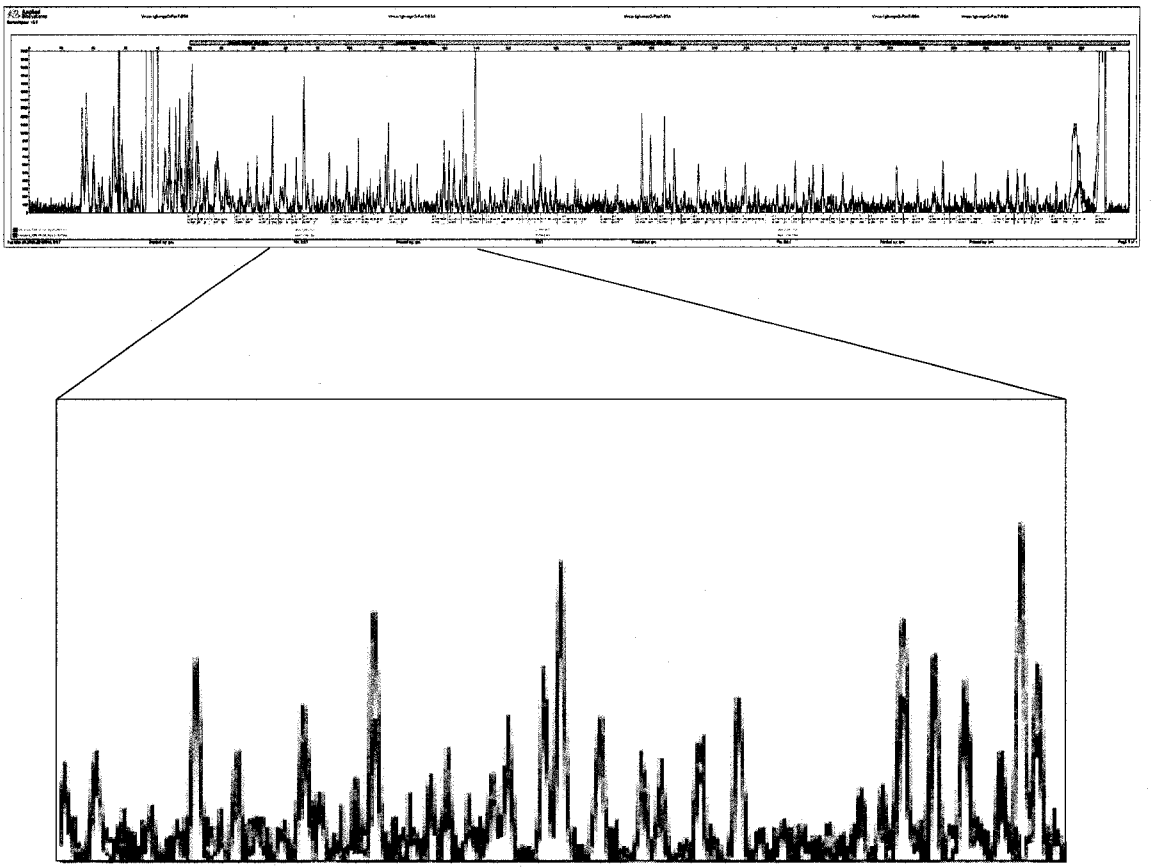


Figure S1. DNase I footprint analysis of the *myogenin* promoter. A segment of the *myogenin* promoter was amplified from C57Bl/6 genomic DNA. Primers mapped to -426 kb from *myogenin* transcptional start site and amplicon size was 336 bp. Forward primers were FAM labeled. Following *in vitro* binding reactions with purified recombinant Pax7, DNA was digested with DNAse I. Fluorescent fragments were resolved on an automated capillary electrophoresis instrument (red trace). BSA instead of Pax7 was used as a negative control (blue trace). Resulting electropherogram shows identical DNAse I cleavage patterns for both Pax7 and BSA binding reactions showing the absence of any non-specific binding from the purified Pax7 protein.