Regulation of alternative splicing of *Gtf2ird1* and its impact on slow muscle promoter activity

Enoch S. E. TAY*[1], Kim L. GUVEN*[1], Nanthakumar SUBRAMANIAM*, Patsie POLLY*, Laura L. ISSA*, Peter W. GUNNING† and Edna C. HARDEMAN*[2]

*Muscle Development Unit, Children's Medical Research Institute, Locked Bag 23, Westmead, NSW 2145, Australia, and †Oncology Research Unit, The Children's Hospital at Westmead, PO Box 3515, Parramatta, NSW 2124, Australia

A human MusTRD [muscle TFII-I repeat domain (RD)-containing protein] isoform was originally identified in a yeast one-hybrid screen as a protein that binds the slow fibre-specific enhancer of the muscle gene troponin I slow [O'Mahoney, Guven, Lin, Joya, Robinson, Wade and Hardeman (1998) Mol. Cell. Biol. **18**, 6641–6652]. MusTRD shares homology with the general transcription factor TFII-I by the presence of diagnostic I-RDs [Roy (2001) Gene **274**, 1–13]. The human gene encoding MusTRD, *GTF2IRD1* (*WBSCR11/GTF3*), was subsequently located on chromosome 7q11.23, a region deleted in the neurodegenerative disease, Williams–Beuren Syndrome [Osborne, Campbell, Daradich, Scherer, Tsui, Franke, Peoples, Francke, Voit, Kramer et al. (1999) Genomics **57**, 279–284; Franke, Peoples and Francke (1999) Cytogenet. Cell. Genet. **86**, 296–304; Tassabehji, Carette, Wilmot, Donnai, Read and Metcalfe (1999) Eur. J. Hum. Genet. **7**, 737–747]. The haploinsufficiency of MusTRD has been implicated in the myopathic aspect of this disease, which manifests itself in symptoms such as lowered resistance to fatigue, kyphoscoliosis, an abnormal gait and joint contractures [Tassabehji, Carette, Wilmot, Donnai, Read and Metcalfe (1999) Eur. J. Hum. Genet. **7**, 737–747]. Here, we report the identification of 11 isoforms of MusTRD in mouse skeletal muscles. These isoforms were isolated from a mouse skeletal muscle cDNA library and reverse transcription–PCR on RNA from various adult and embryonic muscles. The variability in these isoforms arises from alternative splicing of a combination of four cassettes and two mutually exclusive exons, all in the 3′ region of the primary transcript of *Gtf2ird1*, the homologous mouse gene. The expression of some of these isoforms is differentially regulated spatially, suggesting individual regulation of the expression of these isoforms. Co-transfection studies in C2C12 muscle cell cultures reveal that isoforms differentially regulate muscle fibre-type-specific promoters. This indicates that the presence of different domains of MusTRD influences the activity exerted by this molecule on multiple promoters active in skeletal muscle.

Key words: alternative splicing, BEN protein, GTF3 protein, muscle TFII-I repeat domain-containing protein (MusTRD), transcription factor TFII-I.

INTRODUCTION

Skeletal muscle is composed predominantly of two major fibre types: slow and fast twitch. The fibre type composition of each muscle determines its physiological capabilities, slow oxidative fibres being required for maintenance of posture and tasks involving stamina, whereas fast glycolytic fibres are required for movements involving strength and speed. These fibre types are defined in large part by the expression of various isoforms of the muscle structural proteins [1,2]. Fibre specification is achieved at a relatively late stage in skeletal muscle development, and involves a combination of programming in the muscle progenitor cells that gives rise to fibres and a response by fibres to motoneuron contact [3,4]. Mature skeletal muscle possesses the ability to switch fibre type, termed 'plasticity', in response to changes in innervation status and the sensing of biomechanical load [5–8]. There is a substantial body of literature delineating the molecules and defining their roles in the differentiation and early development of skeletal muscle [9]. Relatively little is known about the mechanisms that determine and maintain skeletal muscle-fibre types. Recently, the calcineurin [10,11] and Ras-mitogen-activated protein kinase [12] signalling pathways have been shown to be involved in the maintenance and/or nerve-dependent expression of slow fibre-specific genes in adult muscle.

In order to identify possible regulators that directly control muscle fibre-specific genes, we isolated a novel transcription factor using the enhancer of a representative slow fibre-specific gene, TnI_{slow} (troponin I slow), as bait in a yeast one-hybrid screen of a human skeletal muscle cDNA library [13]. This molecule was named MusTRD1 [muscle TFII-I repeat domain (RD)-containing protein 1], because of its homology with the general transcription factor TFII-I. Both transcription factors contain homologous I-RDs (70 % homology) and a leucine zipper in the extreme N-terminus [14]. In accordance with our hypothesis that MusTRD

Abbreviations used: DBD, DNA-binding domain; ECU, extensor carpii ulnaris; EDL, extensor digitorium longus; EOM, extraocular muscle; FCR, flexor carpii radialis; FCU, flexor carpii ulnaris; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RD, repeat domain; GTF2IRD1, GTF2I RD-containing protein 1; HLH, helix–loop–helix; MHC(-I), myosin heavy chain (type I); $MLC2_{slow}$, myosin light chain 2 slow; (m/h)MusTRD, (mouse/human) muscle TFII-I RD-containing protein; NLS, nuclear localization signal; RACE, rapid amplification of cDNA ends; RT, reverse transcription; TnI_{slow}, troponin I slow; UTR, untranslated region; WBS, Williams–Beuren Syndrome.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed (e-mail ehardeman@cmri.usyd.edu.au).

The nucleotide sequence data reported for mMusTRD1α1, mMusTRD3β3, mMusTRD1α4, mMusTRD3α5, mMusTRD1β1, mMusTRD3α7, mMusTRD1β4, mMusTRD3β5, mMusTRD2α5, mMusTRD3β7, mMusTRD3α3 and BEN will appear/have appeared in DDBJ, EMBL, GenBank® and GSDB Nucleotide Sequence Databases under the accession numbers AF497637, AF497641, AF497638, AF497642, AF247161, AF343348, AF497639, AF343350, AF497640, AF343351, AF343349 and AY030287 respectively.

has a role in regulating the expression of *TnI$_{slow}$*, we have shown that MusTRD1 can act as a repressor of the TnI$_{slow}$ enhancer (P. Polly, L. M. Haddadi, L. L. Issa, N. Subramaniam, S. J. Palmer, E. S. E. Tay and E. C. Hardeman, unpublished work).

The human gene encoding MusTRD, *GTF2IRD1/WBSCR11/GTF3* (where GTF2IRD1 is GTF2I RD-containing protein 1), is located on chromosome 7q11.23, a region deleted in the multifactorial neurodegenerative disease, Williams–Beuren Syndrome (WBS) [15–17]. Patients suffering from WBS display a range of potentially myopathic symptoms, such as an abnormal gait, kyphoscoliosis, joint contractures, muscle weakness, altered muscle fibre size and a low resistance to fatigue [18]. MusTRD's ability to bind an enhancer element within *TnI$_{slow}$*, its involvement in regulating this gene's expression and its haploinsufficiency in WBS suggest its involvement in the myopathic aspect of WBS.

There are currently two known human isoforms from this gene, MusTRD1 and a retinoblastoma-associated nuclear protein, Cream1 [19], that arise from alternative splicing of a 57 bp exon (exon 19) [13] and five mouse isoforms [20–22]. Recently, an isoform was isolated from *Xenopus* that appears to be the orthologue of MusTRD1 [23]. Here, we report the identification of 11 mouse (m)MusTRD isoforms isolated from adult and embryonic skeletal muscles. These isoforms arise from the alternative splicing of *Gtf2ird1*, located on mouse chromosome 5, that involves both single exon and multi-exon cassettes, use of alternative 5′ donor and 3′ acceptor sites (that are mutually dependent) and alternative 3′ terminal exons. In general, these isoforms are expressed ubiquitously in adult and embryonic muscles, but several are subject to spatial regulation. Co-transfection studies of two MusTRD isoforms that differ in their use of alternative C-terminal exons with contractile protein promoters indicate that these isoforms possess differential regulatory capabilities. These differential regulatory capabilities suggest that the MusTRD family of transcription factors have a role in the regulation of multiple muscle-specific promoters.

EXPERIMENTAL

MusTRD antibody

Sheep anti-MusTRD serum was raised against the first 20 amino acids of MusTRD (Mimotopes, Clayton, Victoria, Australia). Hyperimmune serum was affinity-purified using biocytin-tagged N-terminal peptides coupled with M-280 streptavidin Dynabeads (Dynal®, Carlton, Victoria, Australia). Immobilized MusTRD antibody was eluted with 3 M $MgCl_2$, pH 7.2. Antibody preparations were dialysed in a buffer containing 25 mM Tris/HCl, pH 7.2, 0.15 M NaCl and 0.1 % (w/v) BSA (TBS-BSA) using Slide-A-Lyzer® Mini Dialysis Units (10 000 MWCO; Pierce, Rockford, IL, U.S.A.). Purified antibody was stored in TBS-BSA with 0.1 % Tween-20 at 4 °C.

Confocal microscopy

Sections (6 μm) of soleus muscle from 8-week-old male B6D2 mice were fixed in 2 % (w/v) buffered paraformaldehyde for 10 min at 4 °C, washed in ice-cold TBS containing 0.2 % Tween-20, then blocked with 10 % donkey serum in TBST for 30 min at 22 °C. Sections were exposed to purified sheep anti-MusTRD antibody (1:5 to 1:10 dilutions) in TBST containing 10 % donkey serum overnight at 4 °C, followed by donkey anti-sheep IgG conjugated to FITC (1:50 dilution) for 1 h at 22 °C (Sigma). Slides were washed and re-labelled with a mixture of mouse anti-dystrophin (DYS2; Novocastra Laboratories, Benton Lane, Newcastle, U.K.) and mouse antibody for myosin heavy-chain (MHC) type I (MHC-I) (BAF8) [24] overnight at 4 °C. Goat anti-mouse Rhodamine Red-X conjugated antibody was applied for 90 min at 22 °C (Jackson ImmunoResearch, West Grove, PA, U.S.A.). Muscle nuclei were stained with propidium iodide and the slides were mounted in 2.5 % DABCO (1,4-diazadicyclo[2.2.2]octane) in 80 % (v/v) glycerol. Sections were visualized with a Leica confocal laser-scanning microscope.

Isolation and characterization of cDNA clones

A mouse skeletal muscle 5′ plus stretch λgt11 cDNA library (Clontech) was screened with a random primer labelled (Giga Prime Labelling Kit; Geneworks) 1.3 kb *Aat*II/*Pst*I fragment of human (h)MusTRD1 (accession no. NM_005685), 28–1364 bp downstream of the start codon and containing RD1 and RD2. Hybridization was performed in Church buffer [0.5 M Na_2HPO_4 (pH 7.2)/1 % (w/v) BSA/7 % (w/v) SDS/1 mM EDTA (pH 8)] at 50 °C overnight and washed at 60 °C in 0.5 × SSC/0.1 % SDS (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), with three changes, over 30 min.

Rapid amplification of cDNA ends (RACE)

5′ RACE PCR was carried out using a mouse skeletal muscle Marathon-Ready cDNA library (Clontech). Oligonucleotides used were the AP1 forward primer (Marathon cDNA Adaptor primer; Clontech) and reverse primer specific to the MusTRD isoforms located downstream of RD2 (5′-GATCCCACTTCTCTG-ACTTGTCATG-3′). PCR was performed using Advantage HF PCR polymerase mix (Clontech) and MasterAmp PCR Optimization Buffer D (Epicentre Technologies) under the following conditions: 94 °C for 5 s, 72 °C for 4 min for a duration of 5 cycles, 94 °C for 5 s, 70 °C for 4 min for a duration of 5 cycles, 94 °C for 5 s, 68 °C for 4 min for a duration of 30 cycles. Southern blotting was used to confirm the presence of MusTRD sequence in the clones. A hMusTRD1 300 bp *Bam*HI probe [13] or a mouse EST clone 555 547 (accession no. AA111609; 4–103 bp) probe labelled with a random primer labelling kit (Giga Prime Labelling Kit; Geneworks) was used. Blots were hybridized overnight at 65 °C in Church buffer and washed at 65 °C in 0.5 × SSC/0.1 % SDS with three changes over 30 min.

RNA isolation

Total RNA was isolated from differentiated C2C12 cultures, adult extraocular muscle (EOM), flexor carpii radialis (FCR), flexor carpii ulnaris (FCU), extensor carpii ulnaris (ECU), plantaris, quadriceps, soleus and extensor digitorium longus (EDL) muscles excised from fifteen 10-week-old female ARC/S mice using the solution D method [25], whereas total RNA was isolated from the hind-limbs of 20 B6D2 13.5 days post-coitus embryos (E13.5) by Trizol™ (Life Technologies) extraction, according to the manufacturer's instructions.

cDNA synthesis and real-time reverse transcription (RT)-PCR analysis of isoforms

cDNA was generated using Im-prompII™ reverse transcriptase (Promega) according to manufacturer's instructions with 2 μg of total RNA as a template. Real-time RT-PCR reactions were carried out using 5 % of the RT reaction as template, 2 × buffer D (Epicentre Technologies) and SYBR® Green. Amplification reactions were established using a series of forward and reverse

Table 1 Primer sequences

Sequences of primers used in PCR reactions are shown.

Primer	Sequence
Forward	
5′UTR	5′-CAACCAGAGGCGACTGGATC-3′
Ex19	5′-GACCGTCTTGTGGACGAGACC-3′
Ex18/20	5′-CTGGACACTCAAGAAAATTACAAC-3′
Ex23	5′-ACCAGACCAAGGAGACTGCAACAG-3′
Ex22/24	5′-CAAGGACTTATCCCAAAGCCTGAT-3′
Ex20/24	5-′CAAGAAATACGATGAGGATGATG-3′
GapDHF	5′-GCTGGCATTGCTCTCAATGACAAC-3′
Reverse	
Ex30	5′-GGAGGTTGAGTTTCGTCACGTGA-3′
Ex31	5′-TGGCGGCAGGAATATAGTG-3′
482R	5′-GGCTCCCTGAGCAGCCTCTCATAG-3′
GapDHR	5′-GGGTGCAGCGAACTTTATTGATGG-3′

primers (listed in Table 1). Reactions were subjected to an initial denaturation step at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 1 min and 72 °C for 1 min in a Rotor-Gene™ real-time DNA-amplification system (Corbett Research, Australia). SYBR® Green fluorescence was measured after each extension step. RT-PCR products were subjected to measurement of their annealing temperatures and analysis by gel electrophoresis to verify their specificity and identity. A standard regression curve was drawn for each target mRNA to estimate its relative amount and quantification was derived from the exponential phase of the reaction. PCR products were purified using Qiaquick gel extraction kit (Qiagen) and cloned into pGEM-T Easy vector (Promega). Forward and reverse primers specific for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) were designed (Table 1) and used as controls to normalize all RT-PCR experiments.

Plasmid constructs for C2C12 transfections

The cDNAs for human h1α1 and mouse isoforms m3α7, m1β1, and m3β7 were subcloned into the *Age*I site of the CMV (cytomegalovirus) promoter driven pcDNA3.1 *myc/his* expression vector (Invitrogen, Sydney, Australia). Reporter plasmid USE-TK-Luciferase contains the $hTnI_{slow}$ upstream enhancer (USE; 157 bp from −1031 to −875). (−3500) βMHC-pGL3 and (−2554 to +13) IIβMHC-pGL3 reporters containing regions of the MHC-I gene were kindly provided by Ken Baldwin and colleagues [26,27], and the luciferase reporter construct −800MLC2s was kindly provided by Karyn Esser et al. [28].

Cell culture and transient transfections

The C2C12 line is a subclone of the mouse skeletal muscle cell line C2 [29,30]. Monolayer C2C12 cells were maintained in DMEM (Dulbecco's modified Eagle medium) supplemented with 20 % (v/v) fetal-calf serum and 0.5 % chicken embryo extract (both from Gibco BRL) at 37 °C in 10 % CO_2. C2C12 cells were plated at 30–50 % confluence in 24-well plates, cultured for 12 h and then co-transfected using LIPOFECTAMINE™ according to the manufacturer's protocol (Gibco BRL). Plasmid combinations transfected into cells included 100 ng of pcDNA3.1 or pcDNA3.1 containing cDNA for human or mouse isoform expression plasmids, 250 ng of reporter plasmids and 25 ng of RSV-β-galactosidase plasmid as an internal standard. After 6 h incubation at 37 °C in 10 % CO_2, the medium containing the DNA–LIPOFECTAMINE™ mix was removed, and fresh medium supplemented with 2 % horse serum (differentiation medium) was added. After incubation for 36 h in differentiation medium, the cells were lysed and luciferase activity was determined in a TopCount Microplate Scintillation & Luminescence counter (Packard). Values were normalized to β-galactosidase activities. Transfection efficiency varied from 30–40 %. Each experimental condition was measured in triplicate, and the values shown represent the means ± S.D. for three independent experiments. The luciferase activity for each promoter was compared by ANOVA (analysis of variance) and, if found significant, individual comparisons were performed by unpaired t tests. The level of significance was set at $P < 0.05$.

RESULTS

MusTRD is located in the nuclei of mouse muscle fibres

Transverse sections of an adult soleus muscle were incubated with an anti-MusTRD antibody, which was raised against an N-terminal peptide and should, therefore, recognize all known isoforms. Anti-MHC-I and anti-dystrophin antibodies were co-incubated with the anti-MusTRD antibody in order to define fast- and slow-twitch myofibres, and to delineate the position of the muscle fibre plasmalemma. MusTRD proteins could be detected inside muscle fibres, as defined by the ring of dystrophin staining, localizing to the nuclei of both fast and slow muscle fibres (Figure 1).

Isolation, characterization and nomenclature of mMusTRD isoforms

Screening of a mouse skeletal muscle 5′ plus stretch λgt11 cDNA library (Clontech) resulted in the isolation of seven different mMusTRD clones, the sequencing of which revealed extensive splicing at the C-termini. 5′ RACE PCR on the basis of the conserved 5′ end of the cDNAs yielded a product identical with the mouse EST clone BEN (AY030287). RT-PCR was then conducted on mouse adult and embryonic muscles using forward primer 5′UTR (untranslated region) and reverse primers Ex30 and Ex31. Eight different isoforms were identified that differ only in their C-terminal halves. An additional three isoforms were subsequently isolated from differentiated C2C12 myotube RNA, increasing the number of different isoforms to 11. Extensive sequence analysis and alignments mapped all coding sequences to the mouse *Gtf2ird1* gene (Figure 2A). Several of these isoforms (3β7, 3α7, 1α1, 3α5 and 1α4) correspond to isoforms reported by other laboratories [20–22], whereas the isoforms 2α5, 3α5, 3β5, 1β1, 1β4, 3α3 and 3β3 are previously unreported.

These isoforms are composed of a conserved N-terminus and a variable C-terminus. The variation in the C-terminus of the protein arises from alternative splicing of four cassettes, in addition to two mutually exclusive C-terminal exons (Figure 2B). Two cassettes consist of single exons, 19 and 23, whereas two consist of several exons. The first of these multi-exon cassettes contains exons 21, 22 and 23. The splicing mechanism described for the GTF3 and BEN isoforms [21,22] shows the second multi-exon cassette containing exons 26–28. However, detailed nucleotide sequence alignment between cDNA and genomic sequences demonstrates unambiguously that this cassette is made up of the 3′ region of exon 25, exons 26 and 27 and the 5′ region of exon 28 (Figures 2C and 2D). The two mutually exclusive exons, 30 and 31, provide alternative C-termini (Figure 2C). MusTRD isoforms contain

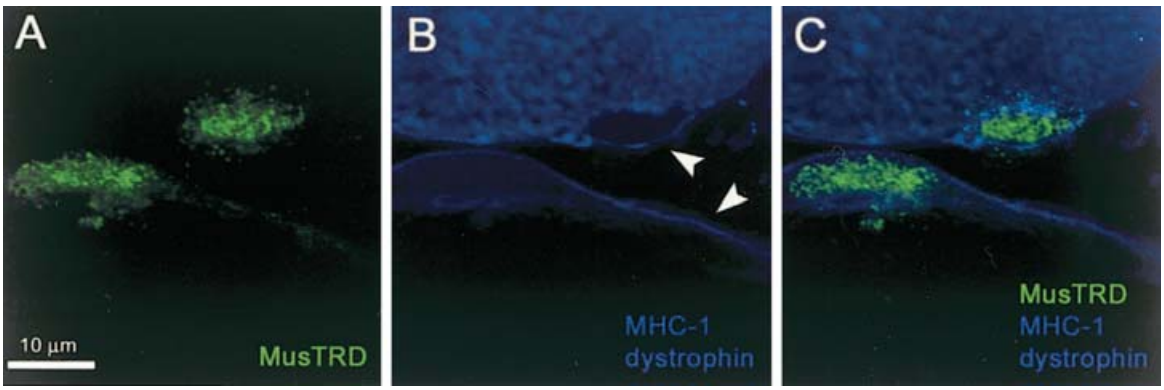

Figure 1 Detection of MusTRD in adult mouse muscle

Confocal images of a mouse soleus muscle showing (**A**) MusTRD expression in the nuclei of two muscle fibres in cross-section. (**B**) Anti-MHC-I staining distinguishes slow- (top; blue cytoplasmic staining) from fast-twitch (bottom) myofibres. Mouse soleus muscle contains only type I and type IIa muscle fibres. Thus all MHC-I negative fibres are fast type IIa. Anti-dystrophin immunostaining delineates the plasma membrane of the muscle fibre (the white arrowheads indicate blue outline of myofibres). (**C**) Merge of (**A**) and (**B**), showing that MusTRD-containing nuclei are within the plasma membrane of myofibres.

repeated stretches of amino acids that are highly homologous with one another called RDs. The cassette containing the 3′ region of exons 25–27 and the 5′ region of exon 28 constitutes the 3′ region of RD5 and the 5′ region of RD6 (Figure 2D). When this cassette is alternatively spliced, the resulting coding sequence contains the 5′ end of RD5 and the 3′ end of RD6, which forms an RD identical in peptide sequence with that of RD6. The isoforms with either of the multi-exon cassettes alternatively spliced have five RDs, similar to the two hMusTRD isoforms [13,19], whereas those with neither alternatively spliced contain six RDs, arranged similarly to those found in TFII-I (Figure 3A) [31]. These RDs bear approx. 70 % similarity to those of TFII-I [32].

Protein motifs in the MusTRD isoforms

A schematic diagram of the 11 mouse MusTRD isoforms is shown in Figure 3(A). The name of the isoform is based on the peptide sequence, and reflects the combination of RDs, C-terminal exon and unique polypeptide cassettes present. The first number indicates the presence of either of these combinations of RDs: 4 + 6 (1), 5 + 6 (2) or 4 + 5 + 6 (3). The isoforms containing exon 30 are termed 'β', whereas those containing exon 31 are termed 'α'. The number that follows indicates the presence of exons 19, 23 and/or 27: specifically, exon 19 alone (1); exon 27 alone (3); exons 19 + 23 (4), exons 19 + 27 (5); or exons 19 + 23 + 27 (7).

Each RD contains a putative helix–loop–helix (HLH)-like motif that could be involved in homodimerization or heterodimerization. Both TFII-I and MusTRD proteins contain a leucine zipper motif at the extreme N-terminus, also believed to be involved in heterodimerization [33]. In addition, the MusTRDs contain a *myc*-type HLH motif involved in heterodimerization [34], and each RD also contains a Leu-Xaa-Xaa-Ile-Leu (LxxIL) motif that is found in most co-activators required for hormone-dependent or -independent nuclear hormone receptor interactions [35,36] (Figure 3B). Interestingly, a unique Ile-Xaa-Xaa-Ile-Ile (IxxII) motif found only in the second RD of MusTRD resembles a core helical motif found in the co-repressors N-CoR and SMRT that is required for nuclear receptor interaction in these proteins [35]. The presence of both types of motifs may allow MusTRD isoforms to play dual roles in the transcriptional regulation of muscle genes. The MusTRDs contain three putative NLSs (nuclear localization signals), one in each of RD2, RD4, and downstream of RD6. Each isoform contains at least two NLSs, whereas the NLS in RD4 is alternatively spliced, and therefore is absent in isoform m2α5. All isoforms, except for m2α5, contain two DBDs (DNA-binding domains), in RD2 and in RD4 (P. Polly, L. M. Haddadi, L. L. Issa, N. Subramaniam, S. J. Palmer, E. S. E. Tay and E. C. Hardeman, unpublished work). The presence of these NLSs and DBDs is consistent with MusTRD's role as a transcriptional regulator located in the nucleus.

Spatial expression of mMusTRD isoforms

The mMusTRD isoforms contain several RDs that are very similar; in particular, RDs 5 and 6 share identical 5′ ends. In isoforms of MusTRD lacking RD5 and containing RDs 4 and 6, the 5′ end of RD6 is spliced to the identical region of RD5 (Figure 2D). Owing to the extent to which sequences in the 3′ region are shared among these isoforms, it is impossible to design primer combinations, cDNA probes or antibodies that exclusively recognize each of the following mMusTRD isoforms m1α1, m1β1, m1α4, m3α5, m3β5, and m3α7. Isoforms m2α5 (lacking RD4), m3α3 (lacking exons 19 and 23, but containing exon 31), m3β3 (lacking exons 19 and 23, but containing exon 30), and m1β4 compared with 3β7 (share exon 23, but can be distinguished from each other by size differences) can be discriminated by RT-PCR.

The presence of these isoforms in various samples representing immature fibres (C_2C_{12}, E13.5 hind-limbs) [37,38], and adult head (EOM), forearm (FCR, FCU, ECU) and hind-limb (plantaris, gastrocnemius, quadriceps, soleus, EDL) muscles was determined using RT-PCR (Figure 4). Amplification of a region common to all mMusTRD isoforms (0–482 bp) using primers 5′ UTR and 482R (Table 1) confirmed that the gene is expressed in all muscle samples tested (Figure 4A). The presence of individual isoforms or subsets thereof was examined using specific RT-PCR primer pairs. Using primer Ex19 in combination with Ex30 or Ex31 resulted in the appearance of several closely grouped bands around 1–2 kb (Figure 4C) and 1.3 kb (Figure 4D) respectively. These groups of bands contain α isoforms m1α1, m1α4, m2α5, m3α5 and m3α7 or β isoforms m1β1, m1β4, m3β5 and m3β7 respectively. The presence of each of these isoforms in amplification reactions was confirmed by the cloning and sequencing of several reactions. Primer combinations Ex20/24

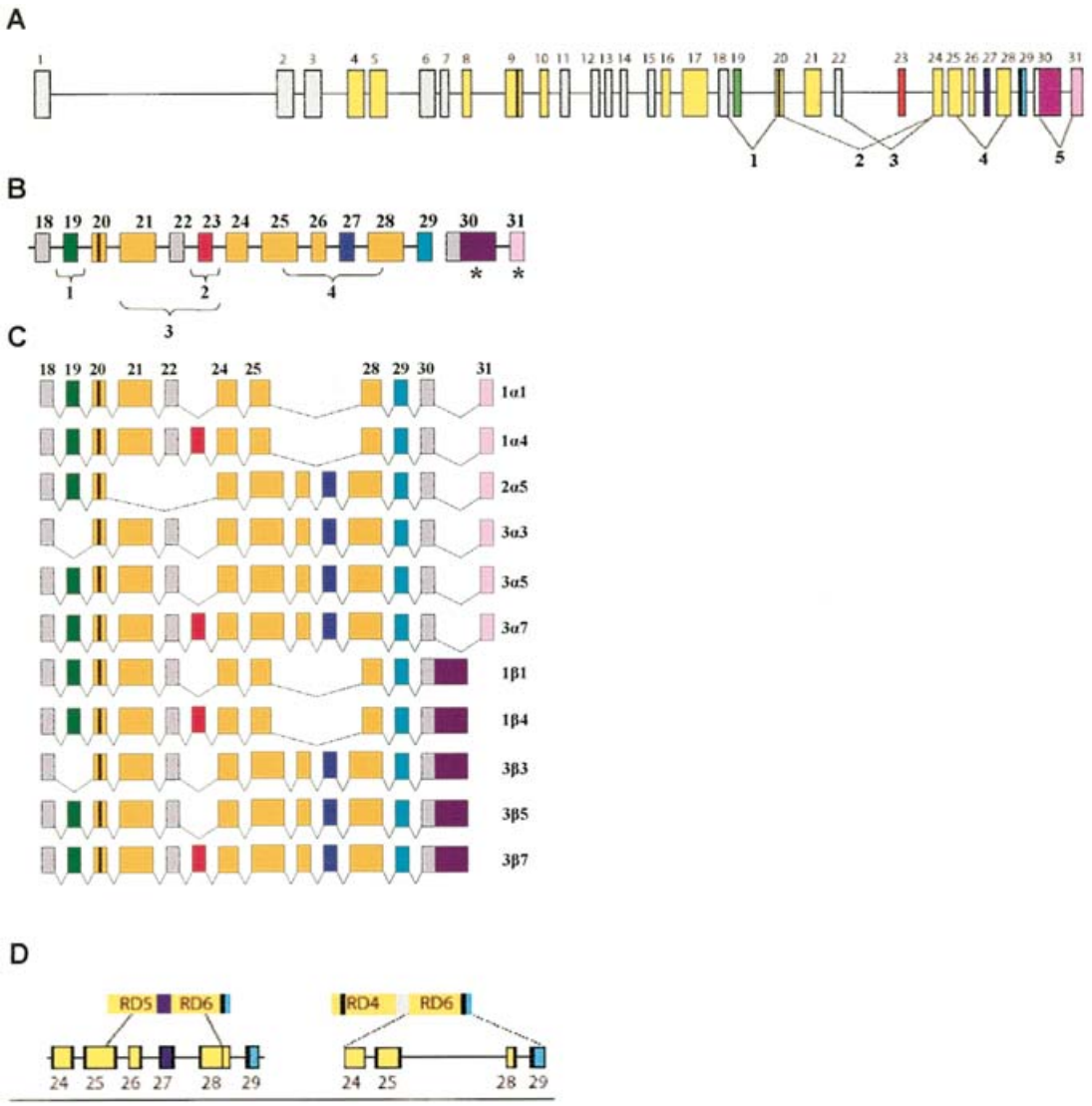

Figure 2 Diagrammatic representation of splicing events in the mouse MusTRD gene (*gtf2ird1*)

(**A**) Schematic representation of *gtf2ird1* showing relative positions of introns and five splicing events in the 3′ region that give rise to the 11 isoforms. (**B**) Exons 18–31 showing the four alternatively spliced cassettes (1–4) and the mutually exclusively spliced regions (5′ region of 30 and exon 31) marked with asterisks. (**C**) Alternative splicing events in the 3′ region of *gtf2ird1* (exons 18–31) that give rise to the 11 mMusTRD isoforms. (**D**) Cassette 4, which constitutes the 5′ region of RD5 and the 3′ region of RD6 in isoforms m2α5, m3α3, m3β3, m3α5, m3α7, m3β5 and m3β7, and the exons that form RD6 in isoforms m1α1, m1α4, m1β1 and m1β4 when this cassette is removed by alternative splicing. Colour-coding for the exons is as follows: yellow, exons present in RDs; green, exon 19; red, exon 23; dark blue, exon 27; light blue, polyserine tract; black, NLSs; purple, exon 30; pink, exon 31.

and Ex31, Ex18/20 and Ex31, Ex18/20 and Ex30 were used to specifically amplify isoforms m2α5, m3α3 and m3β3 respectively (shown in Figures 4B, 4E and 4F respectively). Comparison of Figures 4(C) and 4(D) shows that all samples express both α and β isoforms. Although differences are seen in the relative intensity of bands between different samples in each panel, no quantitative conclusions can be drawn from Figure 4. We conclude from these data that specific isoforms or sets of isoforms are present in each muscle.

In order to provide quantitative assessment of MusTRD isoform expression, real-time RT-PCR reactions were conducted as described in the Experimental section. Expression of total mMusTRD is very high in E13.5 hind-limb and moderately high in EDL, ECU and EOM muscles (Figure 5A). However, the expression of individual isoforms does not, in all cases, follow the same pattern as total MusTRD expression (Figures 5B–5F). The relative expression of mMusTRD2α5 in C2C12 cells is comparable with that in EOM, although the total MusTRD expression in EOM is almost 8-fold that of C2C12 (Figures 5A and 5B). The expression of some of the individual isoforms in the FCR is significantly higher than in the ECU, though the ECU has a far higher expression of total MusTRD (Figures 5A–5F). Therefore some MusTRD isoforms show a degree of spatial regulation.

Regulation of muscle contractile protein promoters/enhancers by two MusTRD isoforms

In order to test the functional capacity of MusTRD isoforms in skeletal muscle, we compared the ability of four isoforms to regulate three muscle slow fibre-specific promoters/enhancers. Pairs of isoforms were chosen that differ by a single alternative

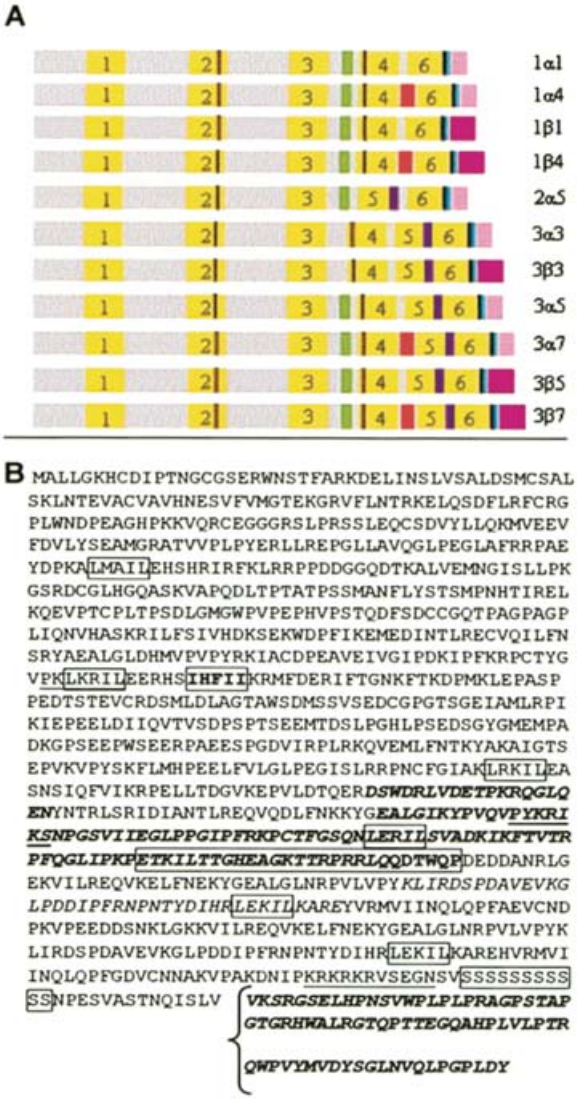

Figure 3 Schematic diagram of 11 mMusTRD proteins (A) and amino acid sequence of MusTRD arising from all possible coding sequences (B)

(**A**) RDs are shown in yellow; polypeptide cassette regions are shown in different colours (exon 19 is green; exon 23 is red; exon 27 is dark blue); the polyserine tract is shown in light blue; NLSs are black. The two mutually exclusive regions at the extreme 5′-termini (5′ region of exons 30 and 31) are indicated in purple and pink respectively. (**B**) Peptide sequences arising from alternatively spliced cassettes are shown italicized and in bold. The alternatively spliced, mutually exclusive 5′ regions of exon 30 and exon 31 are shown in brackets. The amino acid sequence encoded by exon 23 is boxed to distinguish it from the larger cassette. Also shown are the six putative LxxIL motifs sites (boxed), IxxII motif (boxed and emboldened), NLSs (underlined) and the polyserine tract (boxed).

exon at the C-terminus. Expression plasmids encoding h1α1, m1β1, m3α7 or m3β7 were co-transfected with luciferase reporter plasmids containing promoter/enhancer elements from the *MHC-I*, myosin light chain 2 slow (*MLC2*$_{slow}$) and *TnI*$_{slow}$ genes (Figure 6). The core binding motif or Inr-like element (GATTAA) that was defined for 1α1 [13] is present in the MHC-I and TnI$_{slow}$ elements, but not in MLC2$_{slow}$. h1α1 repressed both enhancer elements that contain its binding site and did not regulate MLC2$_{slow}$. m1β1, which differs from 1α1 by only one exon (30, cf. 31), did not regulate any of the fibre-specific elements. Similarly, isoforms m3α7 and m3β7 were compared. The regulatory capabilities of these two isoforms differed one

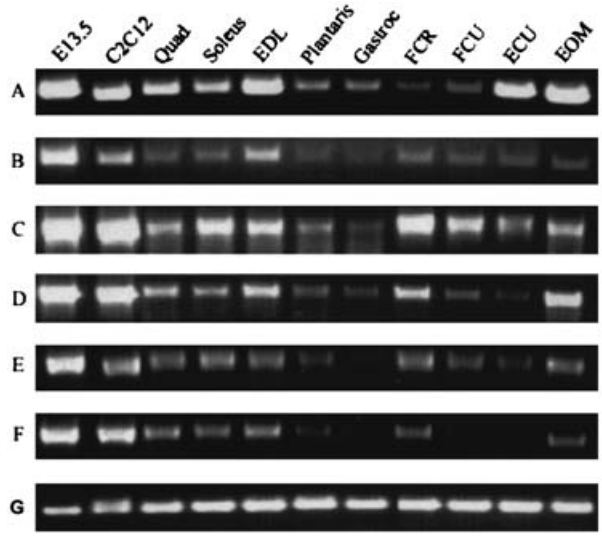

Figure 4 RT-PCR analysis of MusTRD isoform expression in skeletal muscle

RT-PCR was performed on RNA isolated from differentiated C2C12 myotubes (C2C12) and mouse 13.5 days post-coitus embryonic hind-limbs (E13.5), adult quadriceps (Quad), soleus, EDL, plantaris, gastrocnemius (Gastroc), FCR, FCU, ECU and EOM muscles and the PCR products were run on agarose gels. Different primer combinations were used to amplify (**A**) all MusTRD isoforms, (**B**) m2α5, (**C**) m1α1, m1α4, m3α5 and m3α7, (**D**) m1β1, m1β4, 3β5 and m3β7, (**E**) m3α3, (**F**) m3β3 and (**G**) GAPDH.

from the other for each slow fibre-specific element tested. m3β7, but not m3α7, expression significantly repressed MHC-I and m3α7, but not m3β7, significantly repressed TnI$_{slow}$. m3α7 expression significantly elevated activity from the MLC2$_{slow}$ enhancer, whereas m3β7 had no significant effect. We conclude that different MusTRD isoforms can elicit different responses from the same gene element. At least one isoform, m3α7, can act in the absence of the cognate MusTRD binding site. In addition, a single isoform of MusTRD is capable of acting as both an activator and a repressor of muscle-specific genes, depending upon the promoter/enhancer, and alternative splicing of a single exon can modify that activity.

DISCUSSION

Structure–function of the mMusTRD family

The MusTRD isoforms are members of a novel gene family that includes the TFII-I family of transcription factors. This family is characterized by a signature arrangement of homologous RDs containing an HLH-like motif, speculated to be involved in protein–protein interactions [14]. The presence of these HLH-like motifs, a leucine zipper and DBDs, in addition to the *myc*-like HLH motif in MusTRD, implicate these proteins in DNA–protein interactions, as well as multiple interactions with other proteins. Even though motifs may be conserved, it does not mean that they serve the same function in the two gene families. Both MusTRD and TFII-I have a putative leucine zipper at amino acid positions 23–44 in the first RD that may be involved in DNA binding. Cheriyath and Roy [31] have shown that DNA binding by TFII-I requires the first 90 amino acids. In contrast, MusTRD isoforms have two DBDs located in the N- and C-termini between amino acids 408–420 (DBD1) and 738–765 (DBD2) respectively (P. Polly, L. M. Haddadi, L. L. Issa, N. Subramaniam, S. J. Palmer, E. S. E. Tay and E. C. Hardeman, unpublished work).

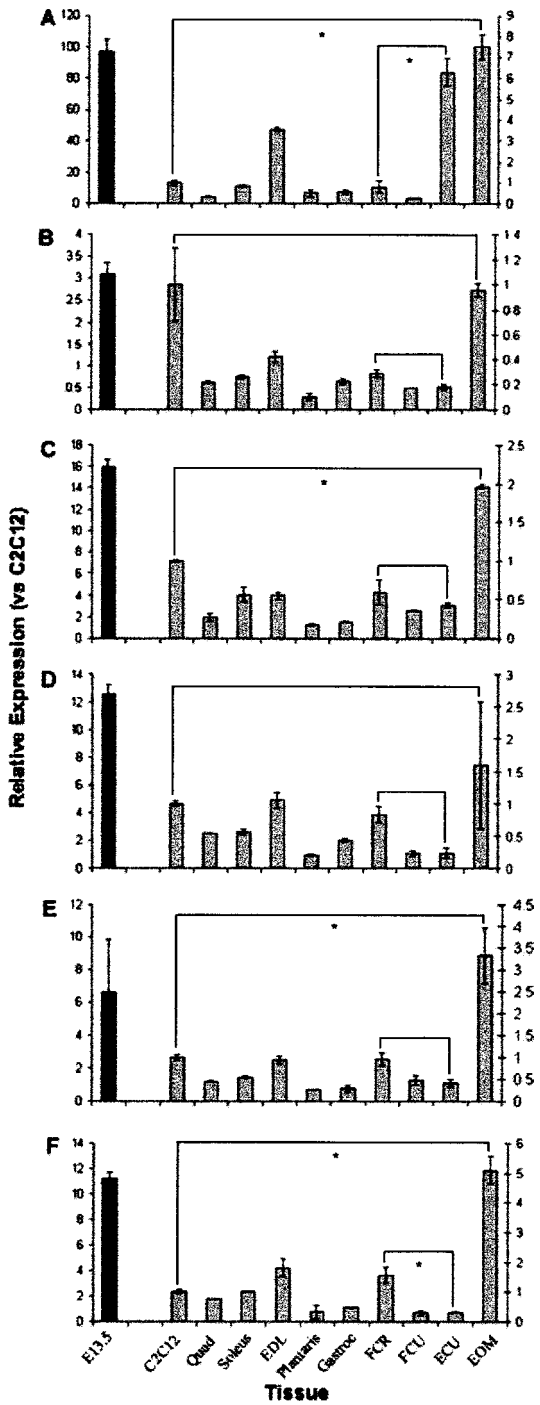

Figure 5 Real-time RT-PCR analysis of MusTRD isoform expression

Expression of (**A**) all mMusTRD isoforms, (**B**) m2α5, (**C**) 3α5, (**D**) 3β5, (**E**) 3α3 and (**F**) 3β3 was quantified from duplicate RT-PCR reactions performed on RNA isolated from differentiated C2C12 myotubes (C2C12) and mouse 13.5 days post-coitus embryonic hind-limbs (E13.5), adult quadriceps (Quad), soleus, EDL, plantaris, gastrocnemius (Gastroc), FCR, FCU, ECU and EOM muscles. Values are normalized to a GAPDH control and expressed relative to the expression in C2C12 cultures (set at 1). Relative E13.5 levels are plotted against the left *y*-axis, whereas those from all other sources are plotted against the right *y*-axis. $^*P < 0.05$.

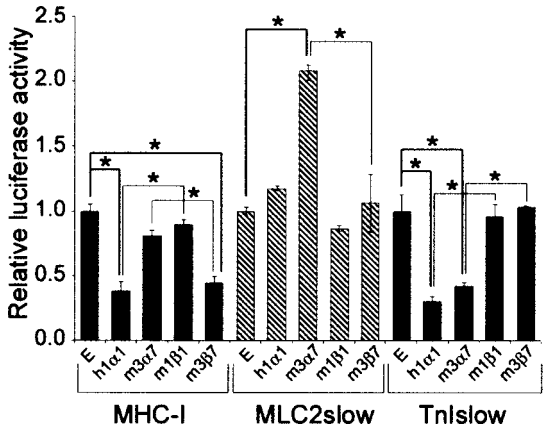

Figure 6 Regulation of slow and fast fibre-specific promoters by MusTRD isoforms in C2C12 cells

h1α1, m3α7, m1β1 or m3β7 expression plasmids were co-transfected with expression plasmids containing either -3500 to $+462$ of the *MHC-I* gene, -800 to $+12$ of the $MLC2_{slow}$ gene, or the 157 bp TnI_{slow} enhancer linked to luciferase. Luciferase activity was determined in lysates from three different transfection cultures per plasmid combination. MusTRD-mediated repression or activation of transcriptional activity is expressed as the fold difference from the basal/empty construct activity that was set at 1. Values shown represent the means $\pm$ S.D. from three separate experiments. E, empty vector (pcDNA3.1 backbone vector). $^*P < 0.05$.

Table 2 Detection of mMusTRD isoforms from different tissue samples

mMusTRD isoforms were detected and identified in a gt11 quad library (Clontech) and in total RNA isolated from different tissues using RT-PCR, cloning and sequencing. RT-PCR using forward primers Ex19, Ex18/20, Ex23, Ex22/24 and Ex20/24 were used in combination with reverse primers Ex30 and Ex31 to identify individual isoforms. '+' indicates the presence of isoform; ND, not detected.

	mMusTRD isoform										
	1α1	1α4	2α5	3α3	3α5	3α7	1β1	1β4	3β3	3β5	3β7
E13.5	ND	ND	+	+	+	+	ND	ND	+	+	+
C2C12	+	+	+	+	+	+	ND	ND	+	+	+
Quadriceps	ND	ND	+	+	+	+	+	+	+	+	+
Soleus	ND	ND	+	+	+	+	ND	ND	+	+	+
EDL	ND	ND	+	+	+	+	ND	ND	+	+	+
Plantaris	ND	ND	+	+	+	+	ND	ND	+	+	+
Gastrocnemius	ND	ND	+	+	+	+	ND	ND	+	+	+
FCR	ND	ND	+	+	+	+	ND	ND	+	+	+
FCU	ND	ND	+	+	+	+	ND	ND	+	+	+
ECU	ND	ND	+	+	+	+	ND	ND	+	+	+
EOM	ND	ND	+	+	+	+	ND	ND	+	+	+

Although MusTRD and TFII-I share a high level of homology in several regions, alternative splicing patterns bear no similarity in pattern or region. Only two cassette exons are alternatively spliced in TFII-I, exons 'a' and 'b', located between RD1 and RD2. 'Cassette-style' alternative splicing of these exons gives rise to four human isoforms, three homologues of which are present in the mouse [39]. In contrast, the alternative splicing of mMusTRD is far more variable, involving four cassette regions and two mutually exclusive regions, resulting in 11 isoforms present in skeletal muscle alone. The relatively narrow range of physical variation observed in TFII-I compared with MusTRD suggests that these two related transcription factors do not share completely overlapping roles in transcriptional regulation.

There are currently two known hMusTRD isoforms, which are 944 amino acids and 959 amino acids in size. The variation between them arises from the alternative splicing of exon 19 [13,19], which is also alternatively spliced in the mouse. The identification of 11 mouse MusTRD isoforms with highly variable C-terminal ends suggests that this is a complex family of transcriptional regulators. The m3α7 isoform is identical with BEN1, which

was isolated from a brain cDNA library and shown to bind the early enhancer region of the *Hoxc8* gene [22]. It has also been shown that the C-terminus of Cream1 is capable of binding retinoblastoma [19]. Together with the data in this manuscript demonstrating differential regulation of muscle fibre-specific genes by MusTRD isoforms, these data support the hypothesis that MusTRD isoforms have multiple roles in transcriptional regulation that is mediated, at least in part, by alternative splicing in the C-terminal half of the molecule.

In addition to facilitating interactions with other proteins, HLH domains have been shown to facilitate interaction between regions of the same protein. This may result in varying tertiary structures between MusTRD isoforms with five RDs and those with six RDs. As well as possible variations in affinity of each isoform for other proteins, these mMusTRD isoforms may also interact with each other, producing a large range of homodimers and/or heterodimers with the ability to recognize a diverse range of DNA motifs. The presence of 11 different isoforms indicates the potential for a wide range of tertiary and quaternary structures with affinity for numerous transcriptional co-regulators and DNA motifs, giving rise to a family with a high level of functional variability. This is supported by the observation that a mutually exclusive alternative splice at the C-terminus can determine the functional capacity of two MusTRD isoforms.

Expression patterns of the mMusTRD family

The classification of MusTRD1/BEN/GTF3/XWBSCR11 as a single, ubiquitously expressed transcription factor in some studies is on the basis of detection methods that recognize multiple, if not all, isoforms from this gene [15,16,33,40]. With the discovery of at least 11 isoforms present in skeletal muscle alone, and with the potential for 24 splice variants in total, it is evident that care must be taken when defining expression patterns and function. We have determined that several individual isoforms (m2α5, m3α3 and m3β3) and groups of these isoforms (m1α1, m1α4, m3α5 and m3α7, and m1β1, m1β4, m3β5 and m3β7) can be discriminated using RT-PCR, but this does not allow unambiguous measurements of each individual isoform. Real-time RT-PCR analysis shows a wide range of total MusTRD expression between different tissues and at different developmental time points. However, this expression profile of different isoforms varies between different muscle samples, such as the EOM and C2C12 cells and FCR and ECU (Figures 5B–5F). These variations demonstrate that the regulation of expression of individual isoforms is independent of the regulation of total MusTRD expression, because different tissues show differential expression patterns for individual isoforms. Variations in MusTRD isoform regulation between muscles, such as the FCR and ECU, may result from different requirements in controlling fibre type-specific gene expression. However, gene knock-out studies will be required to test whether individual isoforms make specific contributions to the regulation of individual muscle fibre types.

Differential regulation of skeletal muscle gene promoters/enhancers

The four isoforms tested in co-transfection studies had differential regulatory effects on the promoters/enhancers of three muscle slow fibre-specific genes. This study has revealed that the regulatory properties of these isoforms may be due to specific combinations of protein regions, rather than a single signature motif in each isoform. For example, isoform m3α7 stimulates $MLC2_{slow}$, whereas m3β7 does not. Since m3α7 differs from m3β7 by the C-terminal exon only, the impact of m3α7 on $MLC2_{slow}$ expression is due to this single element. Additionally, h1α1 and m1β1 contain the same motifs, except for the C-termini. h1α1 represses the TnI_{slow} enhancer, whereas m1β1 has no effect. In addition, this study has revealed that the sequences present in a specific gene dictate the regulatory capability of a MusTRD isoform. The presence of both IxxIL and IxxII motifs in these isoforms may allow them to behave as both activators and repressors under different circumstances, as seen by the ability of m3α7 to activate $MLC2_{slow}$, but repress TnI_{slow}. h1α1 represses *MHC-I* and TnI_{slow}, but does not regulate $MLC2_{slow}$. *MHC-I* and TnI_{slow} contain the Inr-like element that binds h1α1, and this is absent in $MLC2_{slow}$. This provides a means for regulating muscle slow fibre-specific genes independently, as well as in a co-ordinated manner.

Considering the variability in the C-terminal half of these 11 mouse isoforms and in their expression patterns, it is possible that they may have different roles in regulating the expression patterns of various muscle-specific genes and, therefore, in fibre type determination. The fact that they co-exist in different muscles adds to the potential diversity and complexity of their regulatory capacities. Further studies of the effect on gene expression exerted by these isoforms and the biochemical analysis of its protein partners, in addition to knock-out studies, will lead to a greater understanding of the function of MusTRD in regulating gene expression, its role in defining muscle fibre type and potential pathology of its haploinsufficiency in WBS.

Alternative pre-mRNA splicing can create a family of proteins with related, but varying, biochemical properties from a single locus. It is not unexpected, therefore, to find that MusTRD isoforms have differing levels of activity on the target enhancer of *TnIslow*. Quantitative RT-PCR expression analysis suggests that variations in *Gtf2ird1* transcripts do occur between muscle groups. Therefore it is possible that modulation of the MusTRD isoform content within muscle fibres provides a means of differentially regulating downstream target genes in muscles of different fibre composition. It is widely accepted that alternative pre-mRNA splicing is the most important way of achieving greater protein diversity from the surprisingly small numbers of mammalian genes [41,42]. For the MusTRDs, the extent of isoform complexity could relate to the diversity of genes involving MusTRD regulation and/or the diversity of co-regulatory proteins with which the MusTRDs interact.

We thank C. Stephen Robinson and Kata Popovic for excellent technical assistance, Gregory Pelka, Anthony Kee and Nan Yang for many helpful discussions, Stephen Palmer and Mark Corbett for a critical reading of the manuscript, and James Newcombe and Mehan Siriwardhane for assistance in screening the λ library. P. P. is the recipient of a Howard Florey Centenary fellowship from the National Health and Medical Research Council (NH&MRC) of Australia. This work was supported by an NH&MRC grant to E. C. H.

REFERENCES

1 Gunning, P. and Hardeman, E. (1991) Multiple mechanisms regulate muscle fibre diversity. FASEB J. **5**, 3064–3070
2 Schiaffino, S. and Reggiani, C. (1996) Molecular diversity of myofibrillar proteins: gene regulation and functional significance. Physiol. Rev. **76**, 371–423
3 Hughes, S. M. and Salinas, P. C. (1999) Control of muscle fibre and motoneuron diversification. Curr. Opin. Neurobiol. **9**, 54–64
4 Nikovits, Jr, W., Cann, G. M., Huang, R., Christ, B. and Stockdale, F. E. (2001) Patterning of fast and slow fibres within embryonic muscles is established independently of signals from the surrounding mesenchyme. Development **128**, 2537–2544
5 Esser, K. A. and Hardeman, E. C. (1995) Changes in contractile protein mRNA accumulation in response to spaceflight. Am. J. Physiol. **268**, C466–C471
6 Pette, D. and Staron, R. S. (2001) Transitions of muscle fiber phenotypic profiles. Histochem. Cell Biol. **115**, 359–372
7 Criswell, D. S., Hodgson, V. R., Hardeman, E. C. and Booth, F. W. (1998) J. Appl. Physiol. **84**, 1083–1087

8 Booth, F. W. and Baldwin, K. M. (1996) Muscle plasticity: energy demanding and supply processes. In Handbook of Physiology (Peachey, L. D., Adrian, R. H. and Geiger, S. R., eds.), pp. 1075–1123, Williams and Wilkins, Baltimore
9 Buckingham, M. (2001) Skeletal muscle formation in vertebrates. Curr. Opin. Genet. Dev. **11**, 440–448
10 Chin, E. R., Lindinger, M. I. and Heigenhauser, G. J. (1997) Distribution of lactate and other ions in inactive skeletal muscle: influence of hyperkalemic lactacidosis. Can. J. Physiol. Pharmacol. **75**, 1375–1386
11 Serrano, A. L., Murgia, M., Pallafacchina, G., Calabria, E., Coniglio, P., Lomo, T. and Schiaffino, S. (2001) Calcineurin controls nerve activity-dependent specification of slow skeletal muscle fibers but not muscle growth. Proc. Natl. Acad. Sci. U.S.A. **98**, 13108–13113
12 Murgia, M., Serrano, A. L., Calabria, E., Pallafacchina, G., Lomo, T. and Schiaffino, S. (2000) Ras is involved in nerve-activity-dependent regulation of muscle genes. Nat. Cell. Biol. **2**, 142–147
13 O'Mahoney, J. V., Guven, K. L., Lin, J., Joya, J. E., Robinson, C. S., Wade, R. P. and Hardeman, E. C. (1998) Identification of a novel slow-muscle-fiber enhancer binding protein, MusTRD1. Mol. Cell. Biol. **18**, 6641–6652
14 Roy, A. L. (2001) Biochemistry and biology of the inducible multifunctional transcription factor TFII-I. Gene **274**, 1–13
15 Osborne, L. R., Campbell, T., Daradich, A., Scherer, S. W., Tsui, L. C., Franke, Y., Peoples, R. J., Francke, U., Voit, T., Kramer, H. et al. (1999) Identification of GTF2IRD1, a putative transcription factor within the Williams–Beuren syndrome deletion at 7q11.23. Genomics **57**, 279–284
16 Franke, Y., Peoples, R. J. and Francke, U. (1999) Identification of GTF2IRD1, a putative transcription factor within the Williams–Beuren syndrome deletion at 7q11.23. Cytogenet. Cell. Genet. **86**, 296–304
17 Tassabehji, M., Carette, M., Wilmot, C., Donnai, D., Read, A. P. and Metcalfe, K. (1999) A transcription factor involved in skeletal muscle gene expression is deleted in patients with Williams syndrome. Eur. J. Hum. Genet. **7**, 737–747
18 Voit, T., Kramer, H., Thomas, C., Wechsler, W., Reichmann, H. and Lenard, H. G. (1991) Myopathy in Williams–Beuren syndrome. Eur. J. Pediatr. **150**, 521–526
19 Yan, X., Zhao, X., Qian, M., Guo, N., Gong, X. and Zhu, X. (2000) Characterization and gene structure of a novel retinoblastoma-protein-associated protein similar to the transcription regulator TFII-I. Biochem. J. **345**, 749–757
20 Bayarsaihan, D., Dunai, J., Greally, J. M., Kawasaki, K., Sumiyama, K., Enkhmandakh, B., Shimizu, N. and Ruddle, F. H. (2002) Genomic organization of the genes Gtf2ird1, Gtf2i, and Ncf1 at the mouse chromosome 5 region syntenic to the human chromosome 7q11.23 Williams syndrome critical region. Genomics **79**, 137–143
21 Vullhorst, D. and Buonanno, A. (2002) Molecular dissection of DNA sequences and factors involved in slow muscle-specific transcription. J. Biol. Chem. **275**, 26300–26308
22 Bayarsaihan, D. and Ruddle, F. H. (2000) Isolation and characterization of BEN, a member of the TFII-I family of DNA-binding proteins containing distinct helix–loop–helix domains. Proc. Natl. Acad. Sci. U.S.A. **97**, 7342–7347
23 Ring, C., Ogata, S., Meek, L., Song, J., Ohta, T., Miyazono, K. and Cho, K. W. (2002) The role of a Williams-Beuren syndrome-associated helix–loop–helix domain-containing transcription factor in activin/nodal signaling. Genes Dev. **16**, 820–835
24 Borrione, A. C., Zanellato, A. M., Saggin, L., Mazzoli, M., Azzarello, G. and Sartore, S. (1988) Neonatal myosin heavy chains are not expressed in Ni-induced rat rhabdomyosarcoma. Differentiation **38**, 49–59
25 Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. Anal. Biochem. **162**, 156–159
26 di Maso, N. A., Haddad, F., Zeng, M., McCue, S. A. and Baldwin, K. M. (2000) Role of denervation in modulating IIb MHC gene expression in response to T(3) plus unloading state. J. Appl. Physiol. **88**, 682–689
27 Wright, C. E., Bodell, P. W., Haddad, F., Qin, A. X. and Baldwin, K. M. (2001) In vivo regulation of the β-myosin heavy chain gene in hypertensive rodent heart. Am. J. Physiol. Cell. Physiol. **280**, C1262–C1276
28 Esser, K., Nelson, T., Lupa-Kimball, V. and Blough, E. (1999) The CACC box and myocyte enhancer factor-2 sites within the myosin light chain 2 slow promoter cooperate in regulating nerve-specific transcription in skeletal muscle. J. Biol. Chem. **274**, 12095–12102
29 Yaffe, D. and Saxel, O. (1977) Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. Nature (London) **270**, 725–727
30 Blau, H. M., Chiu, C. P. and Webster, C. (1983) Cytoplasmic activation of human nuclear genes in stable heterocaryons. Cell **32**, 1171–1180
31 Cheriyath, V. and Roy, A. L. (2001) Structure–function analysis of TFII-I. Roles of the N-terminal end, basic region and I-repeats. J. Biol. Chem. **276**, 8377–8383
32 Roy, A. L., Du, H., Gregor, P. D., Novina, C. D., Martinez, E. and Roeder, R. G. (1997) Cloning of an inr- and E-box-binding protein, TFII-I, that interacts physically and functionally with USF1. EMBO J. **16**, 7091–7104
33 Tussie-Luna, M. I., Bayarsaihan, D., Ruddle, F. H. and Roy, A. L. (2001) Repression of TFII-I-dependent transcription by nuclear exclusion. Proc. Natl. Acad. Sci. U.S.A. **98**, 7789–7794
34 Suzuki, M., Okuyama, S., Okamoto, S., Shirasuna, K., Nakajima, T., Hachiya, T., Nojima, H., Sekiya, S. and Oda, K. (1998) A novel E2F binding protein with Myc-type HLH motif stimulates E2F-dependent transcription by forming a heterodimer. Oncogene **17**, 853–865
35 Sauve, F., McBroom, L. D., Gallant, J., Moraitis, A. N., Labrie, F. and Giguere, V. (2001) CIA, a novel estrogen receptor coactivator with a bifunctional nuclear receptor interacting determinant. Mol. Cell. Biol. **21**, 343–353
36 Li, D., Wang, F. and Samuels, H. H. (2001) Domain structure of the NRIF3 family of coregulators suggests potential dual roles in transcriptional regulation. Mol. Cell. Biol. **21**, 8371–8384
37 Sutherland, C. J., Esser, K. A., Elsom, V. L., Gordon, M. L. and Hardeman, E. C. (1993) Identification of a program of contractile protein gene expression initiated upon skeletal muscle differentiation. Dev. Dyn. **196**, 25–36
38 Ontell, M. and Kozeka, K. (1984) The organogenesis of murine striated muscle: a cytoarchitectural study. Am. J. Anat. **171**, 133–148
39 Cheriyath, V. and Roy, A. L. (2000) Alternatively spliced isoforms of TFII-I. Complex formation, nuclear translocation and differential gene regulation. J. Biol. Chem. **275**, 26300–26308
40 Calvo, S., Vullhorst, D., Venepally, P., Cheng, J., Karavanova, I. and Buonanno, A. (2001) Molecular dissection of DNA sequences and factors involved in slow muscle-specific transcription. Mol. Cell. Biol. **21**, 8490–8503
41 Claverie, J. M. (2001) Gene number. What if there are only 30,000 human genes? Science **291**, 1255–1257
42 Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A. et al. (2001) The sequence of the human genome. Science **291**, 1304–1351

Received 30 January 2003/12 May 2003; accepted 3 June 2003
Published as BJ Immediate Publication 3 June 2003, DOI 10.1042/BJ20030189