Characterizing a Distal Muscle Enhancer in the Mouse Igf2 Locus

1Damir Alzhanov and 1,2Peter Rotwein

1Department of Biochemistry and Molecular Biology,
Oregon Health & Science University, Portland, OR 97239
2Department of Biomedical Sciences,
Paul L. Foster School of Medicine,
Texas Tech University Health Sciences Center,
El Paso, TX 79905

Running head: Mapping Igf2 enhancer activity in skeletal muscle

To whom correspondence should be addressed:

Peter Rotwein
Department of Biomedical Sciences
Paul L. Foster School of Medicine
Texas Tech University Health Sciences Center
5001 El Paso Drive, El Paso, TX, USA
Tel: 915-215-4910
Fax: 915-783-5223
Email: peter.rotwein@ttuhsc.edu

Keywords: muscle differentiation, IGF2, gene transcription, distal enhancer, recombineering
Insulin-like growth factor-2 (IGF2) is highly expressed in skeletal muscle, and was identified as a quantitative trait locus for muscle mass. Yet little is known about mechanisms of its regulation in muscle. Recently, a DNA segment found ~100 kb from the Igf2 gene was identified as a possible muscle transcriptional control element. Here we have developed an in vivo reporter system to assess this putative enhancer by substituting nuclear (n) EGFP for Igf2 coding exons in a bacterial artificial chromosome containing the mouse Igf2-H19 chromosomal locus. After stable transfection into a mesenchymal stem cell line, individual clones were converted to myoblasts, and underwent progressive muscle-specific gene expression and myotube formation in differentiation medium. Transgenic mRNA and nEGFP were produced coincident with endogenous Igf2 mRNA, but only in lines containing an intact distal conserved DNA element. Our results show that a 294-bp DNA fragment containing two E-boxes is a necessary and sufficient long-range enhancer for induction of Igf2 gene transcription during skeletal muscle differentiation, and provide a robust experimental platform for its further functional dissection.
**Introduction**

Muscle regeneration requires a local population of stem cells termed satellite cells (33). Satellite cell activity is controlled by competing signals mediated by cell-cell contact, and by growth factors and hormones (4) that modify muscle genetic programs directed by myogenic transcription factors (16, 31). Among molecules with pro-myogenic actions are the insulin-like growth factors, IGF1 and IGF2 (18, 28). These proteins play critical roles in muscle development and growth in the fetus (21, 27) and post-natally (2, 25, 34), and help coordinate muscle restoration and re-innervation after injury (3, 6, 23, 24).

IGF2 was identified as a key quantitative trait locus for muscle mass in pigs, with more heavily muscled animals bearing a single nucleotide polymorphism in the IGF2 gene (34) that was responsible for stimulating gene activity by interfering with binding of a transcriptional repressor (5, 22). Other studies have shown that secreted IGF2 functions as an autocrine-paracrine factor for muscle differentiation (11, 12, 15, 17, 37, 38). Yet until recently little was known about mechanisms of regulation of IGF2 in muscle cells (1, 14).

IGF2 gene control is complicated because it is part of a conserved autosomal linkage group that undergoes parental imprinting (10). In most mammalian species, *Igf2* is expressed from the paternally derived chromosome, with adjacent *H19* being activated on the maternal chromosome via epigenetic actions of an imprinting control region (ICR) found in inter-genic chromatin (10). DNA within the ICR contains sites for the nuclear protein, CTCF (10, 26, 35). In maternally derived chromatin, CTCF binds to the ICR and facilitates access of distal enhancers to the *H19* promoter while preventing interactions with *Igf2* promoters (26, 35). In paternal chromatin, DNA in the ICR is methylated and CTCF cannot bind, thus allowing the same enhancers to associate with *Igf2* (26, 35).
Previous studies have shown that induction of *Igf2* gene transcription during muscle differentiation was correlated temporally with the physical association between *Igf2* promoters and a putative enhancer located > 100 kb 3’ to the mouse *Igf2* gene and ~24 kb 3’ to *H19* (1, 14, 40). The boundaries of this element have not been established nor has its sufficiency for muscle-specific *Igf2* expression been determined. Here we have tested a series of modified bacterial artificial chromosomes derived from the murine *Igf2 - H19* locus to show that a 294-bp core enhancer encoding two E-boxes is necessary and sufficient for *Igf2* gene transcription in skeletal muscle cells.
Materials and Methods

Materials. G418 was from Sigma-Aldrich (St. Louis, MO), okadaic acid from Alexis Biochemicals (San Diego, CA), Immobilon-FL membranes from Millipore (Billerico, MA), and TransIT LT1-transfection reagent from Mirus (Madison WI). Dulbecco’s modified Eagle’s medium (DMEM), trypsin/EDTA solution, SuperScript III reverse-transcriptase (RT) kit, SYBR Green Platinum qPCR mix, pCR2.1-TOPO, and Hoechst 33258 dye were from Invitrogen-Life Technologies (Carlsbad, CA). Fetal calf serum and phosphate-buffered saline were from Mediatech-Cellgrow (Herndon, VA). The Qia-Quick PCR purification kit was from Qiagen (Valencia, CA), and restriction enzymes, buffers, and polymerases were obtained from New England Biolabs (Beverly, MA). Primary antibodies to troponin-T (clone #CT3) were purchased from the Developmental Studies Hybridoma Bank (Iowa City, IA). Secondary antibodies were from Invitrogen-Life Technologies (goat anti-mouse IgG conjugated with Alexa Fluor 594). Custom oligonucleotides were from Thermo Scientific (Waltham, MA). Bacterial artificial chromosome (BAC) clone BMQ318O12 was from David Adams (Sanger Center, United Kingdom), and materials for recombineering were from the NCI (Frederick, MD). Other chemicals were from commercial sources.

Developing recombinant cell lines containing a mouse Igf2 BAC. BAC BMQ318O12 encodes 201,288 base pairs of mouse chromosome 7, and spans genomic coordinates 142.482.694 - 142.683.981, including the Ins2, Igf2, H19, Ntcp1, Mrpl23, and Tnnt3 genes and flanking regions, and part of Lsp1 (Ensemble Mus musculus genome assembly GRCn38p4; Fig. 1A). A cassette encoding the neomycin resistance gene driven by a combined prokaryotic and eukaryotic promoter was inserted into the BAC by homologous recombination (recombineering) using E. coli strain SW102 and selection on kanamycin-containing agar plates (20, 41). Coding exons 3 - 6 of the Igf2 gene were replaced with a nuclear-targeted EGFP (nEGFP), in which 3 copies of the SV40 nuclear localization sequence (NLS) were inserted in-frame at the COOH-terminus of
EGFP by a positive-negative selected approach employing the *E. coli* galactokinase gene (galK) (36). An analogous strategy was used for other recombinant BACs (Fig. 1B). All DNA changes were validated by PCR and DNA sequence analysis; in all BACs an intact modified *Igf2* gene was verified. Each recombinant BAC was incorporated into the genome of mouse C3H10T1/2 cells (ATCC #CCL226) following transfection and selection in medium containing G418 [400 µg/ml]. Individual colonies were isolated, expanded, and stored in liquid N₂ until use.

Transgene copy number was assessed by quantitative PCR (9). A standard curve was constructed using serial dilutions of BAC DNA and mouse *Igf2*-specific primers by plotting the cycle threshold (Ct) vs. amount of input DNA, and the slope and correlation coefficient were determined (see (8)). Results were judged acceptable if r² was > 0.98.

**Analysis of recombinant cell lines.** Cells were incubated in antibiotic-free DMEM with 10% fetal bovine serum and 200 µg/ml G418 at 37°C in humidified air with 5% CO₂. For studies of muscle differentiation, cells were infected at ~50% of confluent density with a recombinant adenoviruses for MyoD (Ad-MyoD) (39), followed by washing in low-serum medium 18-20 hr later, and incubation in DM for up to 72 hr.

**Analysis of gene expression.** Whole cell RNA was isolated and assessed by agarose gel electrophoresis (38). Total RNA (2 µg) was reverse-transcribed with oligo-dT primers, and cDNA (0.5 µl) used as a template for conventional PCR (1). Pilot studies established cycle numbers that achieved linear amplification (~24 - 28 cycles). Primer pairs are found in Table 1. Products were visualized after agarose gel electrophoresis by staining with ethidium bromide. Results are representative of ≥ 3 independent experiments. Quantitative RT-PCR was performed as described (7) with BioRad Chromo4 Real-Time PCR detection system and the same primer sets (Table 1).

**Immunocytochemistry.** Cells were fixed and incubated with antibodies to troponin-T in
blocking buffer for 16 hr at 4°C (37). Secondary antibody (1:2000 dilution) and Hoechst dye (1:1000) were added in the dark for 1.5 hr. Images were captured with a Nikon DS-Qi1Mc camera and Nikon Eclipse Ti-U inverted microscope using NIS elements 3.1 software.
Results and Discussion

Developing chromosomally integrated mouse Igf2 transgenes. We previously identified a DNA segment within the mouse Igf2 - H19 locus that interacted in chromatin with Igf2 gene promoters more than 100 kb away, but only during muscle differentiation (1). Biochemical studies using promoter-reporter assays revealed that a 294-bp element within this segment could activate a linked Igf2 promoter in differentiating myoblasts, but not in fibroblasts (1). To test the hypothesis that this element was required for Igf2 gene transcription in muscle, we developed recombinant cell lines in which nEGFP-tagged versions of Igf2 in the Igf2 - H19 locus were integrated in chromatin (Fig. 1B). The structure and potential expression of intact Igf2 versus Igf2-nEGFP in muscle are depicted in Fig. 1C. To generate cells in which the minimal muscle enhancer element was deleted, a bar-coded DNA cassette lacking binding sites for transcription factors was engineered to replace the 294-bp DNA fragment (enhancer knock-out [Enh-KO], Fig. 1B, middle). In a third recombinant BAC, enhancer DNA was modified to lack paired E-boxes (Enhancer Ebox knock-out [Enh-Ebox-ko], Fig. 1B, lower middle), which we had found were responsible for high-level activity of the putative enhancer element in promoter-reporter experiments in muscle cells (1). In a fourth BAC, the enhancer DNA was re-inserted further from Igf2 (Enhancer knock-in [Enh-KI], Fig. 1B, bottom).

The mutation in Enh-KO and Enh-KI BACs would potentially disrupt the Nctc1 transcription unit (14). Nctc1 encodes a long non-coding RNA found to be expressed in muscle, though at levels ~0.1% of Igf2 (13). Eun et al have argued that Nctc1 transcription is required for muscle enhancer function in chromatin (14). Our studies provide an independent test of this hypothesis.

Transgenic DNA was integrated into C3H10T1/2 cell chromatin by selection with G418. Individual clones were isolated and expanded, and all lines were tested for an intact modified Igf2 gene, for pGK Neo, and for other components of the modified BAC by PCR and targeted
DNA sequencing. Three to six cell lines of those screened for each construct maintained all of these features, and representative cells were chosen for further study. Copy numbers of recombinant BACs in the lines used here varied from 2 - 4 per haploid genome (Fig. 1D).

**Expression of Igf2-nEGFP during muscle differentiation depends on an intact distal enhancer.** To test the hypothesis that stimulation of Igf2 gene transcription during skeletal muscle differentiation requires the distal enhancer, C3H10T1/2 cells containing wild-type, Enh-KO, or Enh-Ebox-ko BACs were converted to myoblasts by transduction with Ad-MyoD (39), followed by incubation in DM (Fig. 2A). Time-course studies revealed that induction of endogenous myogenin and Igf2 transcripts occurred within 24 hr of addition of DM, that expression was sustained for at least 72 hr, and that Igf2 mRNA levels were equivalent in both cell lines (Fig. 2B, C). Igf2-nEGFP transgenic mRNA was expressed with identical kinetics as endogenous mouse Igf2 in cells encoding the wild-type BAC, but was undetectable in Enh-KO or Enh-Ebox-ko myoblasts (Fig. 2C). Analysis by immunocytochemistry of muscle differentiation revealed that formation of multi-nucleated myotubes proceeded equivalently in all three lines (Fig. 2D). EGFP was seen within the nucleus only in differentiating myoblasts containing the wild-type BAC (Fig. 2D, E). Thus, based on these results, an intact 294-bp distal enhancer appears to be required for induction of Igf2 gene transcription during muscle differentiation.

**An enhancer knock-in restores Igf2-nEGFP fusion gene expression.** Having established that the distal enhancer was necessary for production of chimeric Igf2-nEGFP transcripts and nEGFP during muscle differentiation, we next asked if its replacement elsewhere within the chromosomal locus could restore transgenic gene expression. Cells containing the Enh-KI BAC (Fig. 1B) were converted to myoblasts with Ad-MyoD, followed by addition of DM. Time-course experiments demonstrated that muscle gene and protein expression, and progressive myotube formation occurred in cells containing the Enh-KI BAC with kinetics similar to those of
the wild-type BAC (Fig. 3A, B). However, nEGFP was detected only in differentiating
myoblasts containing wild-type BAC (Fig. 3B), potentially because levels of \textit{Igf2}-EGFP mRNA
transcribed in Enh-KI myoblasts, as measured by qRT-PCR, were only ~25-30% of those
produced by the wild-type BAC transgene (Fig. 3C). Based on these results, the 294-bp
enhancer element seems to have some positional flexibility in promoting \textit{Igf2} gene transcription
during muscle differentiation, but its activity appears to be attenuated in its new location.

\textbf{Transcriptional control at a distance in the \textit{Igf2} - \textit{H19} locus.} Taken collectively, our results
show that a 294-bp DNA fragment, previously identified as a component of a putative distal
muscle enhancer within the mouse \textit{Igf2} - \textit{H19} locus, is both necessary and sufficient for
induction of \textit{Igf2} gene transcription during skeletal muscle differentiation. These observations
extend prior studies demonstrating that a ~20 kb deletion 3’ to \textit{H19} impaired both \textit{Igf2} and \textit{H19}
gene expression in muscle in transgenic mice and in cultured myoblasts (19), and complement
more recent data focusing on the functions of the long non-coding RNA, \textit{Nect1}, which is located
within the deleted region (13, 14). The muscle enhancer defined here maps to the second intron
of \textit{Nect1}, and is part of its transcription unit (14). Eun \textit{et al} have concluded that active
transcription of \textit{Nect1} is required in \textit{cis} along with the 294-bp element for muscle-specific
activation of \textit{Igf2} and \textit{H19} genes (14). Our studies suggest a more nuanced interpretation, in
which the core 294-bp segment drives the majority of enhancer function, perhaps in part through
the actions of myogenic transcription factors such as MyoD, which recognizes E-boxes (16, 31),
and has been found to bind to this region of the genome in muscle cells (see NCBI sequence read
archive, accession number SRP001761). However, maximal activity appears to require that all
components, including \textit{Nect1}, be in their native anatomical orientation. Strategies to test this
hypothesis in more detail can now be initiated, and may be coupled with parallel approaches to
determine if enhanced expression of endogenous IGF2 can improve muscle mass in chronic
diseases accompanied by sarcopenia (30, 32).
Acknowledgements

We thank the Oregon Health & Sciences University DNA Services Core for performing DNA sequencing. Support for the studies in this manuscript is from NIH grant R01 DK042748-25 (to P. R.).

Disclosures

The authors declare that they have no financial conflicts of interest pertinent to the contents of this article.


17. Ge, Y., Y. Sun, and J. Chen. IGF-II is regulated by microRNA-125b in skeletal myogenesis.


34. Van Laere, A. S., M. Nguyen, M. Braunschweig, C. Nezer, C. Collette, L. Moreau, A. L.


Table 1. Primers used for RT-PCR of total RNA

<table>
<thead>
<tr>
<th>Gene</th>
<th>Location</th>
<th>DNA sequence</th>
<th>Product Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Igf2</td>
<td>exon 3, exon 4</td>
<td>5′- GCAAACTGGACATTAGTTCT 5′- ACTGAAGCGTGTCACAAAGCTC</td>
<td>229</td>
</tr>
<tr>
<td>Igf2 - EGFP fusion</td>
<td>exon 3 (5′ end)</td>
<td>5′- GCAAACTGGACATTAGTTCT 5′- GTCGCCGTCCAGCTGACCAG</td>
<td>147</td>
</tr>
<tr>
<td>myogenin</td>
<td>exon 1, exon 3</td>
<td>5′- ATGGAGCTGTATGAGACATCC 5′- TCAGTTGGCATGTTTCGTC</td>
<td>674</td>
</tr>
<tr>
<td>S17</td>
<td>exon 2, exon 3</td>
<td>5′- ATCCCCAGCAAGAAGCTTCGGAAC 5′- TATGGCATAACAGATTAACAGCTC</td>
<td>323</td>
</tr>
</tbody>
</table>
Figure Legends

Figure 1. Development of transgenic cell lines encoding recombinant BACs containing a modified Igf2 gene. A. Schematic of Igf2 – H19 locus and adjacent genes on mouse chromosome 7. Genes encoding proteins are indicated by a black box; genes expressed as non-coding RNAs (H19 and Nctc1) are labeled by a white box. The red oval signifies a putative muscle enhancer. B. Top: Diagram of wild-type (WT) BAC, with its genomic coordinates, illustrating modifications made to replace Igf2 coding exons with nEGFP and to insert a pGK-Neo selection cassette. The gray color of Lsp1 indicates that only part of the gene is present. Middle: Diagram of Enh-KO BAC, demonstrating replacement of putative muscle enhancer with cassette containing bar-coded DNA. The broken outline of Nctc1 indicates disruption of its transcription unit. Bottom middle: Diagram of Enh-Ebox-ko BAC, demonstrating deletion of paired E-boxes within the muscle enhancer. Bottom: Schematic of enhancer knock-in BAC (Enh-KI). The broken outline of Nctc1 indicates that its transcription unit is perturbed. C. Left: Organization of mouse Igf2 gene (6 exons, 5 introns, 3 promoters (29)). Arrows show locations of primers for Fig. 1D. Below are schematics of major muscle Igf2 mRNA (containing exons 3 - 6), and IGF2 protein precursor (SP, signal peptide; IGF2, 67 amino acid mature protein; E peptide, COOH-terminal 89 amino acid extension). Right: Organization of Igf2-nEGFP gene. Arrows show locations of primers for Fig. 1C. Below are Igf2-nEGFP mRNA, and nuclear-localized EGFP protein (EGFP-nls). D. Copy number determination for lines encoding wild-type (WT) BAC, BAC lacking putative muscle enhancer (KO), and BAC in which putative enhancer has been relocated (KI). The left panel shows PCR-amplified endogenous Igf2 DNA and right panel, amplified transgene DNA. Calculated copy numbers: WT - 2; KO - 4; KI - 2; E-ko - 4 per haploid genome (see Methods for details).

Figure 2. Production of nEGFP by the recombinant Igf2 locus during muscle differentiation. A. Experimental scheme. B. Schematic of native and modified mouse Igf2
genes, showing locations of primers for RT-PCR. C. Time course of expression for \( Igf2, Igf2\)-nEGFP, myogenin and \( S17 \) transcripts, measured by RT-PCR. Con indicates negative and positive controls. D. Time course of myotube formation by immunocytochemistry for troponin-T (red) and nuclear staining by Hoechst (blue). Expression of EGFP in nuclei is indicated by green tint. E. Higher magnification view of 48 hr time point shown in D. For D and E, scale bar = 50 μM.

**Figure 3. Restoration of \( Igf2 \) expression during muscle differentiation by knock-in of the distal enhancer.** A. Expression of \( Igf2, Igf2\)-nEGFP, myogenin, and \( S17 \) RNAs, assessed by RT-PCR in differentiating cells encoding either WT or enhancer KI BACs. Con denotes negative and positive controls. B. Time course of myotube formation by immunocytochemistry for troponin-T (red) and nuclear staining by Hoechst (blue). Expression of EGFP in nuclei is indicated by green tint. Scale bar = 50 μM. C. Relative gene expression for \( Igf2\)-nEGFP by quantitative RT-PCR at 0 and 24 hr after incubation in DM for cells encoding either WT, enhancer KO, or enhancer KI BACs. Note log scale on ordinate.
Fig 2

A. 10T1/2 Igf2-BAC Cells
   \[ \xrightarrow{\text{Ad-MyoD}} \] DM
   \[ \xrightarrow{\text{Assess differentiation}} \]

B. Diagram of Igf2 promoter regions
   \( P_1, P_2, P_3 \)
   \( \text{Igf2} \)
   \( \text{nuclear EGFP} \)
   \[ \text{2 kb} \]

C. Gel analysis:
   - Igf2
   - Igf2-EGFP
   - myogenin
   - S17
   
<table>
<thead>
<tr>
<th>hrs in DM:</th>
<th>0</th>
<th>24</th>
<th>48</th>
<th>72</th>
<th>0</th>
<th>24</th>
<th>48</th>
<th>72</th>
<th>0</th>
<th>24</th>
<th>48</th>
<th>72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Con</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

   Wild-type Igf2 BAC | Enhancer KO | Enhancer E-box-ko

D. Immunofluorescence images:
   - Wild-type
   - Enh-KO
   - Enh-Ebox-ko
   
<table>
<thead>
<tr>
<th>hrs in DM:</th>
<th>0</th>
<th>24</th>
<th>48</th>
<th>72</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enh-KO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enh-Ebox-ko</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E. Immunofluorescence images with EGFP marker:
   - Wild-type
   - Enh-KO
   - Enh-Ebox-ko
   
   Scale bars indicate 100 μm.