Molecular and functional analysis of the utrophin promoter

Carina L. Dennis, Jonathon M. Tinsley, Anne E. Deconinck and Kay E. Davies*

Genetics Laboratory, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK

Received January 29, 1996; Revised and Accepted March 19, 1996 EMBL accession nos X95523 and X95524

ABSTRACT

Utrophin is a ubiquitously expressed cytoskeletal protein which is an important structural component of the mammalian neuromuscular junction. It shows extensive sequence similarity to dystrophin leading to postulation that utrophin may be able to compensate for the absence of dystrophin in Duchenne muscular dystrophy (DMD) patients. In order to study the transcriptional control of utrophin expression including its regulation at the neuromuscular junction, and as a first step in the development of a potential DMD therapy, we have cloned the utrophin promoter region from human and mouse. The utrophin promoter is associated with a CpG island at the 5′-end of the gene, and sequence analysis of the 5′-UTR reveals several Sp1 binding sites and the absence of TATA or CAAT motifs. Transcription is initiated at one major and three minor sites. Using deletion constructs, we have defined an active promoter region of 155 bp. The first exon and 900 bp upstream display limited sequence conservation between human and mouse. The core sequence TTCCGG of the N box which regulates synaptic expression of other genes is also present and may be involved in regulating the specific expression of utrophin at the postsynaptic membrane. This study provides the basis for the understanding of the regulatory mechanism that controls utrophin expression and provides the data needed to develop methods for the upregulation of utrophin in DMD patients.

INTRODUCTION

Duchenne muscular dystrophy (DMD) is an X-linked recessive disorder caused by mutations which result in the absence or expression of mutant forms of dystrophin. Dystrophin is a 427 kDa cytoskeletal protein expressed predominately in skeletal, cardiac and smooth muscle, with lower levels in the brain. In normal adult skeletal muscle, dystrophin binds to the sarcolemma by interaction with the dystrophin protein complex (DPC) (1–4). The DPC forms an essential link between the internal cytoskeleton of the muscle cell and the extracellular matrix. In DMD, the absence of dystrophin results in the loss of integrity of the DPC which eventually leads to muscle degeneration. Thus any effective therapeutic strategy for DMD must involve the reconstitution of this protein complex. Many strategies for DMD therapy have involved the introduction of dystrophin minigenes into muscle using viral vectors or direct injection (5). However the efficiency of gene delivery using these methods is relatively inefficient. An alternative approach is to search for related genes which might compensate for the loss of dystrophin.

Utrophin is an autosomally-encoded protein displaying a high degree of sequence similarity to dystrophin (6). The differences in the function of utrophin and dystrophin may lie within their regulatory sequences rather than the primary coding sequence. Although utrophin is expressed in muscle, its overall expression pattern differs from dystrophin. Dystrophin expression is restricted to adult muscle and brain whereas utrophin is widely expressed. In normal adult skeletal muscle, utrophin is localised to the neuromuscular junction (NMJ) and myotendinous junction, however in dystrophin-deficient muscle, utrophin is also localised at the sarcolemma and co-purifies with components of the DPC (7). In some inflammatory myopathies, both dystrophin and utrophin can be found localised to the sarcolemma of mature fibres (8,9). In normal foetal muscle, utrophin localises at the sarcolemma and utrophin levels decrease as dystrophin levels increase during development, suggesting that the two proteins are co-ordinately regulated during muscle development (10,11). The fact that utrophin can in certain circumstances be localised at the sarcolemma suggests that there may be conditions under which utrophin could be relocated to the muscle membrane in DMD patients.

In normal adult muscle, utrophin co-localises with agrin-induced acetylcholine receptor (AChR) clusters at the NMJ (12). It is widely assumed that utrophin links the extracellular agrin-bound DPC to the submembranous actin cytoskeleton (13). The role of utrophin at the NMJ may be to stabilise the mature AChR clusters at the postsynaptic membrane, rather than in the initial stages of AChR cluster formation (14). Although utrophin may be required for the normal development and functioning of the postsynaptic membrane, nothing is known about the mechanism that controls the specific localisation of utrophin at the NMJ or the regulation of utrophin expression.

As yet, no promoter has been characterised for the utrophin transcript. If upregulation of utrophin to replace dystrophin does become a feasible strategy for DMD therapy, then a full characterisation of the utrophin promoter is essential. Here we describe the isolation and characterisation of a utrophin promoter associated with the CpG island at the 5′-end of the gene which drives the expression

* To whom correspondence should be addressed
of the full-length transcript and contains a sequence motif which may drive expression specifically at the NMJ.

**MATERIALS AND METHODS**

**Isolation, characterisation and sequencing of genomic clones**

YAC 4X23E3 was partially digested with the restriction endonuclease MboI, ligated to λGEM-11 XhoI arms and packaged using the Packagene system (Promega). The mouse genomic phage library screened to isolate the mouse utrophin promoter was kindly supplied by Dr D. Picketts, Institute of Molecular Medicine, Oxford. Positive hybridising phage were further subcloned into pGEM-7Zf(+) plasmid (Promega). Sequencing of double stranded plasmid DNA was performed using Sequenase v2.0 (USB).

**Primer extension**

Extension assays were performed using the AMV Reverse Transcriptase Primer Extension System (Promega). Primer U25 (′5′-AGG CAC CAA CTT TGC CAA ACCG, 1067–1088) was end-labelled by kination in the presence of [γ-32P]ATP, 3000 Ci/mmol (ICN). Primer (100 fmol) was annealed to total RNA (30–60 µg) at 58°C for 20 min then extended at 42°C for 30 min. The products were separated on a 6% polyacrylamide gel under denaturing conditions. A sequencing ladder was run simultaneously for sizing extension products.

**5′ RACE analysis**

5′ RACE was carried out using the 5′ AmpliFINDER RACE kit (Clontech). IN157 poly(A)+ RNA (2 µg), prepared using Dynabeads mRNA purification kit (Dynal), was reverse transcribed using primer U25. The anchor-ligated cDNA was then PCR amplified using the anchor primer and an internal primer, U24 (′5′-AAT CGG CTT CTG GAG CCA GAG, 982–1002). PCR amplified products were cloned into pGEM-T vector (Promega) and sequenced using Sequenase v2.0 (USB).

**RNase protection**

Genomic fragments used as RNase protection probes were cloned into pGEM-7Zf(+) (Promega). The C-terminal control probe for Genomic fragments used as RNase protection probes were cloned and sequenced using Sequenase v2.0 (USB).

**Transfections and luciferase estimation**

IN157, HeLa, COS-7 and C2 cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% foetal calf serum (FCS). Cells were grown to 80–90% confluency, trypsinised, washed once with phosphate buffered saline (PBS) and resuspended (3.75 × 10^7 cells/ml) in ice cold PBS. Transient transfections were carried out using 30 µg luciferase reporter construct (promoter fragment cloned into pGGL2-basic; Promega) and 10 µg pSV-β-galactosidase plasmid (Promega). Cells (3 × 10^6) were electroporated at 250 V, 960 µF using a BioRad Gene Pulser. Following electroporation, the samples were placed on ice for 1 min and aliquoted into a 60 mm diameter Petri dishes containing 5 ml of DMEM–10% FCS. Cells were allowed to recover at 37°C for 12 h (HeLa), 24 h (IN157 and COS-7) or 48 h (C2 myoblasts). Cells were washed twice with PBS, harvested into 400 µl Reporter Lysis Buffer (Promega). Cell extract (20 µl) was mixed with 100 µl Luciferase assay reagent (Promega) and light production quantified using a Turner Designs Model 20 luminometer. β-Galactosidase activity was measured using an enzyme assay system (Promega) then analysed using a microplate reader (BioRad) at 420 nm.

**RESULTS**

**Isolation of the promoter region of the utrophin gene**

Yeast artificial chromosome (YAC) clone 4X23E3 encompasses a CpG-island located at the 5′-end of the human utrophin gene (14). The YAC was subcloned into phage and the resultant library screened with a 0.6 kb EcoRI cDNA fragment containing the first exon of utrophin. Positive hybridising phage were further subcloned resulting in a plasmid clone (pPU1) containing a hybridising 9.4 kb BamHI genomic fragment. Subsequent hybridisation of restriction digests of pPU1 using the 0.6 kb probe identified a 1.25 kb HindIII fragment which was subcloned and sequenced (pHH). To ensure the integrity of the cloned region surrounding exon 1, the 9.4 kb genomic clone and normal human genomic DNA were analysed by restriction digests and hybridisation. The sizes of hybridising fragments within the 9.4 kb BamHI clone were found to be identical to corresponding genomic fragments (data not shown). The sequence of the 1.25 kb HindIII clone (pHH), encompassing the first exon and 0.9 kb of upstream sequence, in addition to ~400 bp of downstream 3′ sequence, is shown in Figure 1.

Comparison of regulatory regions of evolutionarily distant organisms may reveal potential regulatory elements that are conserved, implying functional significance. In an attempt to delineate DNA elements that may regulate the activity of the utrophin promoter, the equivalent mouse promoter was isolated. A mouse genomic phage library was screened with the human 0.6 kb probe. Positively hybridising phage were further subcloned into plasmid vectors and a 3 kb EcoRI subclone (M2.2) was sequenced. Figure 1 shows a sequence comparison between the two corresponding sequences, displaying regions of strong similarity upstream of exon 1. Exon 1 has limited overall conservation between human and mouse, which may reflect the lower evolutionary constraints on untranslated regions. The consensus splice site for exon 1 is conserved between human and mouse genomic sequence.

The sequence of the first exon and 5′ flanking sequence has a very high GC content. Database searches of the human utrophin promoter sequence revealed 99% identity to an uncharacterised GC-rich sequence, identified as a CpG island according to the method described by Cross et al. (15). This sequence corresponds to position 988–1212 (Fig. 1) which is in exon 1 of the human
Figure 1. Sequence comparison of the human (top line) and mouse (bottom line) promoter regions and untranslated exon 1. The location of transcription initiation sites in human by primer extension assays and 5′ RACE analysis are indicated by closed and open triangles, respectively. The major start site identified for the mouse transcript by RNase protection is indicated by asterisks. The N box core sequence and conserved E-box are boxed in both sequences. The conserved splice junction for exon 1–intron 1 is marked. The primers U24 and U25 used in 5′ RACE analysis are denoted above the sequence. The numbering corresponds to the human utrophin promoter sequence EMBL accession no. X95523. Mouse utrophin promoter sequence: EMBL accession no. X95524.

utrophin gene and correlates to our previous findings that the 5′-UTR of utrophin is unmethylated (14).

Sequence analysis reveals the absence of TATA and CAAT motifs common to eukaryotic promoters, and the presence of several potential binding sites for the transcription factor Sp1 (Fig. 1). However, the Sp1 binding sites are not conserved between human and mouse and may indicate that these sites are not involved in the transcriptional regulation of utrophin or that there is no constraint for the position to be conserved between species.

A consensus ‘E’ box is conserved between both species (Fig. 1). An E box element is defined by the nucleotides CANNTG and is a helix–loop–helix factor-binding site involved in regulating muscle gene expression (16). The recently identified N-box element with the core sequence TTCCGG, which appears to regulate synaptic expression of the mouse acetylcholine receptor δ-subunit (17), is present in the 5′ flanking sequence of both the human and mouse utrophin gene, although in different positions (Fig. 1).

Utrophin transcription initiation sites

Different levels of utrophin expression were observed in human cell lines by RNase protection using a probe derived from the C-terminus. Utrophin was shown to be expressed at differing levels in the cervical epithelial HeLa cell line, adult kidney CL11T47 cell line, adult myoblast primary culture (M429) and the rhabdomyosarcoma IN157 cell line, from which utrophin was originally cloned. The highest levels were detected in IN157 and CL11T47 cells (Fig. 2).

A 22mer oligonucleotide (U25) with its 5′-end starting at nucleotide 1088 in exon 1 of the human utrophin gene was used for primer extension assays with RNA isolated from HeLa cells, kidney CL11T47 cells and rhabdomyosarcoma IN157 cells (Fig. 3). A product of 191 bp, representing an initiation site at nucleotide 897 (Fig. 1), was observed in HeLa and IN157 cells. This product extends the start of transcription 38 bp upstream from the 5′-end of the most 5′ utrophin cDNA clone isolated from library screens. A smaller product of 135 bp, which corresponds to transcription initiation at nucleotide 953 (Fig. 1), was generated in IN157 and CL11T47 cells. An extension product of 182 bp was also detected at faint levels in the IN157 sample, which correlates to a start site at nucleotide 906. The variation in primer extension products generated for different cell lines may reflect preferential usage of a start site within a particular cell type.

Rapid amplification of cDNA ends (5′ RACE) was employed to isolate the 5′-end of the human transcript. The oligonucleotide U25, described above, was used to synthesize cDNA from poly(A)+ RNA from the IN157 rhabdomyosarcoma cell line. Following anchor-ligation of the cDNA, an internal gene-specific
Utrophin transcript levels in different human cell lines determined by RNase protection analysis. The levels of utrophin were determined in 10µg total RNA from the following cell lines: cervical epithelial (HeLa), primary adult myoblast (M429); adult kidney (CL11T47); and rhabdomyosarcoma (IN157). Yeast RNA was used as a negative control. (A) Human C-terminus utrophin probe, which generates a 180 bp protected fragment, (B) phosphoglycerate kinase (PGK) probe which generates a protected fragment of 121 bp to control for RNA loading. (C) Histogram showing the normalised levels of utrophin in each cell line. The signal from each protected product was quantified by a Molecular Dynamics phosphorimager. The counts for the utrophin-specific probe were normalised to the values for the PGK probe to account for any difference in levels of RNA in each sample. The background count in the lane for the yeast RNA sample was subtracted from the sample values.

primer U24 (complementary to nucleotides 982–1002) was used in PCR amplification. The amplified products were subcloned and hybridised with the 0.6 kb cDNA fragment containing the first exon of utrophin. Sequence analysis of the positive clones indicated a putative start site at position 907 on the human sequence (Fig. 1). This corresponds to a start site predicted from the 182 bp product in the primer extension assays. A 20mer oligonucleotide (U71) with its 5'-end corresponding to 170 bp downstream from the start of exon 2 (within the untranslated region) was also used for RACE analysis with RNA from IN157 cells. The sequence of positive hybridising subcloned products predicted a putative initiation site at nucleotide 1214 and this may represent an additional more 3' transcription start site.

The most 5' site for transcription initiation of the human utrophin transcript predicts a putative cap site at nucleotide 897 bp, which is 588 bp upstream from the translation initiation codon, and estimates the size of exon 1 as 496 bp. The identification of additional putative starts of transcription may represent multiple clustered start sites for the utrophin transcript.

Probably because of the very high GC content, we experienced technical difficulties in applying RNase protection analysis across the 5'-UTR of human utrophin, a problem also described in the analysis of the GC-rich promoter of the human Dp71 dystrophin isoform (18). In order to try to circumvent this problem we used the mouse transcript, which is less GC rich, for RNase protection assays. A complementary RNA probe was synthesised encompassing the SmaI–PstI fragment from positions 743–1051 (Fig. 1), which spans the putative start site of the human transcript. An RNA probe complementary to sequence within the C-terminal domain (which distinguishes between utrophin and G-utrophin) (19) was used as a control for utrophin expression and we demonstrated that the utrophin transcript was abundantly expressed in mouse lung (Fig. 4A). However using the SmaI–PstI RNA probe at hybridisation temperatures of 42 and 50°C we were unable to detect a protected product (Fig. 4B). A major band of ~130 bp only started to appear at a hybridisation temperature of 55°C and increased in intensity with increasing hybridisation temperatures (Fig. 4B).

The SmaI–PstI RNA probe also hybridised to RNA from mouse liver at lower levels. No protected product was detected when hybridising the mouse SmaI–PstI probe to RNA isolated from human rhabdomyosarcoma IN157 cells (data not shown), indicating that the mouse probe is specific to the mouse transcript. This protected band of ~130 bp indicates the start of transcription initiation close to nucleotide 916 on the human sequence, which correlates to the most 5' start sites predicted by RACE and primer extension assays for the human transcript (Fig. 1). These findings suggest that a similar site for transcription initiation is utilised in human and mouse. It is observed that this region of sequence encompassing the most 5' putative start sites is conserved between the two species.

**Functional analysis of the promoter region**

To test whether a promoter is located in the 1.25 kb HindIII fragment, which contains 900 bp of 5' flanking sequences and the
The human and mouse utrophin promoter regions have several putative Sp1 binding sites and are devoid of TA TA or CAA T motifs. By primer extension, 5′ RACE and RNase protection analysis, we have located several putative start sites for the human utrophin transcript. Although we cannot rule out the possibility that these additional products arose due to premature

**DISCUSSION**

We have isolated and characterised a genomic fragment which contains the first exon and 5′ flanking sequence of human and mouse utrophin. The human fragment is active in initiating transcription of a reporter gene in various cell lines, indicating that the utrophin promoter element may be conserved. A series of 5′-deleted fragments of the human utrophin upstream flanking region were generated to determine the minimal promoter element. Provided the 155 bp region containing the promoter element was intact, deletions in the 5′-end did not significantly alter the transcriptional activity. Hence, the CpG-rich 155 bp region characterised here functions as a basal promoter element, driving utrophin transcription in many cell types.

The human and mouse utrophin promoter regions have several putative Sp1 binding sites and are devoid of TATA or CAAT motifs. By primer extension, 5′ RACE and RNase protection analysis, we have located several putative start sites for the full-length utrophin transcript. Although we cannot rule out the possibility that these additional products arose due to premature

**Figure 4.** Identification of the mouse transcription initiation site by RNase protection analysis. (A) C-terminal probe derived from mouse utrophin, which protects a fragment of 88 bp, was hybridised to 20 µg total RNA from mouse lung at increasing hybridisation temperatures as indicated below the lanes. (B) A 5′ probe, synthesised from the mouse sequence extending from 743–1051 (Fig. 1), was hybridised to 20 µg total RNA from mouse lung at increasing hybridisation temperatures. A protected band of ∼130 bp was apparent for hybridisation temperatures above 50°C.

**Figure 5.** Functional analysis of the 1.25 kb HindIII fragment containing the human utrophin promoter in various cell lines. The constructs pHH·F and pHH·R contain the promoter fragment in the forward and reverse orientation, respectively, cloned upstream of the promoterless luciferase gene in pGL2-basic plasmid. The luciferase activity from each construct was normalised to β-galactosidase activity from a cotransfected internal control plasmid pSV40-β-galactosidase. The results are expressed as a percentage of the maximal level of activity obtained by the pHH·F construct in COS-7 cells (100%). The histogram shows the average of three transfection experiments.
Figure 6. Deletion analysis of the 5′ flanking sequence of the human utrophin promoter. (A) Based on the restriction map of the 1.25 kb HindIII human genomic fragment, upstream restriction sites were utilised for the construction of utrophin–luciferase fusion constructs. The major start of transcription is marked with an arrow. (B) Schematic representation of luciferase activity obtained upon transfection of the constructs into human rhabdomyosarcoma IN157 cells. Transfection and luciferase assays were performed as described in Materials and Methods. The histogram on the right shows the average of the results from three to four independent transfection experiments. For each experiment, luciferase (LUC) activity was normalised to β-galactosidase activity and expressed as a percentage of the activity of the pHH·F construct (100%).

termination by reverse transcriptase during primer extension and RACE analysis, these results are consistent with the observation that CpG-rich TATA-less promoters of widely expressed genes usually contain several transcription initiation sites spread over a fairly large region, rather than at a single base position (21). Multiple clustered start sites have also been described for the acetylcholinesterase gene (22), the human N-CAM gene (23), the AChR α-subunit gene (24) and for the dystrophin brain-specific full-length transcript (25).

Genes with CpG-rich promoters were initially considered to express proteins with a ‘housekeeping’ function in the cell. However, several genes devoid of TATA or CAAT motifs in their promoter regions have been shown to encode proteins that are highly regulated (22,26). Expression of the acetylcholinesterase gene, which also has a CpG-rich promoter, is regulated during muscle cell differentiation and is localised specifically at the NMJ (22). The typical ‘housekeeping’ promoter of the Dp71 dystrophin isoform drives expression in specific cell types (18). Utophin, although expressed in all tissues, also appears to be regulated in different cell types. For example, there are relatively abundant levels in adult lung and higher levels in foetal muscle compared to adult skeletal muscle (27,28). Utophin transcripts are also specifically localised during development, with initial accumulation in the neural tube and later becoming abundant at a variety of other sites such as the tendon primordia in the digits, the pituitary, thyroid and adrenal glands, cardiac muscle, and the kidney and lung (29). Taken together, these observations suggest that, although utrophin is expressed widely, there is also developmental and tissue-specific regulation of expression in certain tissues. We have shown here that the utrophin transcript is detected at different levels in several human cell lines and thus specific transcription factors may regulate the different levels of utrophin expression in the various cell types.

Several putative DNA motifs identified in the 5′ flanking sequence may be involved in the control of utrophin muscle expression. We identified a conserved E-box, which is a binding site for helix–loop–helix proteins of the MyoD1 family, including MyoD1 (30), myogenin (31,32), myf5 (33) and MRF4 (34). E-box motifs are found in the promoters of many muscle-specific genes, and enhance the in vitro transcriptional activity of the α, β and γ AChR subunit genes (24,35–37). Given the co-localisation of utrophin with AChRs at the NMJ, it would be of interest to determine whether myogenic factors regulate the expression of utrophin by interaction with this conserved E-box motif. The human and mouse utrophin 5′ flanking region contain the core sequence of the N box, an element shown to direct synapse-specific expression of the mouse acetylcholine receptor δ-subunit gene (17). This TTCCGG motif restricts the expression of the δ-subunit gene to the NMJ by enhancing expression at the endplate and acting as a silencer in extrajunctional areas. Sequences identical to this core sequence of the N box, an element shown to direct synapse-specific expression of the mouse acetylcholine receptor δ-subunit gene (17). This TTCCGG motif restricts the expression of the δ-subunit gene to the NMJ by enhancing expression at the endplate and acting as a silencer in extrajunctional areas. Sequences identical to this core sequence of the N box, an element shown to direct synapse-specific expression of the mouse acetylcholine receptor δ-subunit gene (17). This TTCCGG motif restricts the expression of the δ-subunit gene to the NMJ by enhancing expression at the endplate and acting as a silencer in extrajunctional areas. Sequences identical to this core sequence of the N box, an element shown to direct synapse-specific expression of the mouse acetylcholine receptor δ-subunit gene (17).
with this observation. The mRNA levels of N-CAM, 43K-rapsyn and s-laminin (38) were shown by in situ hybridisation to be concentrated at the synaptic sites. By database searching, we have determined that the core sequence of the N box is present in the 5′ flanking sequence of β2-syntrophin, which is also localised specifically to the NMJ, whereas it is absent from the upstream sequence of α1-syntrophin, which is localised at the sarcolemma (12). We also determined that the element is absent from the sequences upstream of the brain and muscle-specific dystrophin promoters. This suggests that there may be a general mechanism for selective transcription by synaptic nuclei and this may involve the interaction of a transcription factor(s) capable of recognising the N box sequence.

We have used PCR amplification of single-stranded cDNA synthesised from a range of mouse tissues (brain, liver, lung, heart, skeletal muscle, kidney, small intestine, spleen and eye) to demonstrate that the untranslated first exon of the utrophin gene is utilised in all mouse tissues analysed (data not shown), giving preliminary evidence that the promoter element is active in these tissues. This does not exclude the possibility that there are other full-length isoforms of utrophin derived from alternate promoters, with distinct, and possibly overlapping, patterns of expression. The dystrophin gene has at least three promoters at the 5′-end of the gene which direct the tissue-specific expression of the full-length isoforms (39–42). However, in contrast to dystrophin, the utrophin gene has a CpG island. The differences in genomic organisation at the 5′-end of these genes may reflect a different mechanism for regulating expression of the full-length transcripts. Thus there may only be one promoter at the 5′-end of the utrophin gene, as characterised here, and the control of the widely expressed full-length transcript may be regulated via tissue specific transcription factors binding to as yet uncharacterised elements.

In summary, the CpG-rich utrophin promoter contains several Sp1 binding sites and has motifs that may direct muscle and synapse-specific expression. The promoter directs transcription from multiple start sites. No differing first exon has as yet been determined that the core sequence of the N box is present in the 5′-end of the gene which direct the tissue-specific expression of the full-length isoforms (39–42). However, in contrast to dystrophin, the utrophin promoter element is active in these tissues. The dystrophin gene has at least three promoters at the 5′-end, which direct the tissue-specific expression of the full-length isoforms (39–42). However, in contrast to dystrophin, the utrophin gene has a CpG island. The differences in genomic organisation at the 5′-end of these genes may reflect a different mechanism for regulating expression of the full-length transcripts. Thus there may only be one promoter at the 5′-end of the utrophin gene, as characterised here, and the control of the widely expressed full-length transcript may be regulated via tissue specific transcription factors binding to as yet uncharacterised elements.

ACKNOWLEDGEMENTS

This work is supported by research grants from the Medical Research Council, the Muscular Dystrophy Group of Great Britain and Northern Ireland and the Muscular Dystrophy Association USA. C.D. gratefully acknowledges the support of the Sir Robert Menzies Memorial Trust and the Overseas Research Students Awards Scheme. We thank our colleagues at Cold Spring Harbor Laboratories and Oncogene Science Inc. for helpful discussions and also Derek Blake for critical reading of the manuscript.

REFERENCES