

The dystrophin gene is alternatively spliced throughout its coding sequence

M. Sironi^{a,*}, R. Cagliani^a, U. Pozzoli^a, A. Bardoni^a, G.P. Comi^b, R. Giorda^a, N. Bresolin^{a,b}

^aIRCCS E. Medea, Associazione La Nostra Famiglia, Via Don Luigi Monza 20, 23842 Bosisio Parini (LC), Italy

^bCentro Dino Ferrari, Istituto di Clinica Neurologica, Università di Milano, IRCCS Ospedale Maggiore Policlinico, Milan, Italy

Received 6 March 2002; accepted 13 March 2002

First published online 28 March 2002

Edited by Lev Kisselev

Abstract We have analysed splicing patterns in the human dystrophin gene region encoding the rod and cysteine-rich domains in normal skeletal muscle, brain and heart tissues. Sixteen novel alternative transcripts were identified, the majority of them being present in all three tissues. Tissue-specific variants were also identified, suggesting a functional role of transcriptional diversity. Transcript analysis in dystrophinopathic autaptic and bioptic specimens revealed that pre-mRNAs secondary structure formation and relative strength of exon/exon association play little or no role in directing alternative splicing events. This analysis also showed that independent deletion events leading to the loss of the same exons may be associated with transcriptional variability. © 2002 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Alternative splicing; Dystrophin; Duchenne muscular dystrophy; Becker muscular dystrophy; Exon skipping; Intron removal; Exon codon phase

1. Introduction

Dystrophin is the largest known human gene: it extends over 3000 kb on the X chromosome [1] and is transcribed into a 14-kb mRNA [2]. The gene is composed of 79 exons that together account for only 0.6% of the sequence [3]. Its main protein product, dystrophin, a member of the spectrin superfamily, is a rod-shaped 427-kDa protein [3]. In skeletal muscle dystrophin localises to the cytoplasmic surface of the sarcolemma where it is thought to provide a link between cytoskeletal actin and the extracellular matrix.

Three full-length dystrophin isoforms have been described, each controlled by a tissue-specific promoter [4–6]. The muscle isoform is mainly expressed in skeletal muscle but also in smooth and cardiac muscles [7]; brain dystrophin is specific for cortical neurons but can also be detected in heart and cerebellar neurons, while the Purkinje-cell type accounts for nearly all cerebellar dystrophin [6]. Alternative splicing provides a means for dystrophin diversification: the 3' region of the gene undergoes alternative splicing resulting in tissue-specific transcripts in brain neurons, cardiac Purkinje fibres, and smooth muscle cells [7,8] while 12 patterns of alternative splicing have been described in the 5' region of the gene in skeletal muscle [9].

Mutations in the dystrophin gene are responsible for either Duchenne or Becker muscular dystrophy (DMD or BMD). The majority of DMD and BMD patients carry deletions in the gene (60–65% of cases) [7] and a good correlation exists between the severity of the muscle disease and the effect of the deletion on the reading frame [10]. However exceptions to the reading frame rule are found in about 8% of patients [11] and the possibility that alternative splicing events could modify the clinical phenotype of DMD and BMD by editing the translational reading frame has been proposed [12,13]. A major hot spot for DMD/BMD deletions has been identified around exons 45–55; the existence of alternative splicing events in this region of the gene has never been investigated. Here we provide description of dystrophin transcripts arising from alternative splicing events in the region encoding the rod and cysteine-rich domains. In DMD and BMD, muscle involvement can be associated with heart disease and/or mental retardation; in this view, transcript analysis was extended to human brain and heart tissues.

2. Materials and methods

Tissue samples were derived, with informed consent of the family, from either autopsies or muscle biopsies of healthy individuals and DMD/BMD patients. Patient selection followed standard international criteria and dystrophin gene deletion analysis was performed using routine procedures. Total RNA was extracted from frozen specimens using RNeasy (Qiagen) in accordance with the manufacturer's specifications. RT-PCR for cDNA synthesis was performed using random hexamers and 1st strand cDNA synthesis kit (Boehringer). Exons 37–38 of the dystrophin gene were amplified from each sample in order to verify the ability of the cDNA to act as a template for dystrophin amplification. Six nested primer pairs were used to amplify dystrophin exons 17–58, as previously described [14]. Sixteen additional primers were designed to specifically amplify splice variants (sequences are available on request). PCR fragments were separated on 2% agarose gels; after purification, the amplified products were sequenced using the same primers as used for PCR amplification and BigDye[®] terminator cycle sequencing (PE Applied Biosystems) on an ABI PRISM 310 genetic analyzer. Exonic splicing enhancer (ESE) scores were calculated for dystrophin exons using ESE scoring matrices as previously described [15,16]. Matrices were slightly modified as suggested by the authors (personal communication).

3. Results and discussion

Analysis of dystrophin alternative splicing events in the region encompassing exons 17–58 was performed for human skeletal muscle, brain and heart tissues. Spliced products were amplified with nested primer pairs from normal human tissues and directly sequenced in order to determine their exact exon

*Corresponding author. Fax: (39)-031-877499.
E-mail address: msironi@bp.inf.it (M. Sironi).

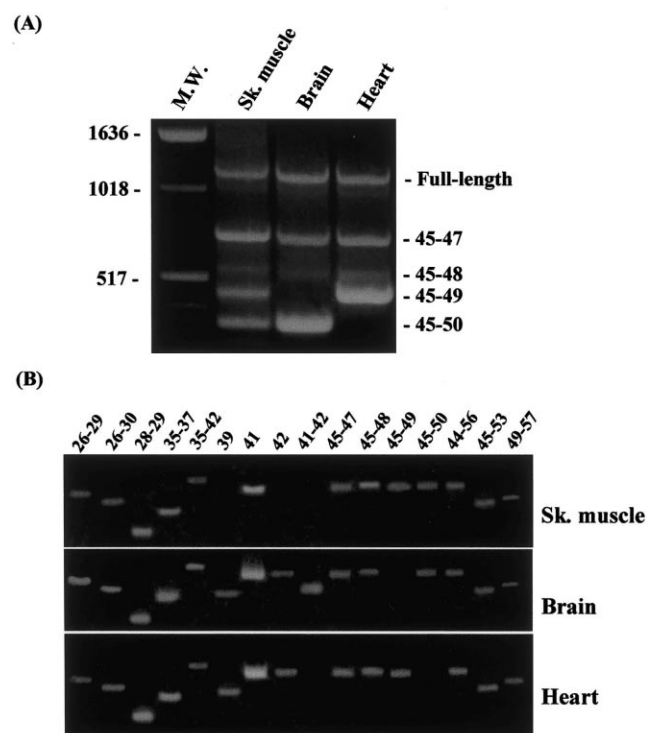


Fig. 1. RT-PCR amplification of dystrophin alternatively spliced transcripts in human skeletal muscle, brain and heart. A: Amplification of splice variants using nested PCR set 7 (exons 43–51). B: Selective amplifications using transcript-specific primers. Lane headings indicate skipped exons.

composition. As nested PCR amplifications often resulted in a complex transcript pattern (Fig. 1A), specific primers were designed on the basis of the novel exonic junctions and new selective amplifications were carried out to confirm the presence of the alternatively spliced variants and to better assess their tissue specificity. Selective amplifications for skeletal muscle, brain and heart tissues are shown in Fig. 1B. The presence of each alternative transcript was confirmed in tissues from at least two unrelated subjects. A total of 16 alternative transcripts were identified, the majority of them being present in all three tissues (Table 1). Tissue-specific splice variants were also identified: skipping of exons 39 and 42 was specific for brain and heart, while transcripts lacking exons 45–49 were only detectable in skeletal and cardiac muscles. Loss of exons 41–42 was restricted to brain while omission from dystrophin mRNA of exons 45–50 was specific for brain and muscle and was never detected in heart. Spliced variants displayed a great variability with respect to the number of missing exons with some of them skipping up to 13 exons.

Interestingly, when the reading frame of all alternatively spliced gene products was determined, only one of them, namely the transcript lacking exons 45–50, was predicted to be translated into a truncated dystrophin, while all the other products preserved the reading frame. An extensive study of alternative splicing events in the 5' region of the gene [9] revealed that only 50% of alternative products maintained the transcriptional reading frame of the full-length dystrophin transcript. Reading frame analysis of alternative transcripts might provide information concerning whether they are likely to serve any physiological role or they are merely the results

of splicing errors. If the latter were the case, one could expect to find in- or out-of-frame transcripts in a proportion reflecting exon codon phases in the gene, that is intron insertion patterns with respect to the codon position in reading frames. Since the only phase constraint is between two consecutive exons, reading frames of products derived from omission of one, two or more exons can be calculated and are expected to be a result of the number of different exon phases and of exon distribution along the gene. Frames of all possible products originating from skipping of one up to 10 consecutive exons were calculated for the regions of the dystrophin gene encompassing exons 1–18 and exons 19–58 (Table 2); a different expectation of in-frame products is observed in the two gene regions with the 5' portion presenting a percentage of in-frame products ranging from 44.4% to 75% and the central part a percentage ranging from 57.5% to 84.6%, depending on the number of skipped exons. These gene features might partially account for the different ratio of in-frame vs out-of-frame products detectable in different gene regions. Yet, our data indicate that, in the region under analysis, frame-shifted transcripts are definitely under-represented even when exon codon phases are considered. These findings might suggest that alternative splicing does not proceed at random and that alternative variants might serve specialised cellular functions, maybe irrespective of frame conservation or alteration. Indeed, the hypothesis cannot be rejected that out-of-frame transcript might have a biological role as well. Remarkably, here we show evidence that skipping of exons 45–50, which originates an out-of-frame product, is restricted to muscle and brain. Previously described dystrophin transcripts originating from alternative splicing events were always shown to be in-frame when displaying tissue specificity [7,8]; this is consistent with the notion that various molecular species might serve diverse roles in different tissues. The identification of an out-of-frame transcript displaying some degree of tissue specificity might suggest that production of frame-shifted transcripts is a regulated process and not the result of random processing. Interestingly, the widest range of different dystrophin transcripts could be detected in brain (Table 1); it is tempting to speculate that the high complexity, in terms of biological processes, of this tissue also carries along the need for multiple dystrophin gene products.

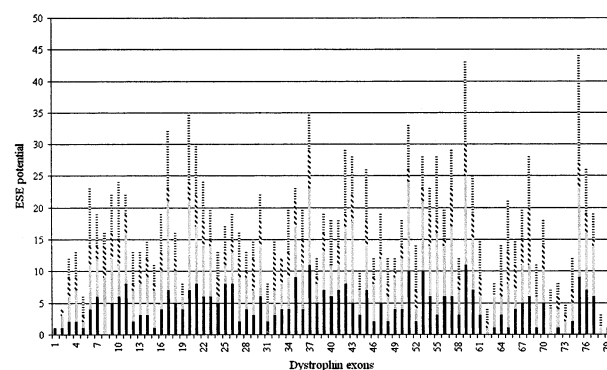


Fig. 2. ESE representation in single dystrophin exons. The entire bar length indicate total ESE potential; relative length of different segments indicate scores for four different SR protein recognition sequences; dark grey: SF/ASF; light grey: SRp40; hatched: SRp55; horizontal lines: SC35. ESE positions in dystrophin introns were not included, as they are not pertinent to this study, but are available on request.

Table 1
Analysis of alternative splicing events

PCR set ^a	skipped exons	Sk. muscle	Brain	Heart																						
4 (ex. 25-33)	26-29	+	+	+																						
	28-29	+	+	+																						
	26-30	+	+	+																						
5 (ex 33-40)	35-37	+	+	+																						
	39	-	+	+																						
5/6 (ex 33-44)	35-42	+	+	+	Pt 1 (del 44-52)	Pt 2 Pt 3 Pt 4 (del 45-47)			Pt 5 Pt 6 Pt 7 (del 45-48)			Pt 8 (del 46-48)		Pt 9 Pt 10 (del 48)		Pt 11 (del 48-50)		Pt 12 (del 48-51)		Pt 13 (del 45-53)		Pt 14 (del 50-52)		Pt 15 (del 45-52)		
	41	+	+	+																						
6 (ex 40-44)	42	-	+	+	Pt 1	Pt 2	Pt 3	Pt 4	Pt 5	Pt 6	Pt 7	Pt 8	Pt 9	Pt 10	Pt 11	Pt 12	Pt 13	Pt 14	Musc.	Brain	Heart					
	41-42	-	+	-																						
7 (ex 43-51)	45-47	+	+	+		+	+	+		+	+	+		+	+			+								
	45-48	+	+	+		-	-	-	+	+	+	-	+	+				+								
	45-49	+	-	+		+	+	+	+	+	+	-	+	+												
	45-50	+	+	-		+	+	+	+	+	+	+	+	+												
	49-57	+	+	+			+	+										+								
7/8 (ex 43-58)	44-56	+	+	+	+	-	+	+	-	-	-		+	-	+	-	-	+	+	+	+					
	45-53	+	+	+		-	-	-	+	+	+		-	-	+	+	+	+	-	+	+					

^aPCR sets followed a previous designation [14]; spanned exons are in parentheses.

Extensive studies of splicing mechanisms in human genes have allowed many of the principal elements that determine exon inclusion in the mature transcript to be elucidated. We considered 3' and 5' consensus values and U1 snRNA annealing free energies for the dystrophin gene region under analysis (data not shown) as previously calculated [17]; we found that these parameters provide no reasonable explanation concerning exon skipping in alternative transcripts. Recently, SR proteins, which bind ESE, have been shown to promote exon/exon associations [18] suggesting that they might be important for the selection of exon pairs during alternative splicing. We applied ESE scoring matrices [15,16] to single dystrophin exons in order to define their potential for SR protein binding. In particular, ESE analysis scored binding potentials for the different SR proteins identified to date; it is evident from Fig. 2 that ESE representation in different exons cannot account for the observed splicing patterns: in many cases exons that were found to be skipped in at least one alternative transcript displayed high SR protein binding potentials. Nonetheless,

different molecular mechanisms, which do not rely on specific sequence elements, might account for exon skipping in mature transcripts. mRNA secondary structure formation might have a role in this process, maybe bringing into close proximity splice sites separated by huge physical distances or sequestering skipped exons into a loop. Also, exon skipping might be expected to be influenced by the order of intron removal in a given gene region, in particular when skipping of multiple adjacent exons is considered. Intron removal is orderly but not processive from the 5' to the 3' end of precursor mRNA [19,20]; slow and fast splicing events exist that give rise to different major and minor pathways with a preferred order of intron removal. Further information concerning splicing mechanisms might be gained by the analysis of precursor mRNAs carrying variations in exon composition. We selected 14 patients carrying deletions in the dystrophin gene and verified the presence of alternative products in their skeletal muscles. Deletions involved exons either skipped in alternative transcripts or immediately flanking alternatively spliced

Table 2
Frame of all possible products deriving from skipping of one up to 10 consecutive exons

	Skipped exons									
	1	2	3	4	5	6	7	8	9	10
IF% ^a (exons 1–18)	52.9%	50.0%	46.7%	50.0%	69.2%	66.7%	45.5%	50.0%	44.4%	75.0%
IF% (exons 19–58)	57.5%	84.6%	60.5%	70.3%	66.7%	71.4%	70.6%	69.7%	75.0%	64.5%

^aIF%: percentage of in-frame products.

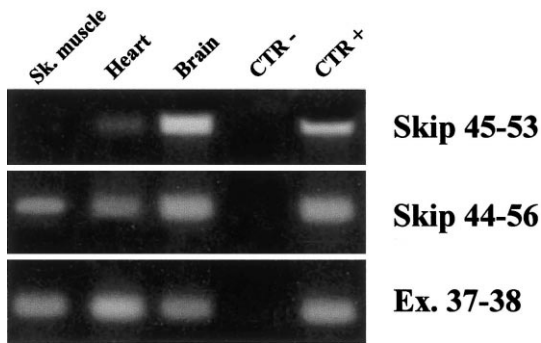


Fig. 3. Transcript analysis in tissues from patient 15. Splice variant-specific primers were used. Skipped exons are indicated on the right together with the exons amplified in the control reaction (exons 37–38).

regions. Data concerning splicing patterns in dystrophic muscles are reported in Table 1. When possible, two or three patients carrying the same deletion were analysed. Remarkably, splicing patterns differed according to the location or extension of deletion mutations in the dystrophin gene and, in some instances, also among patients carrying the same deletion. Surprisingly short deletions were found to abolish splicing variants that longer, overlapping deletions did not. In particular, deletions involving exons 46–48 determined loss of transcripts lacking exons 45–49 and 45–50, while deletion of exons 45–48 had no effect on these products. Similarly, deletions of exon 48 and of exons 45–47 abolished skipping of exons 45–53 while deletions of exons 45–48 preserved this product. These data suggest that secondary structure formation on dystrophin pre-mRNAs plays no or little role in directing alternative splicing events. Moreover, loss of some splicing variants as a consequence of mutations deleting part of skipped exons indicates that factors other than the relative strength of exon/exon association might be involved in directing splicing patterns. Conversely, these evidences might be consistent with the hypothesis that the different splicing products reflect minor and major splicing pathways in the dystrophin gene.

In two cases patients carrying the same exonic deletion displayed different splicing behaviours with respect to the loss or preservation of alternative transcript lacking exons 44–56 (see Table 1). This is not surprising since apparently similar deletions are supposed to have different extension in flanking introns and, consequently, to have a diverse impact on splicing mechanisms. Indeed, given the tissue specificity of splicing patterns, the same deletion might also be expected to originate diverse effects in different tissues. Transcript analysis was extended to skeletal muscle, brain and heart tissues derived from a DMD patient carrying a deletion of exons 45–52: a different pattern of alternative transcripts was detectable in the three tissues (Table 1 and Fig. 3). In particular, brain and heart tissues of this patient preserved both the transcript lacking exons 44–56 and the one lacking exons 45–53, while this latter product could not be detected in skeletal muscle.

For both BMD and DMD, considerable variability in disease presentation is found across patients carrying the same

mutation and, in both pathologies, muscle involvement can be associated with heart disease and/or mental retardation; nonetheless these manifestations seem not to be consistent with the severity of the muscular phenotype. Our data indicate that alternative splicing events are differentially regulated in different organs and that deletions involving the same exons can determine diverse splicing behaviours in different patients or even in different tissues of the same individual. In this view, allelic differences and tissue specificity in splicing factors should be regarded as possible determinants of disease expression and differential organ involvement.

Acknowledgements: We are grateful to Dr Maria Teresa Bassi for useful discussion about the paper.

References

- [1] Nishio, H., Takeshima, Y., Narita, N., Yanagawa, H., Suzuki, Y., Ishikawa, Y., Ishikawa, Y., Minami, R., Nakamura, H. and Matsuo, M. (1994) *J. Clin. Invest.* 94, 1037–1042.
- [2] Koenig, M., Monaco, A.P. and Kunkel, L.M. (1988) *Cell* 53, 219–228.
- [3] Ahn, A.H. and Kunkel, L.M. (1993) *Nat. Genet.* 3, 283–291.
- [4] Klamut, H.J., Gangopadhyay, S.B., Worton, R.G. and Ray, P.M. (1990) *Mol. Cell. Biol.* 10, 193–205.
- [5] Nudel, U., Zuk, D., Einat, P., Zeelon, E., Levy, Z., Neuman, S. and Yaffe, D. (1989) *Nature* 337, 76–78.
- [6] Gorecki, D.C., Monaco, A.P., Derry, J.M.J., Walker, A.P., Barnard, E.A. and Barnard, P.J. (1992) *Hum. Mol. Genet.* 1, 505–510.
- [7] Bies, R.D., Phelps, S.F., Cortez, M.D., Roberts, R., Caskey, C.T. and Chamberlain, J.S. (1992) *Nucleic Acids Res.* 20, 1725–1731.
- [8] Feener, C.A., Koenig, M. and Kunkel, L.M. (1989) *Nature* 338, 509–511.
- [9] Surono, A., Takeshima, Y., Pramono, Z.A. and Matsuo, M. (1997) *Biochem. Biophys. Res. Commun.* 239, 895–899.
- [10] Koenig, M., Beggs, A.H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Muller, C.R., Lindlof, M., Kaariainen, H., de la Chapelle, A., Kiuru, A., Savontaus, M.L., Gilgenkrantz, H., Recan, D., Chelly, J., Kaplan, J.C., Covone, A.E., Archidiacono, N., Romeo, C., Liechti-Gallati, S., Schneider, V., Braga, S., Moser, H., Darras, B., Wrogemann, K., Blonden, L.A., van Paassen, H.M.B., van Ommen, G.J.B. and Kunkel, L.M. (1989) *Am. J. Hum. Genet.* 45, 498–506.
- [11] Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H. and Kunkel, L.M. (1988) *Genomics* 2, 90–95.
- [12] Winnard, A.V., Klein, C.J., Coovert, D.D., Prior, T., Papp, A., Snyder, P., Bulmann, D.E., Ray, P.A., McAndrew, P., King, W., Moxley, R., Mendell, R. and Burghes, H.M. (1993) *Hum. Mol. Genet.* 2, 737–744.
- [13] Winnard, A.V., Mendell, J.R., Prior, T., Florence, J. and Burghes, H.M. (1995) *Am. J. Hum. Genet.* 56, 158–166.
- [14] Roberts, R.G., Barby, T.F., Manners, E., Bobrow, M. and Bentley, D.R. (1991) *Am. J. Hum. Genet.* 49, 298–310.
- [15] Liu, H.X., Zhang, M. and Krainer, A.R. (1998) *Genes Dev.* 12, 1998–2012.
- [16] Liu, H.X., Chew, S.L., Cartegni, L., Zhang, M.Q. and Krainer, A.R. (2000) *Mol. Cell. Biol.* 20, 1063–1071.
- [17] Sironi, M., Pozzoli, U., Cagliani, R., Comi, G.P., Bardoni, A. and Bresolin, N. (2001) *Hum. Genet.* 109, 73–84.
- [18] Stark, J.M., Cooper, T.A. and Roth, M.B. (1999) *J. Biol. Chem.* 274, 29838–29842.
- [19] Gudas, J.M., Knight, G.B. and Pardee, A.B. (1990) *Mol. Cell. Biol.* 10, 5591–5595.
- [20] Kessler, O., Jiang, Y. and Chasin, L.A. (1993) *Mol. Cell. Biol.* 13, 6211–6222.