

Identification of a Novel First Exon in the Human Dystrophin Gene and of a New Promoter Located More Than 500 kb Upstream of the Nearest Known Promoter

Hisahide Nishio, Yasuhiro Takeshima,* Naoko Narita,* Hiroshi Yanagawa, Yoshiyuki Suzuki,[‡] Yukistoshi Ishikawa,[§] Yuka Ishikawa,[§] Ryoji Minami,[§] Hajime Nakamura, and Masafumi Matsuo*

*Department of Pediatrics and Division of Genetics, International Center for Medical Research, Kobe University School of Medicine, Kobe 650, Japan; [‡]Department of Clinical Genetics, The Tokyo Metropolitan Institute of Medical Science, Tokyo 113, Japan; and [§]Department of Pediatrics, National Sanatorium Yakumo Hospital, Hokkaido 04931, Japan

Abstract

The dystrophin gene, which is mutated in patients with Duchenne and Becker muscular dystrophies, is the largest known human gene. Five alternative promoters have been characterized until now. Here we show that a novel dystrophin isoform with a different first exon can be produced through transcription initiation at a previously unidentified alternative promoter. The case study presented is that of a patient with Duchenne muscular dystrophy who had a deletion extending from the 5' end of the dystrophin gene to exon 2, including all promoters previously mapped in the 5' part of the gene. Transcripts from lymphoblastoid cells were found to contain sequences corresponding to exon 3, indicating the presence of new promoter upstream of this exon. The nucleotide sequence of amplified cDNA corresponding to the 5' end of the new transcript indicated that the 5' end of exon 3 was extended by 9 codons, only the last (most 3') of which codes for methionine. The genomic nucleotide sequence upstream from the new exon, as determined using inverse polymerase chain reaction, revealed the presence of sequences similar to a TATA box, an octamer motif and an MEF-2 element. The identified promoter/exon did not map to intron 2, as might have been expected, but to a position more than 500 kb upstream of the most 5' of the previously identified promoters, thereby adding 500 kb to the dystrophin gene. The sequence of part of the new promoter region is very similar to that of certain medium reiteration frequency repetitive sequences. These findings may help us understand the molecular evolution of the dystrophin gene. (*J. Clin. Invest.* 1994. 94:1037–1042.) **Key words:** dystrophin • Duchenne muscular dystrophy • promoter • mRNA

Address correspondence to Masafumi Matsuo, M.D., Ph.D., Division of Genetics, International Center for Medical Research, Kobe University School of Medicine, 7-5-1 Kusunokicho, Chuo, Kobe 650, Japan.

Received for publication 24 February 1994 and in revised form 18 May 1994.

1. *Abbreviations used in this paper:* BMD, Becker muscular dystrophy; DMD, Duchenne muscular dystrophy; 5'RACE, rapid amplification of 5' cDNA end.

J. Clin. Invest.

© The American Society for Clinical Investigation, Inc.

0021-9738/94/091037/06 \$2.00

Volume 94, September 1994, 1037–1042

Introduction

Together, Duchenne and Becker muscular dystrophies (DMD and BMD, respectively)¹ are the most common inherited muscular disease. They are caused by mutations in the dystrophin gene, which extends over 2,500 kb of the X chromosome and is thus the largest known human gene. Dystrophin, a cytoskeletal protein localized to the sarcolemma, is encoded by a 14-kb RNA transcript that is processed from 79 exons (1). The amino-terminal actin binding domain of dystrophin is thought to link its intracellular domain to the membrane cytoskeleton, while its carboxyl terminus may interact with transmembranous dystrophin-related glycoproteins (2). Since the discovery of the gene, much attention has been focused on its large size, which is considered to be one of the factors responsible for unusually high incidence of DMD/BMD in all human populations. Up to 70% of DMD/BMD patients have a deletion in the dystrophin gene, and nearly 40% of all deletion breakpoints map to the large intron, number 44 (3).

In addition, it has been variously suggested that another gene might be located within an intron of the dystrophin gene (3), and that many isoforms of dystrophin could be encoded by this mammoth gene (4). Indeed, differential splicing in the 3' coding region has been shown to generate isoforms that may possess diverse functional properties (5–7). Until now, five promoters have been mapped in the dystrophin gene (2). Three of these promoters allow expression of a full-length dystrophin transcript (8–10), while the two other promoters located near the 3' end of the gene allow the expression of a truncated transcript coding for a carboxyl-terminal domain of the protein (4, 11). All of these promoters are considered to be activated in a tissue-specific manner (2) or at a particular stage in development (12). Dystrophin isoforms are named after their specific tissue distribution: C-, M-, P-, S-, and G-dystrophins are mainly produced in the cortex, muscle cells, Purkinje cells, Schwann cells and general tissues, respectively. Thus, the form of dystrophin which is defective in DMD/BMD patients is now classified as M-dystrophin. The three promoters located in the 5' region of the gene each code for a processed transcript with a unique first exon spliced to a common exon 2. Physical mapping has shown that the C-, M-, P-promoters are arranged in tandem in the 5' region of the gene (see Fig. 1).

One third of DMD/BMD patients show a variable degree of mental retardation (13), suggesting the existence of a previously unsuspected form dystrophin essential for normal brain development. Therefore, the possible existence of additional, tissue-specific dystrophin gene promoters should not be discounted. The identification of such promoters would represent an im-

portant step towards a better understanding of the unusual structure of this gene. In this paper we describe the results of cloning and physical mapping of the sixth dystrophin promoter region identified from the transcript in lymphoblastoid cell lines.

Methods

Cell line established from a case of DMD. An Epstein Barr virus (EBV)-transformed lymphoblastoid cell line was established from a Japanese DMD patient. Southern blot analysis revealed the presence of a large deletion extending from the 754 locus to exon 2 of the dystrophin gene (see Fig. 1, case A) (14). We previously demonstrated the absence of both C- and M-promoters by analyzing the dystrophin transcript in this cell line (15).

Genomic DNA analysis. Genomic DNA extracted from cultured EBV-transformed lymphoblastoid cells or peripheral lymphocytes (1×10^5 cells) as described previously (16) was used as a template for polymerase chain reaction (PCR) amplification. The primer sequences for amplification of the C-promoter and exon 3 were reported previously (9, 17). The amplification of the new promoter region was carried out using the following primers: 5'-GTATAGCCTAAGTGTA CAGTG-3' (see Fig. 2 b, F1) and 5'-CTGTCAGAGATTCTAACACAATGG-3' (see Fig. 2 b, R1). The conditions for PCR and detection of amplified DNA were as described previously (16).

mRNA analysis. Total RNA was extracted from cultured EBV-transformed lymphoblastoid cells or peripheral lymphocytes (1×10^6 cells) by the method of Chomczynski and Sacchi (18) and cDNA was prepared as described previously (16). Two rounds of PCR amplification were necessary in order to visualize the product by ethidium bromide staining after electrophoresis in an agarose gel as described before (15). To identify a new transcript, the initial amplification of cDNA encompassing exons L1 to 8 was followed by a second round of amplification of cDNA encompassing exons L1 to 4. The sequences of the primer sets were as follows: outer primer set: 5'-ACTGACACATA GAGTAACTT-3' (Fig. 2, F2) and 5'-CTGTTGAGAATAGTGCATTTGATG-3' (MR8: complementary to exon 8 sequence), inner primer set: 5'-CTAGTCTCCAAGCTCTACTC-3' (Fig. 2, F3) and 5'-GTTCAGGGC-ATGAACCTCTTG-3' (MR4: complementary to exon 4 sequence).

Identification of the 5' end of the transcript. To determine the 5' end of the transcript, rapid amplification of cDNA ends (5' RACE) was carried out using the GIBCO BRL 5' RACE System (GIBCO BRL, Gaithersburg, MD). The first PCR amplification was performed with an anchor primer and a primer of MR8. Then, 2 μ l of the first PCR product was subjected to 35 rounds of a second PCR amplification with a universal amplification primer and an exon 3-specific reverse primer (MR3: 5'-CTGTAGGTCAGTGAAGAGGTT-3').

Identification of the promoter region. Inverse PCR was used to clone the region flanking the new exon. 2 μ g of genomic DNA from the patient (case A) was digested with *Sau3A* I (Takara Shuzo Co., Kyoto, Japan) and the digested products were circularized using a DNA ligation kit under the conditions recommended by the supplier (Takara Shuzo Co.). The ligated DNA was used for a template for amplification by PCR using exon L1-specific primers. The primers used were R2 (5'-GAAGTTACTCTATGTGTCAAGTGA-3') and F3 in Fig. 2. The conditions for PCR were as previously described (16).

Sequencing of the 5' end of cDNA and the promoter region. The product amplified either by 5' RACE or by inverse PCR was subcloned into pT7Blue (R) (Novagen, Madison, WI). The sequences of the inserted DNA from four or more clones were determined by automated DNA sequencing using the Taq dye primer cycle sequencing kit (Applied Biosystems Inc., Foster City, CA) with a model 373A machine (Applied Biosystems Inc.).

Results

The dystrophin transcript containing the exon 3 sequence was analyzed in cDNA prepared from an EBV-transformed lymph-

oblastoid cell line derived from a DMD patient with a deletion which removed the 5' promoter region of the dystrophin gene (Fig. 1, case A). We did not expect this cell line to have a dystrophin transcript containing exon 3, but we were in fact able to use nested PCR to amplify a fragment corresponding to exons 3 to 8 (data not shown). The size of amplified product was the same as that amplified from control (wild type) cells, and we did not detect any other fragments which might have indicated alternative splicing in the transcript in the DMD lymphoblastoid cell line. These results suggested the presence of a novel transcription initiation site upstream of exon 3, in addition to three proximal promoters known to be present in the wild-type dystrophin gene.

5' RACE was used to characterize the 5' end of the new dystrophin transcript detected in the cell line. A single amplified DNA product was obtained. Sequencing of this fragment led to the identification of 46 unknown nucleotides upstream of exon 3 (Fig. 2 a). The only ATG codon in the newly-identified sequence which was in-frame with exon 3 was at the junction with the exon. Thus, the new dystrophin transcript (which we will call the L-dystrophin transcript) presumably codes for a novel form of dystrophin with a methionine residue followed by the sequence normally encoded by exon 3. The presence of an in-frame stop codon 30 bp from the 3' end of the new sequence suggests that translation does not initiate elsewhere, and therefore that it corresponds to the 5' untranslated region of the new dystrophin transcript. Transcription presumably starts 43 bp upstream of translational start point, although we can not rule out the possibility that this site arises due to premature termination by reverse transcriptase.

Although L-dystrophin is considered to be another dystrophin isoform, we could not ignore the possibility that it is the product of a novel gene, unique to the patient, created by fusing the 5'-truncated dystrophin gene to the 3' end of a 3'-truncated gene located upstream. To test this possibility, a fragment encompassing exon L1 to exon 4 was amplified from cDNAs prepared from peripheral lymphocytes of seven control subjects and two control lymphoblastoid cell lines by nested PCR. A product corresponding to 156 bp was obtained from cDNA of case A, and the same size product with some uncharacterized bands was also obtained when control cDNAs were amplified (Fig. 3 a). To determine whether these 156-bp bands corresponded to the L-dystrophin transcript, the single DNA band from case A and the co-migrating band from the control (156 bp) were cut out from the agarose gel and subjected to further amplification using an inner reverse primer as described in Fig. 3. A single clear band corresponding to the region comprising exons L1 and 3 (83 bp) was obtained in both cases (Fig. 3 b). To further rule out the possibility that this fragment was a PCR byproduct, three additional experiments were performed. The results showed the following (data not shown): (a) a fragment corresponding to 156 bp was not amplified when control cDNA was omitted from the reaction mixture or when cDNA synthesis step was skipped; (b) no amplified product was obtained when cDNA from a patient with deletion of exons 3 and 4 was used to amplify the 156-bp fragment; (c) an expected size product encompassing L1 to 8 could be amplified when different sets of primers were used at nested PCR. These indicates that the L-dystrophin transcript exists in control peripheral lymphocytes and lymphoblastoid cell lines, and therefore that it is not the product of a gene fusion.

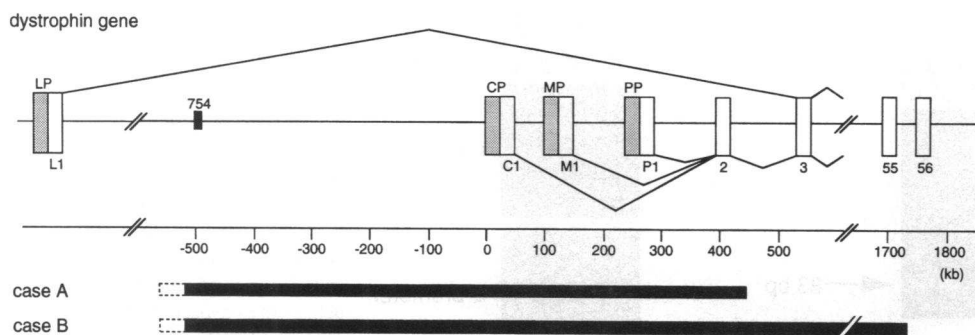


Figure 1. Physical map of the 5' region of the dystrophin gene. Open, dotted, and black boxes represent exons, promoters and the 754 locus, respectively. Thin horizontal lines joining these boxes indicate introns. Diagonal lines indicate splicing patterns. CP, MP, and PP represent the C-, M-, and P-promoters, respectively. LP, the promoter identified in this paper, is located upstream of 754 locus. Numbers under each open box indicate the exon number.

ber. The two deleted segments of the dystrophin gene are shown below the physical map. The upper shaded bar represents the segment deleted from the dystrophin gene in DMD case A. It extends from an imprecisely-mapped point upstream from the 754 locus through to exon 2 (14). The lower shaded bar represents the deleted region in DMD case B. It spans the region from an undetermined site upstream of the 754 locus to a site between exons 55 and 56.

A product encompassing exons L1, 2, 3, and 4 (218 bp) was not amplified from the control cDNA, but a product encompassing exons L1, 3 and 4 (156 bp) was amplified (Fig. 3 a). Though larger bands amplified from controls seemed to correspond to alternative splice product containing exon 2 (Fig. 3 a), no amplified product was obtained when primer complementary to exon 2 was used at nested PCR step. This indicates that the L-dystrophin transcript is naturally devoid of exon 2, and thus that the use of the alternative (L) promoter is accompanied by alternate splicing which deletes this exon. It is quite conceivable that the biological role of L-dystrophin might be different from that of other dystrophins because it lacks the 20

amino acids of the actin binding domain which are encoded normally by exon 2.

To further characterize L-dystrophin, the region flanking exon L1 was cloned using inverse PCR and then sequenced (Fig. 2 b). Since the transcription initiation site was not confirmed, the position of the 5' end of the transcript (position at +1 in the sequence of Fig. 2 b) must be considered as tentative. A total of 400 bp containing a single exon was identified. A systematic analysis for homologies using GenBank data base disclosed that a part of the identified sequence (from position -31 to position +73 in Fig. 2 b) shared 80% identity with a medium reiteration frequency repetitive sequence, MER17A1 (19) (Fig. 2 c). Fur-

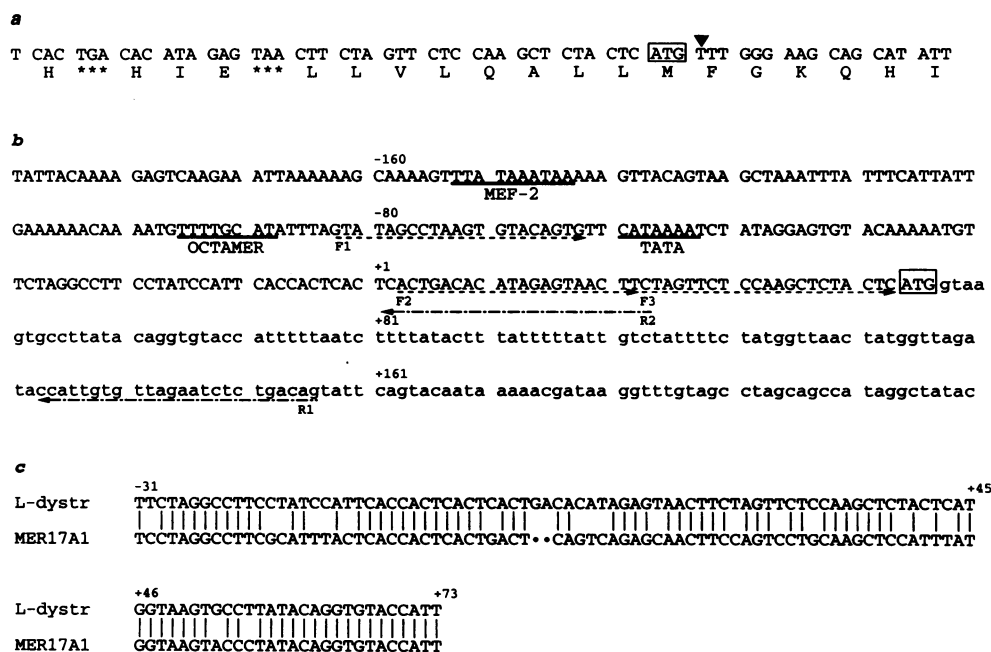


Figure 2. Nucleotide sequence of the region of the dystrophin gene under study. (a) DNA and predicted amino acid sequence of the exon L1-exon 3 region of L-dystrophin obtained by 5' RACE. The arrowhead indicates the nucleotide corresponding to the 5' border of exon 3. The only initiation codon in-frame with the downstream dystrophin sequences is highlighted. (b) Genomic DNA sequence of exon L1 and its flanking regions. The 5' flanking and exon sequences are in upper case letters, and the intron sequence is in lower case letters. Bases are numbered relative to the 5' end of exon L1 (+1). Positions of consensus sequence homologous for an MEF-2 element, an octamer motif and a TATA box are underlined. Dotted lines indicate sites of primers used to amplify the alternative first exon from genomic DNA or to amplify the L-dystrophin transcript. F and R indicate the forward and reverse primers, respectively. The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence data base under accession number D32048. (c) Comparison of L-dystrophin promoter/exon and MER17A1 nucleotide sequences. All nucleotides sequence of MER17A1 (19), a medium reiteration frequency repetitive sequence, is described and is compared with that of L-dystrophin promoter/exon. The vertical bars point out identical nucleotides.

script. F and R indicate the forward and reverse primers, respectively. The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence data base under accession number D32048. (c) Comparison of L-dystrophin promoter/exon and MER17A1 nucleotide sequences. All nucleotides sequence of MER17A1 (19), a medium reiteration frequency repetitive sequence, is described and is compared with that of L-dystrophin promoter/exon. The vertical bars point out identical nucleotides.

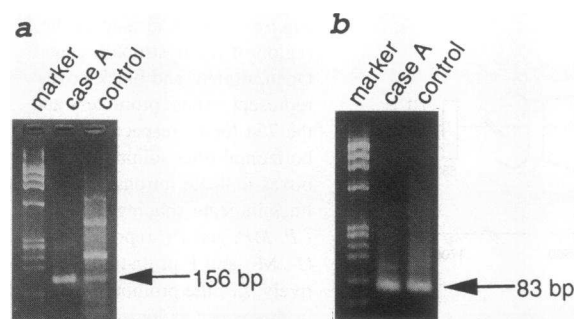


Figure 3. Amplification of a cDNA fragment spanning exon L1 to exon 4. (a) Product amplified from cDNA by nested PCR. The transcripts encompassing exons L1 to 4 from case A and from a normal control were examined by RT-nested PCR. One clear band was obtained from case A. This band comigrated with a band (156 bp) amplified from the control, but several uncharacterized bands were amplified from control cDNA. (b) Further PCR amplification of amplified products. The 156-bp bands in *a* were cut out and reamplified by using an inner primer located in exon 3 (5'-TGTCAGGCCTTCGAGGAGGTCTA-3'). A single, clear band (83 bp) corresponding exon L1 and 3 was amplified in both cases. The marker lane was loaded with HaeIII digested ϕ X174 phage DNA.

ther, L-promoter region could be amplified from all of 20 control DNAs. This indicated the promoter region is present in every genomic DNA.

The 3' flanking sequence of the exon is consistent with it being a splice junction (20), indicating that this sequence is an independent exon. The sequence upstream of exon L1 includes a potential basal promoter element consisting of a TATA box and an octamer motif located 60 and 96 bp upstream of the 5' end of exon L1, respectively (Fig. 2 *b*). Significantly, an MEF-2 element, which is essential for the activity of the muscle-specific promoter, was identified at -153 bp. Therefore, this region was considered to be the promoter region for exon L1. This is the sixth promoter/exon identified in the dystrophin gene. The functional role of L-dystrophin will be analyzed further by characterizing the interactions between these *cis*-elements and *trans*-effectors.

Three facts suggested that promoter region of L-dystrophin might be present in intron 2. First, the case in whom L-dystrophin was identified (Fig. 1, case A) had a 3' deletion endpoint in intron 2 (150 kb) (9). Second, the second amino acid of L-dystrophin is encoded by exon 3 (Fig. 2 *a*). Third, the L-dystrophin transcript found by analysis of control cDNA lacked sequence from exon 2 (Fig. 3). To investigate the relative position of L-dystrophin promoter to exon 3, L- and C-promoter regions and the exon 3 region were separately amplified from genomic DNA from a control case and from a DMD case who had been shown to have deletion stretching from locus 754 to exon 55 (Fig. 1, Case B). All three regions were clearly amplified from the control DNA, whereas only the L-promoter and exon 3 regions were amplified from the genomic DNA of case A (Fig. 4). Contrary to our expectations, case B must in fact have the L-dystrophin promoter, since we were able to amplify a product corresponding to the L-promoter region from the genomic DNA of this patient. This suggested that L-promoter is not located in intron 2 but is instead located upstream of the 754 locus, which is in the 770 kb Sfi I fragment (21) and ~ 500

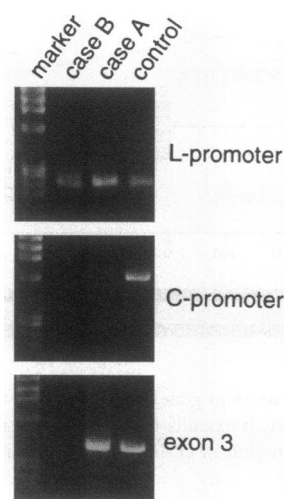


Figure 4. Amplification of regions of the L- and C-promoters and exon 3 from genomic DNA. Three regions including the L- and C-promoters and exon 3 were amplified from genomic DNA obtained from cases A and B and from a normal control. All three regions could be amplified from the control. The size of amplified products were as expected. Only the L-promoter region could be amplified from case B, while both the L-promoter and exon 3 regions were amplified from case A. The marker lane was loaded with HaeIII digested ϕ X174 phage DNA.

kb upstream from the C-promoter (Fig. 1). Therefore, ~ 500 kb have to be added to the previously estimated size of the dystrophin gene (2.5 Mb) (22). The alternative promoter/exon completely avoid exon 2 and are spliced to exon 3 in L-dystrophin. The size of the intron is calculated to be > 1,000 kb, which is larger than any other known human intron.

Discussion

The human dystrophin gene has hitherto been characterized by the following points: (a) at 2.5 megabases, it is the largest known human gene; (b) it has multiple promoters (at least five); (c) the exon length to intron length ratio is small (1:180); (d) it contains the largest known intron in a human gene, intron C1 (400 kb) (9). Our findings modify these characteristics somewhat: the gene is shown here to comprise more than three megabases, it has at least six alternative promoters, and it contains another intron (intron L1) which, at more than 1,000 kb, is even larger than intron C1. The large stretches of DNA between the alternative promoters (there are more than 500 kb between the L-dystrophin promoter and the C-dystrophin promoter; Fig. 1) may play an important role in delimiting distinct domains of control. The six promoters could conceivably be under the control of alternative enhancers or silencers. During maturation of mRNA precursor, three promoters (C, M, and P) and exon 2 are all spliced out to give the L-dystrophin transcript (Fig. 1). This raises the important question of whether enhanced expression from the L-dystrophin promoter is sufficient to facilitate the splicing pattern required to produce the L-dystrophin transcript, or whether additional factors are required. Further work will be required to clarify this point.

Prior to this study, it was known that the five independent dystrophin gene promoters specify the transcription of their respective alternative first exons in a cell-specific and developmentally controlled manner (2). Transcription from three of these promoters results in the production of full-length dystrophin. We are currently investigating the size of L-dystrophin transcript. However, low abundance of L-dystrophin transcript makes it quite difficult to determine its size. The transcript appears to be of low abundance even in lymphocytes because it could only be detected after nested PCR.

Exon 2 is exclusively skipped during splicing in the L-dystrophin transcript (Fig. 2a), even though the L promoter is upstream from the exon 2 sequence in the genomic DNA. This strongly suggests that the transcript lacking exon 2 is a real product from this promoter, because the exclusion of exon 2 is necessary for the transcript to maintain translational reading frame. This structural difference may provide a unique function to L-dystrophin. C-, M-, P-dystrophins can all bind to actin through their amino-terminal actin binding domain and all three forms can be stabilized in the sarcolemma. However, L-dystrophin lacks the amino terminal 31 amino acids encoded by exons M1 and 2 but has instead a single amino acid residue (Met) encoded by exon L1. Thus, we presume that L-dystrophin lacks actin binding activity. Two other dystrophin isoforms (G- and S-) lacking the amino terminal actin binding domain have also been identified; their functions are considered to be different from those of other full-length dystrophins (2). The function of L-dystrophin is supposed to be different from those of other dystrophin isoforms. Though its function would be estimated by analyzing its tissue distribution at protein or mRNA levels, unavailability of fresh human tissues hampered this study.

This is not the first example of a human gene which uses more than two promoters to produce transcripts with alternative 5' exons. Such phenomena could result in the synthesis of different or identical proteins (23, 24). When two isozymes are encoded by mRNAs transcribed from different promoters, they may both be tissue specific (25), they may both be widely distributed housekeeping proteins (26), or one can be housekeeping protein while the other is tissue-specific (24, 25, 27, 28). The first reported example of a gene with three optional promoters: one muscle-specific and two housekeeping, was the aldolase A gene (23). Recently, four promoters have been identified in the aromatase cytochrome P450 gene and these multiple promoters are subject to tissue-specific and development-dependent regulation (29). Therefore, L-dystrophin may also be expressed in either a tissue-specific or a development-dependent fashion.

Examination of the cis-element of each promoter region provides clues to the identity of the tissues in which they may be specifically expressed. The M-dystrophin promoter contains a TATA box, a G+C-rich motif, and CArG and MCAT boxes which are found in other muscle-specific genes (8). A housekeeping type promoter consisting of G+C-rich region containing three putative Sp-1 factor binding sites has been identified in the promoter region of G-dystrophin (4, 30), and an A+T-rich region has been identified in the promoter region of C-dystrophin (9). Byers et al. could not identify any specific sequences in the S promoter region that is specifically activated in peripheral nerve cells to give S-dystrophin (11).

The region of L-promoter studied here contains a TATA box, an octamer motif and an MEF-2 element (Fig. 2b). The sequence of the TATA box region (CATAAAA) is identical to that found in the β -globin gene (31). The presence of both octamer motif and TATA box has been reported to be sufficient for lymphoid-specific promoter activity (32). This suggests that L-dystrophin might be the product of a specific dystrophin transcript expressed in lymphoblastoid cells or in lymphocytes. The myocyte-specific enhancer-binding factor MEF-2 is undetectable in proliferating myoblasts but appears rapidly when myoblasts are induced to differentiate. If MEF-2 is a major trans-activator controlling the expression of the L-dystrophin

transcript, the L-dystrophin might be produced transiently during muscle fiber development. We are now underway to characterize the L-dystrophin promoter function.

Alternative promoters have apparently evolved to regulate tissue-specific protein synthesis. In general, transcription from multiple promoters affords flexibility in the control of gene expression by conferring either tissue- or developmental stage-specificity. One of the possible mechanisms by which a gene may acquire multiple promoters is the insertion of a new promoter carried on a transposable element (33). In a previous report we have showed an insertion mutation of a retrotransposon into the dystrophin gene in a case of DMD (34). The sequences of the L-promoter and of the first exon region are homologous to one of the medium reiteration frequency repetitive sequences, MER17A1, which itself is homologous to an intergenic sequence in the human crystalline gene (19). In addition, a fragment corresponding to exon L1 sequence is amplified by using PCR in rat genomic DNA (data not shown). This suggests that L-dystrophin may result from the insertion of a promoter/exon derived from a transposable element. This observation may contribute to a better understanding of the molecular evolution of the dystrophin gene.

Acknowledgments

We thank Dr. A. P. Pugsley for advice and for critically reading the manuscript.

This work was supported by grants from the Ministry of Education, Science and Culture of Japan and the National Center of Neurology and Psychiatry (NCNP) of the Ministry of Health and Welfare of Japan, the Hyogo Science and Technology Association, the Foundation for Therapeutics for Epilepsy and the Nakamoto Science Foundation.

References

1. Roberts, R. G., A. J. Coffey, M. Bobrow, and D. R. Bentley. 1993. Exon structure of the human dystrophin gene. *Genomics*. 16:536-538.
2. Ahn, A. H., and L. M. Kunkel. 1993. The structural and functional diversity of dystrophin. *Nature Genet.* 3:283-291.
3. Hoffman, E. P., and L. M. Kunkel. 1989. Dystrophin abnormalities in Duchenne/Becker muscular dystrophy. *Neuron*. 2:1019-1029.
4. Lederfein, D., D. Yaffe, and U. Nudel. 1993. A housekeeping type promoter, located in the 3' region of the Duchenne muscular dystrophy gene, controls the expression of Dp71, a major product of the gene. *Hum. Mol. Genet.* 2:1883-1888.
5. Feener, C. A., M. Koenig, and L. M. Kunkel. 1989. Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy terminus. *Nature (Lond.)*. 338:509-511.
6. Nudel, U., D. Zuk, P. Einat, E. Zeelon, Z. Levy, S. Neuman, and D. Yaffe. 1989. Duchenne muscular dystrophy gene product is not identical in muscle and brain. *Nature (Lond.)*. 337:76-78.
7. Chelly, J., J.-C. Kaplan, P. Maire, S. Gautron, and A. Kahn. 1988. Transcription of the dystrophin gene in human muscle and non-muscle tissues. *Nature (Lond.)*. 333:858-860.
8. Klamut, H. J., S. B. Gangopadhyay, R. G. Worton, and P. N. Ray. 1990. Molecular and functional analysis of the muscle specific promoter region of the Duchenne muscular dystrophy gene. *Mol. Cell. Biol.* 10:193-205.
9. Boyce, F. M., A. H. Beggs, C. Feener, and L. M. Kunkel. 1991. Dystrophin is transcribed in brain from a distant upstream promoter. *Proc. Natl. Acad. Sci. USA*. 88:1276-1280.
10. Gorecki, D. C., A. P. Monaco, J. M. J. Derry, A. P. Walker, E. A. Barnard, and P. J. Barnard. 1992. Expression of four alternative dystrophin transcripts in brain regions regulated by different promoters. *Hum. Mol. Genet.* 1:505-510.
11. Byers, T. J., H. G. W. Lidov, and L. M. Kunkel. 1993. An alternative dystrophin transcript specific to peripheral nerve. *Nature Genet.* 4:77-81.
12. Rapaport, D., O. Fuchs, U. Nudel, and D. Yaffe. 1992. Expression of the Duchenne muscular dystrophy gene products in embryonic stem cells and their differentiated derivatives. *J. Biol. Chem.* 267:21289-21292.

13. Emery, A. E. H. 1993. Duchenne muscular dystrophy. Oxford University Press. Oxford. 392.
14. Shimmoto, M., A. Oshima, R.-C. Yang, H. Sakuraba, Y. Nomura, M. Segawa, and Y. Suzuku. 1990. Atypical X-linked muscular dystrophy with mental retardation: Analysis of a large deletion at the proximal region of dystrophin gene. *Brain Dysfunct.* 3:72-79.
15. Nishio, H., Y. Kitoh, N. Narita, M. Shimmoto, Y. Suzuki, and H. Nakamura. 1993. Brain and muscle-type promoters of the dystrophin gene are selected in peripheral lymphocytes and Epstein Barr virus-transformed lymphoblastoid cells. *J. Neurol.* 241:81-86.
16. Matsuo, M., T. Masumura, H. Nishio, T. Nakajima, Y. Kitoh, T. Takumi, J. Koga, and H. Nakamura. 1991. Exon skipping during splicing of dystrophin mRNA precursor due to an intraexon deletion in the dystrophin gene of Duchenne muscular dystrophy Kobe J. *Clin. Invest.* 87:2127-2131.
17. Beggs, A. H., M. Koenig, F. M. Boyce, and L. M. Kunkel. 1990. Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. *Hum. Genet.* 86:45-48.
18. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
19. Kaplan, D. J., J. Jurka, J. F. Solus, and C. H. Duncan. 1991. Medium reiteration frequency repetitive sequences in the human genome. *Nucleic Acids Res.* 19:4731-4738.
20. Shapiro, M. B., and P. Senapathy. 1987. RNA splice junctions of different classes of eukaryotes: sequence statics and functional implications in gene expression. *Nucleic Acids Res.* 15:7155-7174.
21. van Ommen, G. J. B., J. M. H. Verkerk, M. H. Hofker, A. P. Monaco, L. M. Kunkel, P. Ray, R. Worton, B. Wieringa, E. Bakker, and P. L. Pearson. 1986. A physical map of 4 million bp around the Duchenne muscular dystrophy gene on the human X-chromosome. *Cell.* 47:499-504.
22. den Dunnen, J. T., L. Casula, A. Makover, B. Bakker, D. Yaffe, U. Nudel, and G. J. B. van Ommen. 1991. Mapping of dystrophin brain promoter: a deletion of this regions compatible with normal intellect. *Neuromuscular Disorders.* 5:327-331.
23. Maire, P., S. Gautron, V. Hakim, C. Gregori, F. Menecier, and A. Kahn. 1987. Characterization of three optional promoters in the 5' region of the human aldolase A gene. *J. Mol. Biol.* 197:425-438.
24. Chretien, S., A. Dubart, D. Beaupain, N. Raich, B. Grandchamp, J. Rosa, M. Goossens, and P.-H. Romeo. 1988. Alternative transcription and splicing of the human porphobilinogen gene either in tissue-specific or in housekeeping expression. *Proc. Natl. Acad. Sci. USA.* 85:6-10.
25. Perlino, M., R. Cortese, and G. Ciliberto. 1987. The human-antitrypsin gene is transcribed from two different promoters in macrophages and hepatocytes. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2767-2771.
26. Shtivelman, E., B. Lifshitz, R. P. Gale, B. A. Roe, and E. Canaani. 1986. Alternative splicing of RNAs transcribed from the human abl gene and from the bcl-abl fused gene. *Cell.* 47:277-284.
27. Takadera, T., S. Leung, A. Gernone, Y. Koga, Y. Takihara, N. G. Miyamoto, and T. W. Mak. 1989. Structure of the two promoters of the human lck gene: differential accumulation of two classes of lck transcripts in T cells. *Mol. Cell. Biol.* 9:2173-2180.
28. Shapiro, L. H., R. A. Ashmun, W. M. Roberts, and A. T. Look. 1991. Separate promoters control transcription of the human aminopeptidase N gene in myeloid and intestinal epithelial cells. *J. Biol. Chem.* 266:11999-12007.
29. Mahendroo, M. S., C. R. Mendelson, and E. R. Simpson. 1993. Tissue-specific and hormonally controlled alternative promoters regulate aromatase cytochrome P450 gene expression in human adipose tissue. *J. Biol. Chem.* 268:19463-19470.
30. Hugnot, J. P., H. Gilgenkrantz, N. Vincent, P. Chafey, G. E. Morris, A. P. Monaco, Y. Berwald-Netter, A. Koulakoff, J. C. Kaplan, A. Kahn, and J. Chelly. 1992. Distal transcript of the dystrophin gene initiated from an alternative first exon and encoding a 75-kDa protein widely distributed in nonmuscle tissues. *Proc. Natl. Acad. Sci. USA.* 89:7506-7510.
31. Kazazian, H. H. 1990. The thalassemia syndromes: molecular basis and prenatal diagnosis in 1990. *Seminars Hemat.* 27:209-228.
32. Wirth, T., L. Staudt, and D. Baltimore. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature (Lond.).* 329:174-178.
33. Schibler, U., and F. Sierra. 1987. Alternative promoters in developmental gene expression. *Annu. Rev. Genet.* 21:237-257.
34. Narita, N., H. Nishio, Y. Kitoh, Y. Ishikawa, Y. Ishikawa, R. Minami, H. Nakamura, and M. Matsuo. 1993. Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. *J. Clin. Invest.* 91:1862-1867.