Human Molecular Genetics, 2000, Vol. 9, No. 4 605–616

Constitutive and regulated modes of splicing produce six major myotonic dystrophy protein kinase (DMPK) isoforms with distinct properties

Patricia J.T.A. Groenen[+], Derick G. Wansink, Marga Coerwinkel, Walther van den Broek, Gert Jansen[§] and Bé Wieringa[¶]

Department of Cell Biology, Medical Faculty, University of Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands

Received 28 October 1999; Revised and Accepted 17 December 1999

Myotonic dystrophy (DM) is the most prevalent inherited neuromuscular disease in adults. The genetic defect is a CTG triplet repeat expansion in the 3′-untranslated region of the myotonic dystrophy protein kinase (*DMPK*) gene, consisting of 15 exons. Using a transgenic *DMPK*-overexpressor mouse model, we demonstrate here that the endogenous mouse *DMPK* gene and the human *DMPK* transgene produce six major alternatively spliced mRNAs which have almost identical cell type-dependent distribution frequencies and expression patterns. Use of a cryptic 5′ splice site in exon 8, which results in absence or presence of 15 nucleotides specifying a VSGGG peptide motif, and/or use of a cryptic 3′ splice site in exon 14, which leads to a frameshift in the mRNA reading frame, occur as independent stochastic events in all tissues examined. In contrast, the excision of exons 13/14 that causes a frameshift and creates a C-terminally truncated protein is clearly cell type dependent and occurs predominantly in smooth muscle. We generated all six full-length mouse cDNAs that result from combinations of these three major splicing events and show that their transfection into cells in culture leads to production of four different ~74 kDa full-length (heart-, skeletal muscle- or brain-specific) and two C-terminally truncated ~68 kDa (smooth muscle-specific) isoforms. Information on *DMPK* mRNA and protein isoform expression patterns will be useful for recognizing differential effects of $(CTG)_n$ expansion in DM manifestation.

INTRODUCTION

Myotonic dystrophy (DM) is an autosomal dominant multisystemic disorder involving myotonia and progressive wasting of skeletal muscles, cardiac conduction defects, cataracts, endocrine disturbance, mental retardation and cognitive deficiencies (1). The disease belongs to the class of trinucleotide expansion disorders and is caused by the amplification of an unstable $(CTG)_n$ repeat in the 3′-untranslated region (3′-UTR) of the DM protein kinase gene (*DMPK*) (2–5). An increase in $(CTG)_n$ repeat size correlates with earlier onset and more pronounced disease manifestation and forms the basis for genetic anticipation in DM families. Unfortunately, the molecular and cellular consequences of $(CTG)_n$ repeat expansion are still poorly understood. Mouse models for abnormal DMPK expression display only mild myopathies (10,11), suggesting that DM is not caused by a simple loss or gain of DMPK function alone, but that the presence of a $(CTG)_n$ repeat is critical and has pleiotrophic consequences. It is now believed that the DM mutation results in *cis* or *trans* effects on processing or nuclear routing of *DMPK* and other mRNAs. In turn, this may result in disregulation of DMPK protein kinase activity, but the expression of other proteins such as cardiac troponin T may also be affected (12) via binding by, or titration of, CUG-binding proteins (13). The situation becomes even more complex when one takes into consideration that besides *DMPK*, at least two other genes, *DMAHP/SIX5* and *DMWD* (formerly called gene 59 and DMR-N9 in man and mouse, respectively), may be involved via direct interference with the transcriptional process (6–9, and references therein).

Computer analysis of the *DMPK* gene sequence of man and mouse, which is composed of 15 exons, predicts protein product(s) with a composite domain structure (Fig. 1B). In the full-length polypeptide, five distinct domains with distinct functional roles can be distinguished. The 40 amino acid N-terminal stretch is particularly leucine rich and may play a role as an aggregation or routing signal in the cell. Exons 2–8 specify a kinase domain with the 11 subregions characteristic for members of the serine/threonine-type subfamily of kinases (14). Immediately following the kinase domain there is a five amino acid VSGGG peptide sequence with unknown function. Exons 10–12 encode an α-helical domain with homology to domains of myofibrillar and filamentous proteins, which may be involved in the formation of coiled-coil structures or self-association of the enzyme. Exons 12–15 can encode various C-termini, possibly acting as self-association domains (15) or as membrane anchors. Based on the sequence homology which extends to beyond the kinase domain (16,17) and on the identification of specific

[+]Present address: Department of Pathology, University Hospital Nijmegen, PO Box 9101, 6500 HB Nijmegen, The Netherlands
[§]Present address: Department of Molecular Biology, Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, The Netherlands
[¶]To whom correspondence should be addressed. Tel: +31 24 3614329; Fax: +31 24 3540525; Email: b.wieringa@celbi.kun.nl

binding partners for DMPK (P.J.T.A. Groenen, unpublished data), DMPK can be considered as a member of the subfamily of Rho-kinases. Members of this family have a putative role in cell shape determination and in the regulation of actin–myosin contractility. Moreover, we have found that the *DMPK* gene product(s) may act as a modulator of the activity of voltage-gated ion channels (18). Previously, it was demonstrated that *DMPK* transcripts and proteins can be found in a wide range of tissues, with the highest expression in organs containing smooth muscle cell linings (stomach and colon) and in cardiac and skeletal muscles (most prominent in tongue, oesophagus and diaphragm). In the brain, a moderate overall level of expression is found, but this may be high in certain subregions of the central nervous system and negative in others (10,19). Subcellular fractionation and immunohistochemical studies described that the DMPK protein was localized at neuromuscular and myotendinous junctions (20,21) and terminal cisternae of the sarcoplasmic reticulum of skeletal muscle cells (22–24). Furthermore, the protein was found at intercalated disks (20,21) and the corbular and junctional sarcoplasmic reticulum of cardiac muscle cells (25). Although the data are not all consistent, it is clear that the localization of DMPK(s) is at or near sites of dense channel clustering.

To understand better the biological significance of the protein, its relevance for disease aetiology and its intracellular partitioning behaviour, basic knowledge about the primary structure and expression behaviour of DMPK(s) is a prerequisite. Initially, a variety of alternatively spliced *DMPK* mRNA isoforms were identified in mouse and man (4) (Fig. 1A). However, subsequent studies, based on sequencing of cDNAs in libraries prepared from various tissues (4,26–28), indicated that not all products are equally prevalent. Based on available data, the splicing pattern in mouse may be more complex than in humans. Here, we report the use of a transgenic mouse model, containing a 15 kb human genomic DNA fragment encompassing the entire human *DMPK* gene, to make a detailed comparison between isoform products of the human and the endogenous mouse *DMPK* gene. The results presented indicate that the products of both species behave similarly with regard to splice preference, cell type dependency and frequency of occurrence. Altogether, six distinct isoforms make up the DMPK profile in different cell types. Interestingly, both the VSGGG segment and the C-terminus confer distinct properties on DMPK with regard to modification and routing behaviour. The availability of the complete set of *DMPK* gene products will help us in further exploiting cellular and animal models for DM disease manifestation, in studying the cell signalling role of distinct isoforms and in recognizing the *cis* and *trans* effects that $(CTG)_n$ expansion may have on RNA processing in different cell types.

RESULTS

The existence of various *DMPK* mRNA isoforms due to alternative splicing has been shown by several groups, including ours (2–4). Figure 1A gives a summary of cDNAs identified in libraries prepared from various human and mouse tissues. In mouse, but not in human, the alternative splicing of an 87 nucleotide sequence within exon 8 (region I, encoding 29 amino acids that contribute to the protein kinase domain) was observed. Furthermore, the use of a cryptic 5′ splice site in exon 8 has been reported, resulting in the inclusion or deletion of the last 15 nucleotides from this exon (region II). This segment spans the short VSGGG amino acid motif which is a putative glycosaminoglycan addition site. Alternative splicing of exon 10 (region III) occurs only in mouse. This event leads to a frameshift and introduces a premature stop codon. Also, insertion of intron sequences (regions IV and V) is predicted based on the identification of alternative transcripts in a mouse brain cDNA library (4). Again, these events would result in the use of premature stop codons and yield C-terminally truncated protein products. Another DMPK variant in both mouse and human arises as a result of alternative skipping of exons 13 and 14, and involves the direct fusion of exons 12 and 15 (region VI). This causes a frameshift which terminates translation directly at the beginning of exon 15 and therefore creates a short C-terminal end. Finally, the activation of a cryptic splice acceptor site in exon 14 (region VII), located close to its upstream border, leads to removal of an additional four internal nucleotides from the mRNA, and is thus associated with a frameshift in the reading frame. The predicted isoform has a C-terminal tail with similar length to the other long isoform, but is less hydrophobic (4).

We decided to make a more detailed comparison of the splicing behaviour of the human and mouse genes by utilizing a mouse model that carries multiple copies of the normal (i.e. non-expanded) human *DMPK* transgene (developed by our group; see ref. 10). This overexpressor mouse provides easy access to a broad range of fresh tissues and enables us to study the products of human and mouse *DMPK* splice events in parallel. To obtain insight into the consequences of alternative splicing at the protein level, we prepared western blots loaded with representative human tissue extracts and different tissue extracts of *hDMPK* transgenic, wild-type and *DMPK* knock-out mice (10) as negative controls and incubated these with one of our polyclonal rabbit antisera (antiserum B79) directed against recombinant full-length mDMPK C (Fig. 1B). It should be stressed that although our B79 antiserum was raised against one purified recombinant isoform (DMPK C, with tail b) of DMPK from mouse, it adequately detects other mouse and human DMPK isoforms when expressed in COS-1 cells (Figs 2 and 3B, lanes 5 and 6), but is not absolutely monospecific (Fig. 3A). We know that this phenomenon is observed commonly with different independently generated antibodies such as the anti-pep2 Ab (20), the anti-DMK Ab (provided by R.G. Korneluk; see refs 10,21,29) or the anti-DMPK mouse monoclonal antibodies (30). As all antibody preparations mentioned have been tested on protein blots from tissues of our mouse models, we know that larger and/or smaller polypeptides (~200, ~55 and ~45 kDa) are always being recognized. We conclude that the cross-reacting proteins share ubiquitous sequence similarity with the DMPK catalytic or coiled-coil domains, for which several strongly related homologues can be found nowadays in the protein databases. Antiserum B79 was chosen because it consistently yielded the most intense signals on extracts of cells in culture that express *DMPK* gene products (data not shown). As shown in Figure 3A, this antiserum detects a broad protein band that clearly displays heterogeneity, with an apparent molecular mass of ~74 (70–78) kDa in tongue, heart, skeletal muscle and brain of *hDMPK* transgenic and wild-type mice. The heterogeneous ~74 kDa signal is greatly increased in transgenic overexpressor compared with wild-type and is absent from the knock-out mouse tissues, confirming that the ~74 kDa products are authentic DMPK proteins. Moreover, the antiserum is also

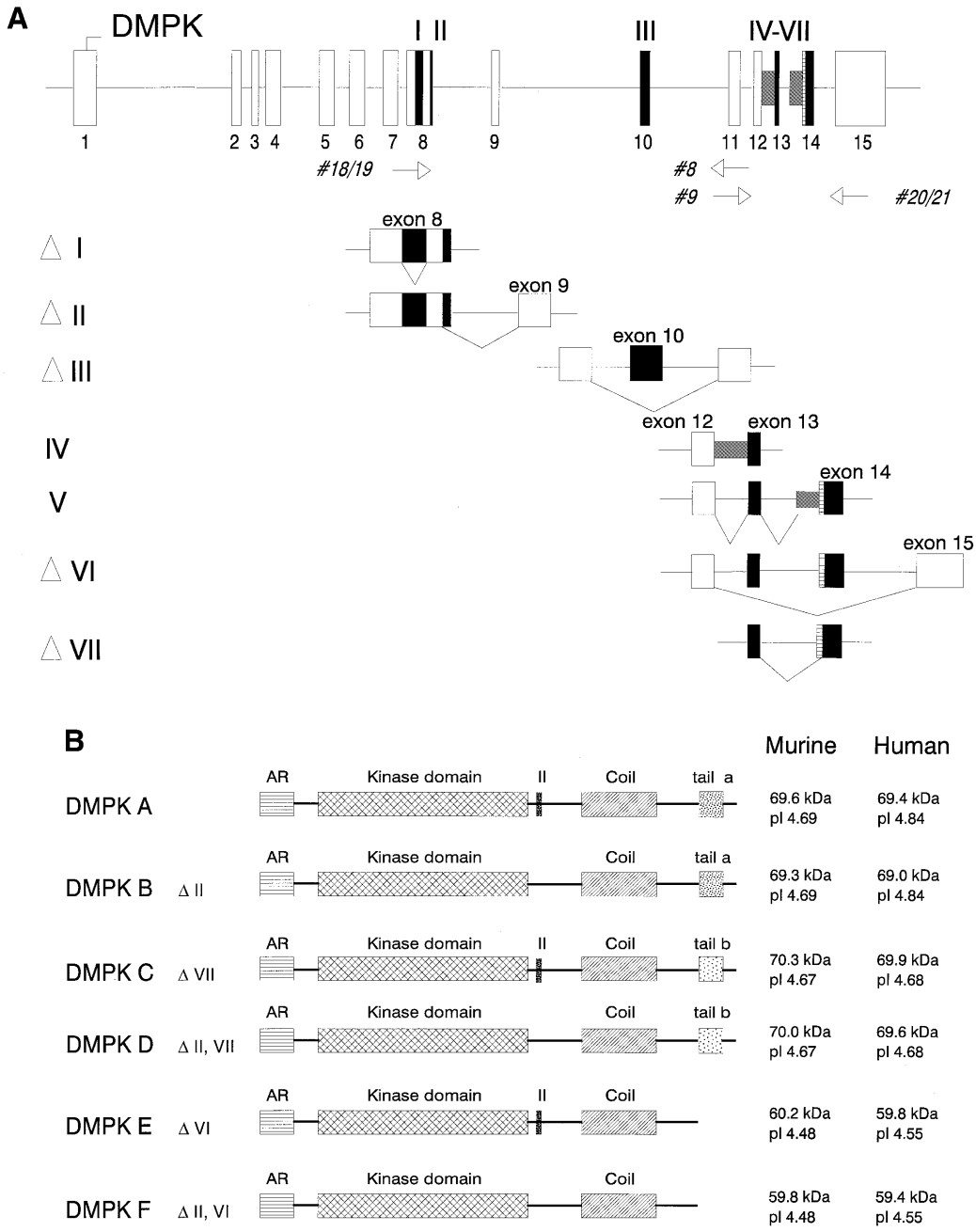

Figure 1. (**A**) Intron–exon organization of the *DMPK* gene and possibilities for alternative use of splice sites based on cDNA information. Exons are depicted as open boxes or, if alternatively spliced, as black boxes. Cryptic intron segments are shown as grey boxes. Splicing of regions I (deletion of nucleotides 983–1069 of exon 8), III (deletion of nucleotides 1233–1344 of exon 10), IV (insertion of complete intron 12) and V (insertion of part of intron 13 between nucleotides 1653 and 1654) is presumably mouse specific. Alternative use of regions II (deletion of nucleotides 1132–1146 of exon 8), VI (complete deletion of exons 13 and 14) and VII (deletion of nucleotides 1654–1657 of exon 14) occurs in both man and mouse. Nucleotide positions are given as cDNA positions, the numbering of nucleotides beginning at the translation start (27). (**B**) Structural domain organization of the major mouse and human DMPK isoforms. All predicted DMPK isoforms (A–F) contain an N-terminal leucine-rich stretch, a kinase domain encoded by exons 2–8 and an α-helical domain that shows significant homology to protein segments capable of forming coiled-coil structures. DMPK variants differ in the absence or presence of the VSGGG motif encoded by region II, a hydrophobic (tail a) or less hydrophobic C-terminus (tail b), encoded by splice mode VII, or a truncated C-terminal end (splice mode VI). The molecular masses and the predicted pIs of the mouse and human DMPK isoforms are given on the right.

highly reactive against ~74 kDa protein(s) in human diaphragm (Fig. 3B, lane 1) and against DMPK produced from a full-length human cDNA (hDMPK A) in transfected COS-1 cells (Fig. 3B, lane 5). In mouse stomach (Fig. 3A), a broad signal of ~68 kDa is predominantly present. This protein band is also seen, albeit in a lower amount, in heart. In human uterus, mainly composed of smooth muscle cells, the ~68 kDa band is prominent as well (Fig. 3B, lane 2). Like the ~74 kDa protein band(s), the ~68 kDa band(s) clearly displays heterogeneity in all tissue extracts examined. As the ~68 kDa band(s) is

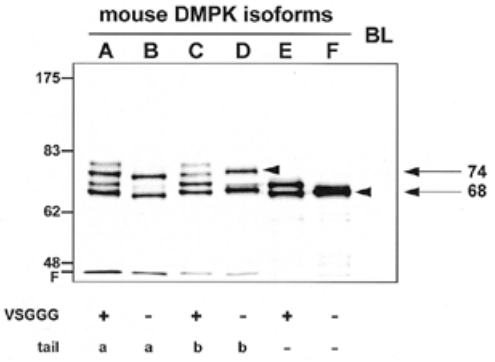

Figure 2. Mouse DMPK isoforms expressed in COS-1 cells. Homogenates of COS-1 cells expressing individual mouse *DMPK* cDNA products were separated via SDS–PAGE and transferred to nitrocellulose. A typical example of a western blot, incubated with the B79 anti-DMPK polyclonal antiserum, is shown. Multiple DMPK proteins are observed. The (apparent) molecular masses of the major products in each lane were determined using the markers shown on the left (in kDa; F, front). Shown are: DMPK A, ~78, ~73, ~70 and ~67 kDa (lane A); DMPK B, ~72 and ~66 kDa (lane B); DMPK C, ~78, ~73, ~70 and ~67 kDa (lane C); DMPK D, ~74 (arrowhead) and ~68 kDa (lane D); DMPK E, ~70 and ~67 kDa (lane E); and DMPK F, ~67 kDa (arrowhead, lane F). No DMPK-reactive proteins are observed in COS-1 extracts expressing an unrelated cDNA product (lane BL). Note that the largest proteins in the lanes of isoforms DMPK A, C and E (~78 kDa in lanes A and C; ~70 kDa in lane E) result from the presence of the VSGGG motif (i.e. similar products are not seen in lanes B, D and F, respectively). The products of DMPK E and F are less heterogeneous and do not produce smaller sized (degradation) products (lanes E and F).

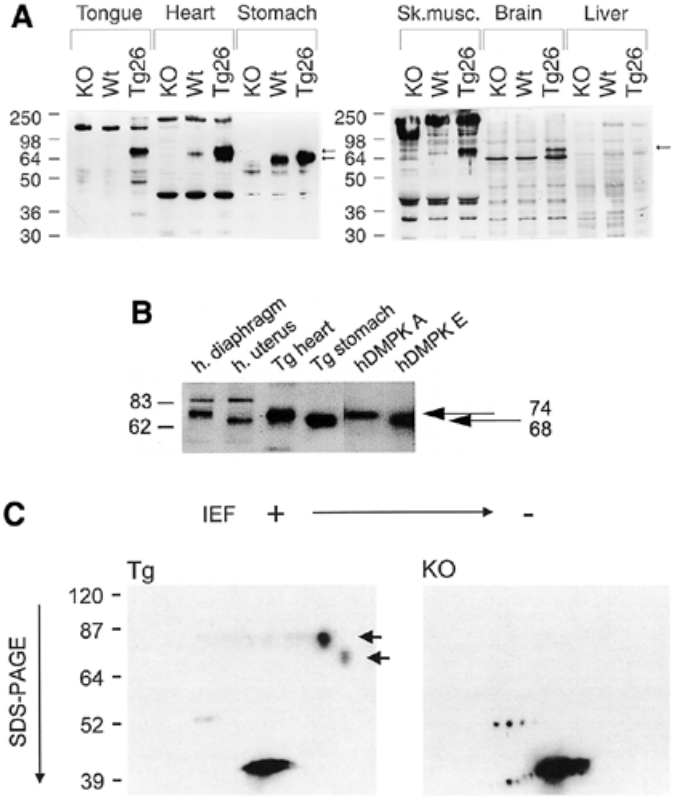

Figure 3. DMPK protein expression analysed on western blots using the B79 anti-DMPK polyclonal antiserum. (**A**) Western blots with tissue extracts from *DMPK* null mutant (KO), wild-type (Wt) and overexpressor (Tg) mice: tongue and stomach (2.5 mg), heart (3.0 mg), brain (7.5 mg), liver (5.0 mg) and skeletal muscle (9.0 mg). An ~74 kDa protein band, showing heterogeneity (upper arrow) is detected in tongue, heart, skeletal muscle and brain (and at a very low level in stomach) of the overexpressor mouse and to a lesser extent in the wild-type mouse tissues. In wild-type skeletal muscle and brain, this band is barely above the threshold of detection. Overexpressor and wild-type mice express a strong broad ~68 kDa band (lower arrow), again displaying heterogeneity, in the stomach. As anticipated, no DMPK proteins are observed in the liver. (**B**) Tissue homogenates of human diaphragm and uterus (2.5 mg), transgenic heart (1.5 g) and stomach (1.2 mg), and COS-1 cell extracts (<1 mg) expressing human DMPK A (+ I, +II, +III, –IV, –V, +VI, +VII) and DMPK E (+ I, +II, +III, –IV, –V, –VI, –VII). A heterogeneous ~74 kDa protein band (upper arrow) is detected in human diaphragm and in transgenic heart. This band(s) co-migrates with the full-length human DMPK A protein expressed in COS-1 cells. A heterogeneous ~68 kDa band (lower arrow), which co-migrates with a C-terminally truncated isoform (hDMPK E) expressed in COS-1 cells, is present in human uterus and in transgenic stomach. (**C**) Two-dimensional electrophoresis pattern of anti-DMPK-reactive proteins. Mixed tissue homogenates of heart and stomach (100 and 70 mg, respectively) of an overexpressor (Tg) mouse (left) and a *DMPK* null mutant (KO) mouse (right) were applied to the first dimension. Isoelectric focusing is in the horizontal direction, with the basic side on the left and the acidic side on the right. SDS–PAGE is in the vertical direction, from top to bottom. Protein dots due to cross-reactivity of the anti-DMPK polyclonal antibody are seen in both KO and Tg tissues, whereas DMPK-specific proteins (indicated by arrows) are only present in Tg tissues. Both the ~68 kDa (with lowest pI, lower arrow) and the ~74 kDa dots (higher pI, upper arrow) display heterogeneity; however, individual isoforms cannot be resolved. Molecular mass markers (in kDa) are indicated.

completely absent from tissue extracts of a *DMPK* knock-out mouse and co-migrates with the C-terminally truncated human DMPK protein expressed in COS-1 cells (hDMPK E; Fig. 3B, lane 6), we consider them to be authentic DMPK proteins.

In Figure 1B, we list several of the DMPK isoforms predicted on the basis of computer analyses of spliced cDNAs. Strikingly, none of these products has a size of 74 kDa, but instead almost similar molecular masses of 69–70 kDa are inferred for four different full-length isoforms, namely for the full-length product of the *DMPK* mRNA that contains all alternatively spliced segments and for products of mRNAs that have lost segments VII, II or II + VII. Likewise, both products from mRNAs that have lost either region VI or regions II + VI have an almost identical predicted size of 59–60 kDa for both mouse and man. To resolve the presumed heterogeneity in the (experimentally determined) ~74 and ~68 kDa proteins, and confirm identity with the (theoretically predicted) 69–70 and 59–60 kDa isoforms listed in Figure 1B, we decided to perform two-dimensional gel electrophoresis in combination with western blotting on mixed heart and stomach extracts of the overexpressor mouse (Fig. 3C, left). The ~74 and ~68 kDa categories of products do not yield sharp and spot-like signals but clearly show microheterogeneity, in particular in the vertical dimension, suggesting that they are made up of multiple DMPK isoforms. The pI of the proteins belonging to the ~68 kDa size class is more acidic than that of the ~74 kDa products, consistent with computer predictions for DMPK proteins with truncated and long C-terminal tails as listed in Figure 1B. From these pI data and from the migration behaviour of human DMPK proteins expressed in COS-1 cells (Fig. 3B, lanes 5 and 6), we conclude that the ~68 kDa signal represents the theoretically predicted 59–60 kDa DMPK isoforms with truncated C-terminal ends, and that the ~74 kDa signal represents the theoretically predicted group of 69–70 kDa DMPK isoforms with long C-terminal tails. Unfortunately, presumably due to the very small differences in pIs as predicted for DMPK isoforms

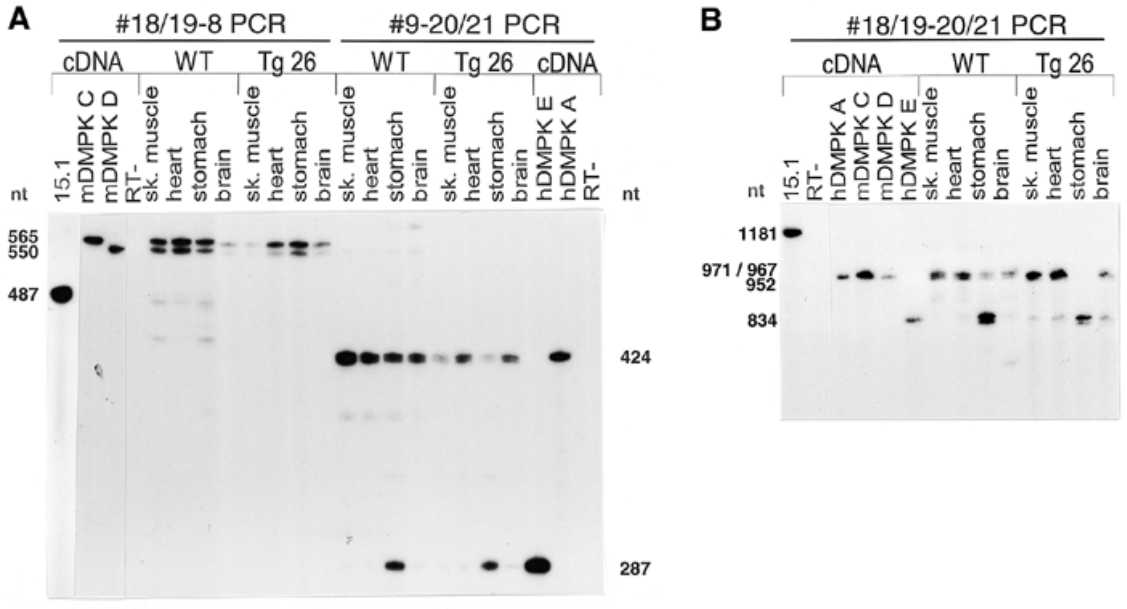

Figure 4. RT–PCR analysis of alternative splicing events in mouse and human *DMPK* mRNAs. RNAs were prepared from tissues of wild-type and transgenic *hDMPK* overexpressor mice. cDNA reverse transcription starting material was generated with antisense primer #10, but similar results were obtained with oligo(dT) or random hexamer primers. The RT–PCR products from skeletal muscle, heart, stomach and brain RNAs were resolved on sequencing gels and visualized by autoradiography as described in Materials and Methods. (**A**) Differential analysis of splice site use in the 5′ and 3′ half of *DMPK* mRNAs. Primer combination #18 (mouse-specific) or #19 (human-specific) plus #8 (human/mouse-specific) provides information about alternative skipping of regions I, II and III in the mouse (endogenous) and human (transgenic) *DMPK* mRNAs. Products from mouse cDNAs mDMPK C (+I, +II, +III; 565 nucleotides), mDMPK D (+I, −II, +III; 550 nucleotides) and 15.1 mDMPK (–I, +II, +III; 487 nucleotides) were used as markers. Products with (565 nucleotides) and without (550 nucleotides) region II are the most prominent and occur in about equal amounts. Primer combination #9 (human/mouse-specific) plus #20 (mouse-specific) or #21 (human-specific) was used to reveal splicing events involving regions IV, V, VI and VII. Products from human hDMPK A (–IV, –V, + VI, +VII) and hDMPK E (–IV, –V, –VI, –VII) cDNAs were used as markers for the anticipated 424 and 287 nucleotide products, respectively. Alternative splicing of region VI (exons 13 and 14) is most prominent in RNA from stomach. No bands are detected in RT^- reactions (right lane). (**B**) Analysis of combinatorial events in splicing patterns of *DMPK* mRNAs. The possibility of the interdependence of splicing reactions was analysed using the primer combinations #18/20 (mouse-specific) and #19/21 (human-specific) which bracket the largest possible span in the *DMPK* mRNA. Products from control cDNAs hDMPK A (+I, +II, +III, −IV, –V, +VI, +VII), hDMPK E (+I, +II, +III, –IV, –V, –VI, –VII), mDMPK C (+I, +II, +III, –IV, –V, +VI, –VII), mDMPK D (+I, –II, +III, –IV, –V, +VI, –VII) and 15.1 mDMPK (–I, +II, +III, +IV, +V, +VI, +VII) were used as 971, 834, 967, 952 and 1181 nucleotide markers, respectively. For mouse and human *DMPK* mRNAs of skeletal muscle, heart, stomach and brain, a doublet of the 971–952 nucleotide full-length products (± splice VII and ± splice II) is seen. Alternative splicing of exons 13 and 14 (region VI), seen as a doublet (± splice II) at 834 nucleotides, is predominant in stomach.

within one given size class (pI 4.68–4.84 for human proteins in the ~69–70 kDa range; pI 4.55 for both proteins in the 59–60 kDa range), we were not able to infer the number and identity of DMPK isoforms present within each group. Due to reasons unknown, signals of proteins in the predicted pI 4.68–4.84 range always appeared superimposed.

We therefore performed RT–PCR experiments to obtain detailed information at the RNA level, in order to establish unequivocally the distribution frequencies and structural identities of *DMPK* splice products. To this end, we used the primer combination #18/19 (mouse/human, respectively) plus #8 for the RT–PCR in order to distinguish between inclusion or absence of alternatively spliced regions I, II and III. In skeletal muscle, heart, stomach and brain of transgenic overexpressor mice (Fig. 4A), a normally sized product of 565 nucleotides, corresponding to a fragment containing region I, II and III as in the mDMPK C control, was found. In addition, in 40–60% of mRNAs, region II was spliced out, resulting in a product that co-migrates with the control that lacks this 15 nucleotide region (mDMPK D: 550 nucleotides). The same processing with the alternative splicing of region II is found for the endogenous mouse *DMPK* pre-mRNA, as shown by the mouse-specific PCR in Figure 4A. We also observed endogenous faint bands at ~487 nucleotides. One of these co-migrates with the signal derived from the 15.1 mDMPK cDNA that lacks region I. The splitting into two bands at this position reflects the absence or presence of region II in this splice product. An even weaker band at 453 nucleotides, probably reflecting alternative use of region III (112 nucleotides), is also present. Although these mouse-specific isoforms are expressed at a very low level (<2% of the predominant 565/550 nucleotide forms), they are present in all mouse tissues analysed and were identified in cDNA libraries earlier (4). Next, we used primer combination #9 plus #20/21 (mouse/human, respectively) to study the occurrence of splicing in the 3′ part of the pre-mRNA. All tissues from the transgenic mice express two large human-specific fragments of 424 and 420 nucleotides in equal amounts (Fig. 4A). The 424 nucleotide band, which co-migrates with the fragment amplified from a hDMPK A control cDNA, represents the full-length open reading frame containing the four nucleotides (region VII), whereas the 420 nucleotide band represents the isoform in which the four nucleotides are spliced out. In stomach, we found an additional prominent RT–PCR product of 287 nucleotides. This product co-migrates with the cDNA insert in the human smooth muscle DMPK (hDMPK E) and therefore represents the isoform that lacks exons 13 and 14 (region VI). The expression distribution for the endogenous mouse *DMPK* mRNA isoforms is similar to that of the human *DMPK* (Fig. 4A), except for the smaller isoform lacking region VI (predominantly in stomach),

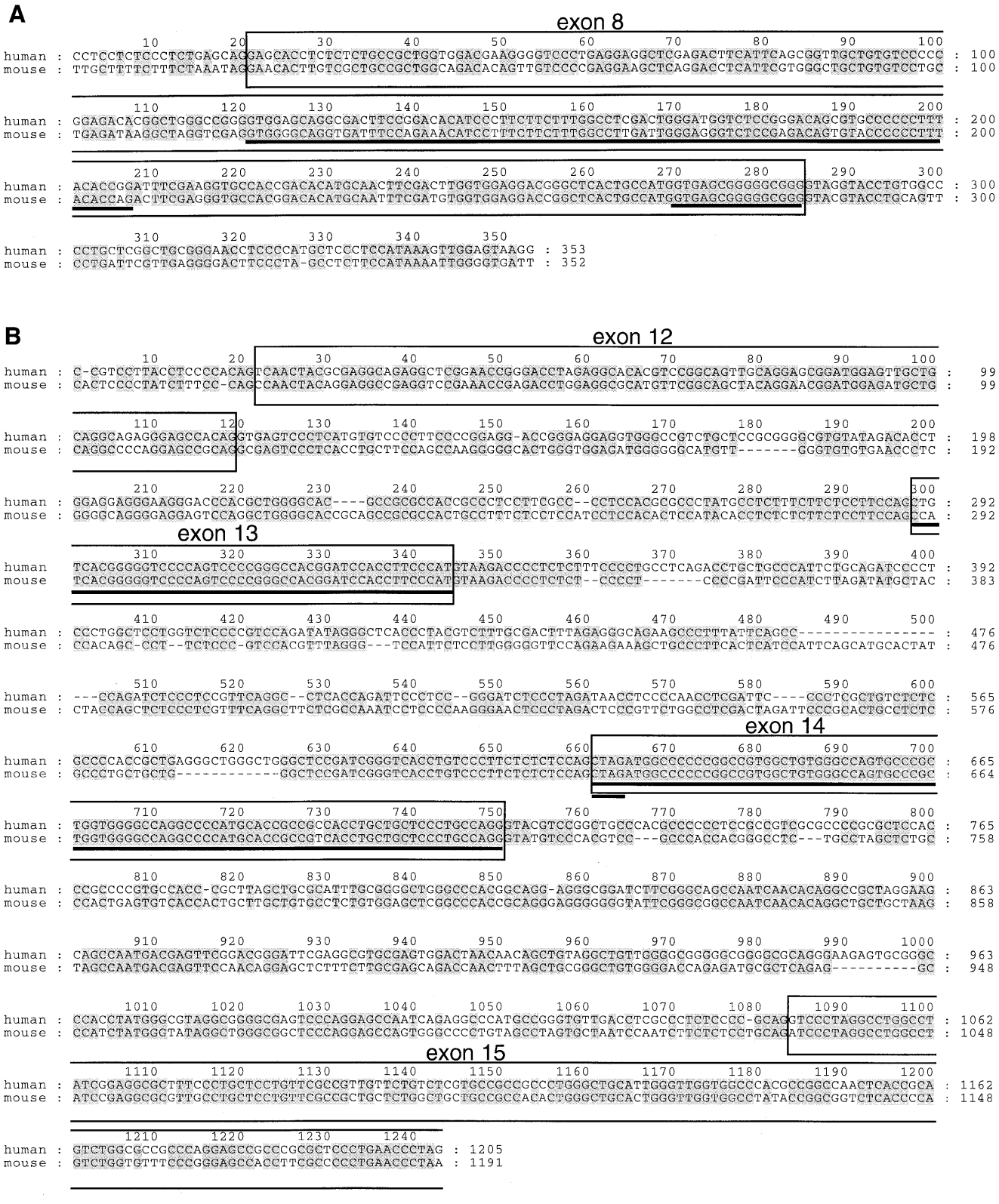

Figure 5. Alignment of genomic sequences across the relevant alternatively spliced segments in human and mouse *DMPK* genes: (**A**) exon 8 area and (**B**) exons 12–15 area, showing a high amount of sequence conservation. Exons are depicted in boxes; conserved positions are shaded. Alternatively spliced regions are underlined. It is of note that splicing of cryptic intron region I, the 87 nucleotide region within exon 8 which is flanked by 5′ G-A/G-G-g-t-g-g-g/a and 3′ a-c-c-a/g-g-A-C/T-T motifs, occurs at very low frequency and can only be used in mouse (4,27). Alternative splicing of the last 15 nucleotides of exon 8 (region II), the entire area containing exons 13 plus 14 (region VI, smooth muscle-specific) or the four nucleotide segment at the border of exon 14 (region VII) are frequent events in both human and mouse *DMPK* (this study).

which is expressed more prominently by the human than the mouse gene. It is of note that larger amplified fragments resulting from insertion of intron regions IV and/or V were not (or very weakly, in brain) observed among the mouse gene products.

Taken together, the results in Figure 4A demonstrate that the option to undergo splicing at the alternative cryptic 5′ splice site of region II (loss of 15 nucleotides), the cryptic 3′ splice acceptor site (region VII, four nucleotides) or the exon 13 and 14 regions is conserved in human and mouse *DMPK* pre-mRNA. This is in line with the strong homology found in the exon and intron sequences surrounding region II and exons 13 and 14 (Fig. 5) and indicates that other mRNA segments that differ drastically between the human and mouse pre-mRNAs (e.g. introns 12, 13 and 14 and exon 15 untranslated region; data not shown) have no strong *cis* effect. Next, we addressed the question of whether these alternative splice modes and the exon 13/14 skipping event should be considered as independent, or mutually exclusive events. Therefore, we used primer combination #18/20 (mouse specific) and #19/21 (human specific) that bracket the entire relevant mRNA segment. As shown in Figure 4B, RNA preparations of skeletal muscle, heart, stomach and brain yielded large isoform products of 971–952 nucleotides, which co-migrate with the controls hDMPK A (971 nucleotides), mDMPK C (967 nucleotides) and mDMPK D (952 nucleotides), all derived from full-length cDNAs with exons 13/14 included. With our electrophoresis conditions, we could distinguish between products with the 15 nucleotide region II included (mDMPK C, 967 nucleotides) or excluded (mDMPK D, 952 nucleotides), but resolution was not sufficient to see the products with and without the four nucleotides (alternative splice mode VII, see below). Note that besides the 971–952 nucleotide full-length products, skeletal muscle, heart and brain of transgenic and wild-type animals also yielded some minor bands at ~834 nucleotides, which co-migrate with the hDMPK E control that lacks exons 13 and 14. In contrast, stomach RNA from wild-type mice (i.e. endogenous *DMPK* RNA) gave products in which these 834 nucleotide bands, again as doublets caused by additional splicing of region II, were much more prominent (Fig. 4B). In the human *DMPK* RNA from stomach of transgenic animals, these 834/819 nucleotide bands were in fact the only variants expressed (Fig. 4B), which is in line with the PCR results shown in Figure 4A (#9-#20/21 PCR panel). Ultimately, several of the products were cloned and sequenced to confirm their identity (data not shown).

Although we cannot use our autoradiograms for quantitative interpretation, it is possible to draw some general conclusions regarding the frequency distribution of the different *DMPK* mRNA isoforms. Clearly, the excision of the exon 13/14 regions is a tissue cell type-dependent event, and we must assume that the splicing machinery of smooth muscle cells favours direct exon 12–15 joining. We know that also in the small and large intestine, this splice mode is the preferred one (data not shown). In contrast, the alternative use of region II (15 nucleotides) seems to be a default event. The recognition of splice sites in this region and across the exon 13/14 segment are non-related options, which occur with approximately equal likelihood, resulting in the production of doublet bands of (nearly) equal intensity. Finally, from hybridization experiments (Fig. 6) with cDNA primers with or without the four nucleotides (region VII), and from cloning and sequencing of individual RT–PCR products (data not shown) of the 971–952 nucleotide products, we know that each individual signal is in fact composed of co-migrating DNA fragments, with and without the four nucleotides, with approximately equal prevalence. Again, the alternative use of the four nucleotides (region VII) and the 15 nucleotides (region II) occurs independently. This means that the alternative use of the cryptic 3′ splice site in exon 14 is also a random event which is neither tissue type nor species-of-origin related nor coupled to use of other splice modes elsewhere in the pre-mRNA. Finally, it is important to note that the major splicing events are indeed confined to the mRNA segment examined here. Sequencing of 5′ RACE cDNA products from *DMPK* mRNAs yielded no evidence for heterogeneity in the 5′-UTR (data not shown).

Based on the above results, we decided to generate the complete set of six major mouse *DMPK* cDNAs in order to study whether absence or presence of the VSGGG motif, or any of the different C-termini, would confer unique physico-chemical properties on the *DMPK* products. For example, differences in mobility behaviour in SDS–polyacrylamide gels of the various isoforms can explain the heterogeneity observed in extracts from whole tissues (in which different isoforms are expressed simultaneously). To this end, six mouse *DMPK* cDNAs were introduced in the polylinker of a pSG8-derived eukaryotic expression vector (see Materials and Methods) and individually transfected into COS-1 cells. Care was taken in the preparation of extracts, by solubilizing transfected adherent COS-1 cells directly in hot SDS sample buffer followed by analysis on SDS–polyacrylamide gels and western blotting. We found several mDMPK products. A typical example of a blot, showing all possible mouse DMPK isoforms and their modifications, is shown in Figure 2. In lanes with long DMPK isoforms containing the VSGGG motif (i.e. Fig. 2, lanes A and C), we see four major bands of ~78, ~73, ~70 and ~67 kDa. In products without this motif, only two major bands in each lane are observed, namely ~72 and 66 kDa for DMPK B (Fig. 2, lane B), and ~74 and ~68 kDa for DMPK D (lane D). Strikingly, products corresponding to the smooth muscle isoforms (with or without the VSGGG motif) are less heterogeneous (Fig. 2, lanes E and F). Also here the VSGGG motif causes an additional band at higher molecular mass, but the splitting as seen for the products with long C-terminal ends (Fig. 2, lanes A–D) no longer exists.

Taken together, our results suggest that mDMPK isoforms undergo post-translational modification promoted by or at the VSGGG motif, causing a significant portion of the full-length proteins to shift to a higher molecular weight, i.e. from ~73 to ~78 kDa (lanes A and C) or from ~67 to ~70 kDa (lane E). In DMPK isoforms with either one of the two possible long C-terminal tails, partial proteolytic cleavage or a conformational transition is most likely to be involved in producing the ~66–68 kDa proteins (products with the smallest apparent molecular mass in Fig. 2, lanes A–D) from the ~72–74 kDa full-length products. A similar apparent size shift has been observed after expression of a long human DMPK isoform (not specified, but probably type A) in bacteria (21). Since no smaller sized protein bands are seen when expressing the C-terminally truncated DMPK isoforms (DMPK E and F) and since the truncated DMPK A–D isoforms display a migration behaviour in SDS–polyacrylamide gels which is similar to that of DMPK E and F (i.e. ~67 kDa), an eventual cleavage site must reside very close to the C-proximal border of the coiled-coil segment. The ~70 kDa bands observed in DMPK A and C probably reflect *DMPK* products which have lost (or refolded) their C-terminal tail, but have shifted because of a VSGGG motif-related modification (note that they show the same mobility as the

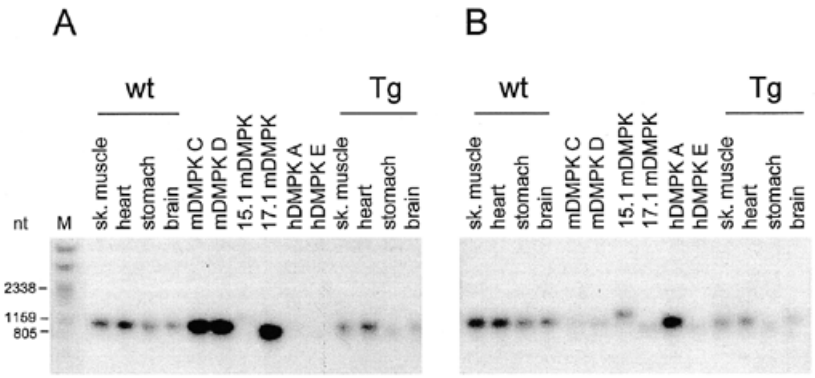

Figure 6. Southern blot for determining the prevalence of amplified RT–PCR fragments with or without the four nucleotide region VII of exon 14. Equal amounts of RT–PCR fragments generated with primer combinations #18/20 and #19/21 from skeletal muscle, heart, stomach and brain were resolved on agarose gels, blotted and subjected to hybridization with ^{32}P-labelled discriminative oligonucleotide probes HMAlt4 and -5 (see Materials and Methods). Hybridization in (**A**) is with an oligonucleotide probe that fits the mouse–human *DMPK* cDNAs without the four nucleotide insert (region VII), and in (**B**) with an oligonucleotide probe that fits the mouse–human *DMPK* cDNAs with the four nucleotides; note that the discriminative ability of the primer is not 100% perfect. Human cDNAs hDMPK A (+I, +II, +III, –IV, –V, +VI, +VII) and hDMPK E (+ I, +II, +III, –IV, –V, –VI, –VII), and mouse cDNAs mDMPK C (+ I, +II, +III, –IV, –V, +VI, –VII), mDMPK D (+I, –II, +III, –IV, –V, +VI, –VII), 15.1 mDMPK (–I, +II, +III, +IV, +V, +VI, +VII) and 17.1 mDMPK (+I, –II, –III, –IV, –V, + VI, –VII) were used as controls. Note that in the wild-type and Tg tissues examined, each individual RT–PCR product is in fact a doublet of co-migrating fragments with and without the four nucleotides, in approximately equal amounts.

presumably modified and shifted DMPK E isoform, the largest product in Fig. 2, lane E).

Interestingly, preliminary data indicate that the different mDMPK isoforms have slightly different subcellular localizations in COS-1 cells (data not shown). This might be due to the presence or absence of the VSGGG moiety, variation in the C-termini and/or differences in post-translational modification. Further studies are necessary to clarify the cell signalling role of each major isoform of DMPK and the translational efficiency of each *DMPK* mRNA.

DISCUSSION

This is the first report describing the tissue type-specific alternative splicing pattern of murine and human DMPK in detail. Based on sequence information from *DMPK* cDNAs identified in brain, muscle and heart cDNA libraries, a variety of alternatively spliced *DMPK* mRNA products, in particular for murine *DMPK*, were predicted (3,4,26–28). By using RT–PCR methodology, although this did not yield absolutely reliable quantitative data, we now show that human and mouse *DMPK* pre-mRNAs display similar and more simpler splicing patterns than anticipated. To study the splice products of both the human and mouse *DMPK* gene in parallel, we made use of our transgenic *DMPK* overexpressor mouse as an easily accessible source of tissues. Analysis of a few available human tissues indicated that indeed the same DMPK protein isoforms were present as in cognate transgenic mouse tissues (Fig. 3B). Hence, we consider this evidence that our transgenic *DMPK* overexpressor mouse provides the correct cellular context and is a faithful model for comparing (human- and mouse-specific) alternative splicing events.

Our findings imply that the *DMPK* gene specifies distinct protein products with different properties and/or functions. Differential use of regions II and VII only already accounts for the production of four of six major isoforms. In almost all tissue types, the full-length (exon 13/14-containing) transcripts are either present exclusively or predominantly. Skipping of exons 13 and 14 is prevalent selectively in smooth muscle, and to a lesser extent in heart [cDNAs corresponding to this truncated mRNA originally were isolated from the latter tissue (27)]. In our earlier work (4), we described the presence of two different splice acceptor sites in exon 14. It is now clear that both sites in exon 14 are used with about equal likelihood in all tissues examined. Likewise, the inclusion or deletion of the 15 nucleotide region in exon 8 are co-existing events. It is of note that these major splice modes are not coupled to the use of other splice events elsewhere in the *DMPK* pre-mRNA. Based on cDNA sequencing data, it was predicted that alternative splicing of region I, which encodes 29 amino acids that contribute to the homology with protein kinases, is a mouse-specific event (4). Now we show that this splice is very infrequent. We may therefore conclude that identification of the pertinent cDNA was a fortuitous finding and that the corresponding protein product has no biological significance, unless used in only a minor subset of specialized cells. An analogous situation was observed for alternative splicing of region III, or insertion of intron sequences from regions IV and V, which again occur at frequencies barely above the level of detection. We may therefore consider these minor mRNA isoforms as aberrant or incomplete splice products, presumably irrelevant for skeletal, heart and smooth muscle functioning.

Competition for recognition of the correct and cryptic 5′ and 3′ splice sites is a fundamental step in the process of constitutive alternative splicing. However, on analysis of the actual sequence information across the different splice sites in the *DMPK* gene, it is hard to categorize different sites as strong or weak binders and predict their effectiveness in splice choice competition. For example, based on their sequence motifs, the cryptic 5′ (GAGguggg) and 3′ (accagACU) splice sites that bracket region I in mouse exon 8 would be recognized as strong competing (see ref. 31 for consensus) splice sites, but they are hardly ever used. (Down)modulation of splice site strength may in this case be caused by the presence of a rather small polypyrimidine tract at an unusual distance upstream of the 3′ splice site. In contrast, the

non-consensus 5′ splice junction of mouse exon 12 (CAGgcgag) is fully functional. Obviously, this latter site is exposed more favourably to the splicing machinery or is not involved in splice site competition at all. The strongly conserved 5′ cryptic splice site in mouse and human exon 8 (region II) (Fig. 5A) is used frequently. It is accepted nowadays that tissue-specific alternative splicing, such as the switch in splicing pattern for *DMPK* between smooth muscle and other cell types observed here, involves both *cis*-acting regulatory RNA elements and *trans*-acting regulatory proteins that recruit the basal splicing machinery. For example, specific *cis*-acting elements essential for the cell type-dependent regulation of α-tropomyosin exon 3 splicing in smooth muscle have been identified. These sequence elements (URE and DRE), one in each of the introns flanking exon 3, are required for selective exon skipping in smooth muscle (32). Still, a sequence comparison revealed no stretches with clear homology to the URE or DRE sequences in the introns flanking exons 13 and 14 (nor in other introns) of human or mouse *DMPK*. As another mechanism, *trans*-regulation could result from the absence or presence of cell-specific regulatory factors such as p55 [as in the case of α-tropomyosin (33)] or from variation in the levels or activities of constitutive factors such as the polypyrimidine tract-binding protein (PTB) (34) or members of the serine/arginine (SR) family of proteins (35,36). Currently, evidence for involvement of any of these factors is lacking, and further study is necessary to elucidate which mechanism is most relevant for DMPK expression.

Our observation that alternative processing of human and mouse *DMPK* mRNA results in six main DMPK protein products, two ~68 kDa proteins and four ~74 kDa isoforms (Figs 1B, 2 and 3), is in line with previous reported data (10,15,21,29,30). Yet, recently published reports still describe the presence of ~43 to ~55 kDa *DMPK* products in rat and human brain (19,21) and in rat and human skeletal muscle (24,25,37) at rather high abundance. This is clearly at odds with our data, as the ~43 to ~55 kDa *DMPK* products are not predicted by the outcome of our RT–PCR experiments for human and mouse, nor by our transfection studies presented here. We also consider it less likely that these products are due to alternatively initiated translation (38), or to proteolytic cleavage of DMPK isoforms, as many of the smaller sized products are still seen in animals which are truly null for *DMPK* mRNA and protein expression (P.J.T.A. Groenen, unpublished data) (Fig. 3A). However, we should not completely rule out effects of cell type-dependent proteolysis, because products of single mouse *DMPK* cDNAs in COS-1 cells clearly undergo partial post-translational changes (Fig. 2). Taking all the evidence together, the appearance of lower molecular weight bands in blots of tissue extracts currently is best explained by the high frequency of aspecific reactivity in antibody preparations raised against DMPK.

Perhaps more interesting is the biological significance of the absence or presence of the distinct protein segments, the VSGGG motif and the different C-terminal ends, in the DMPK isoforms. The presence of the VSGGG motif specified by exon 8 is seen in ~50% of DMPK isoforms in man and mouse. Formerly, by using computer analysis, we predicted a role for this motif as a putative glycosaminoglycan addition site (4). Here, by comparing products from single cDNAs for each of the six major DMPK isoforms, we show that the presence of the motif produces a shift towards a slightly higher molecular mass in a significant portion of the proteins on transfection in COS-1 cells (Fig. 2). What we observe here is probably similar to the splitting of DMPK signals observed in several studies on (transgenic) mouse or human tissue extracts on western blots (Fig. 3) (10,21,22,24,25,29). It is interesting to note that 'modification' results in a still distinct migration behaviour of the *DMPK* product(s). We currently are studying the effects of cell type and growth conditions on the size shift and frequency of occurrence of this modification by transfection of cDNAs into different cell types. We probably will need mass spectroscopic analyses to discriminate between the possibility that the VSGGG motif promotes modification by a glycosaminoglycan moiety with uniform size and shape, undergoes phosphorylation or promotes another type of modification elsewhere in the protein.

We also expect clues on the significance of these modulary elements and differences in location of distinct DMPKs to come from functional studies on homologous proteins and the identification of binding partner proteins for the pertinent domains. It is therefore of note here that preliminary evidence points to a strong structural and functional homology between mouse and human DMPK and the RhoA- and Cdc42-associated kinases, such as p160 ROKα and -β (16,39) and the myotonic dystrophy kinase-related Cdc42-binding kinase [MRCKα and -β (40)], which both play a role in the myosin-based contractile activity (40,41). Besides having a high sequence homology in the N-terminal kinase motif, these proteins also share similarity in the overall domain organization, such as the presence of a coiled-coil α-helical domain in their C-terminal half. Importantly, in p160ROKα, the C-terminal region functions as a negative regulator of Rho kinase activity (41). We should therefore focus future studies on the functional comparison of the smooth muscle (C-terminally truncated) and skeletal muscle isoforms of DMPK. Computer-assisted comparison of the three possible alternative C-termini (data not shown) did not reveal any functional motifs. Whatever its role, whether it is in protein anchoring or scaffolding, intramolecular folding behaviour, membrane clustering or (self)association behaviour (15), we should remain aware of the fact that the C-terminus is among the most divergent segments between mouse and human DMPK isoforms. It may be that species comparison could therefore help to clarify this issue.

In conclusion, we have demonstrated that our transgenic overexpressor produces a 'near-natural' mixture of human and mouse DMPKs, with two ~68 and four ~74 kDa proteins as the major isoforms in heart, skeletal muscle, smooth muscle and brain. This mouse strain should therefore be considered as a useful model for ongoing attempts to unravel structure–function relationships of distinct DMPKs. Ultimately, detailed knowledge on *DMPK* alternative splicing patterns resulting in specific DMPK isoforms may help us in understanding the possible effects of the $(CTG)_n$ expansion on RNA processing. The availability of all distinct *DMPK* cDNAs will enable us also to study the distribution and translational fate (i.e. 'efficiency') of each individual *DMPK* mRNA *in vivo* and *in vitro* in more detail.

MATERIALS AND METHODS

Western blot analysis

Tissue samples were homogenized in lysis buffer as described (10), and proteins in the extract were analysed by electrophoresis on SDS–polyacrylamide (10% w/v) gels and/or two-dimensional gels (with expansion of the pI 4–6.5 range in the first dimension, and 10% SDS–polyacrylamide gels in the second dimension) and

subsequently transferred onto nitrocellulose membranes (Schleicher and Schuell, Keene, NH). After blocking with 5% skimmed milk in phosphate-buffered saline (PBS) containing 0.05% Tween-20 (PBST) for 60 min, blots were incubated in PBST overnight at room temperature with a polyclonal anti-DMPK antibody (B79). This antibody was raised in rabbits against purified recombinant full-length mouse DMPK C protein (+I, +II, +III, –IV, –V, +VI, –VII; see sequence in ref. 4) produced in *Escherichia coli* (pGEX expression system). After washing in PBST, the blots were incubated with peroxidase-conjugated goat anti-rabbit immunoglobulins (Pierce, Rockford, IL) at 1:10 000 dilution for 2 h, washed in PBST again and developed using a chemiluminescence western blotting reagent.

RNA preparation

Tissues were collected from 4-month-old male mice, snap-frozen in liquid nitrogen, thawed and homogenized in 4 M LiCl/8 M urea, and RNA was isolated using standard procedures. RNA concentrations were determined (from $E_{260\ nm}/E_{280\ nm}$ spectrophotometer readings) and all samples were treated with DNase I (Amersham Pharmacia Biotech, Roosendaal, The Netherlands) before use.

RT–PCR

Reverse transcriptase reactions were performed with 1 µg of RNA using M-MLV reverse transcriptase (Superscript; Gibco BRL, Breda, The Netherlands) according to standard protocols using oligo(dT) (100 ng/µl), random hexamers (100 ng/µl) or an antisense DMPK primer (#10, mouse- and human-specific: 5′-GGTGGGGACAGACAATAAA-3′;100 ng/µl) with the ability to hybridize just upstream of the poly(A) stretch. Amplification of specific regions of mouse and human DMPK transcripts was carried out with the forward primers #18 (mouse-specific: 5′-GCCGCTGGCAGACACAGTT-3′) and #19 (human-specific: 5′-GCCGCTGGTGGACGAAGGG-3′) and a reverse primer #8 (mouse- and human-specific: 5′-GCTCAGGCTCTGCCG-GGTGA-3′) to distinguish between the presence or absence of alternative spliced regions I, II and III. Likewise, primer #9 (forward, mouse- and human-specific: 5′-CTCACCCGGCAGAGCCTGA-3′) and reverse primers #20 (mouse-specific: 5′-ACCAGACT-GGGGTGAGACC-3′) and #21 (human specific: 5′-GCCAGACT-GCGGTGAGTTG-3′) were used to analyse alternative splicing of regions IV, V, VI and VII. The primer combination #18/#20 (mouse) or #19/#21 (human), which brackets the largest span of the mRNA, was used to analyse all possible splicing combinations. The approximate locations of primer sequences are given in Figure 1A.

For PCR, ~20 ng of the forward or reverse primer was end-labelled with [γ-^{32}P]ATP (Amersham Pharmacia Biotech) using T4 polynucleotide kinase. cDNA aliquots corresponding to 100 ng of RNA equivalents (one-tenth of the reverse transcriptase reaction volume) were PCR amplified with 100 ng of each primer by a 'hot start' protocol; *Taq* polymerase (0.75 U/reaction) and the radiolabelled primer were added at 80°C after denaturation at 96°C for 10 min. For amplification, 35 cycles consisting of incubation at 96°C for 1 min, 56°C for 1 min and 72°C for 3 min were used. Finally, as the last step, a chase at 72°C (10 min) was performed, all in a Perkin Elmer Cetus (Norwalk, CT) DNA thermal cycler. Amplified DNA fragments were resolved by electrophoresis on a 38 cm denaturing 4% polyacrylamide gel (60 W; for 6–9 h dependent on the fragment sizes) and visualized by autoradiography. Alternatively, Southern blot analyses were performed to determine the amounts of amplified RT–PCR fragments with or without region VII (four nucleotides). To this end, RT–PCR products were loaded onto a 1% agarose gel and oligonucleotides of 16 residues in length, HMAlt4 (5′-GGCCATCTAGATGGGA-3′: +4 nucleotides, reversed) and HMAlt5 (5′-GGGGCCATATGGGATG-3′: –4 nucleotides, reversed), end-labelled with [γ-^{32}P]ATP (Amersham Pharmacia Biotech) using T4 polynucleotide kinase, were used as discriminative probes in a standard hybridization assay (in 5× SSPE/0.3% SDS at 42°C).

Generation of six mouse DMPK cDNAs

Based on the use of 1.1 DMR-B15 cDNA (4) as starting material, we employed a 5′ RACE protocol to generate double-stranded mouse DMPK C cDNA (+I, +II, +III, –IV, –V, +VI, –VII) with a full-length open reading frame. *Bgl*II linkers were positioned just upstream of the ATG start codon and at the end of the 3′-UTR [i.e. at the beginning of the poly(A) tail] in order to clone the entire cDNA into the *Bgl*II site of pBluescript, yielding pBlmDMPK C.

To construct an mDMPK cDNA which is lacking region II but otherwise identical to cDNA C, pBlmDMPK C was used as template in a PCR with forward and reverse primers (5′-AT-GGCCATAGACTCCGTG-3′ and (5′-GCATGTCTGACAGCG-TCTCCATGGCAGTGAGCCGGT-3′, respectively). The reverse primer matches 18 nucleotides upstream and 18 nucleotides downstream of region II, but lacks the internal 15 nucleotide sequence of this region. Then, the resulting PCR product and *Nco*I-digested pBlmDMPK C were used as templates in a second round of PCR using the same forward primer in combination with another reverse primer (5′-CTGTAGTTGGCTGGAGAA-3′). From the product formed, a 297 bp *Nco*I–*Xmn*I fragment (lacking region II) was excised and used to replace the corresponding 312 bp *Nco*I–*Xmn*I fragment (containing region II) in pBlmDMPK C, resulting in the generation of pBlmDMPK D (+I, –II, +III, –IV, –V, +VI, –VII). Subsequently, the cDNA inserts of both pBlmDMPK C and D were excised as *Bgl*II–*Bgl*II fragments and cloned into the *Bgl*II site of the eukaryotic expression vector pSG8DEco [a modified version of pSG5 (42)], resulting in pSGmDMPK C and pSGmDMPK D, respectively.

Plasmids pSGmDMPK A (+I, +II, +III, –IV, –V, +VI, +VII) and B (+I, –II, +III, –IV, –V, +VI, +VII) were generated by four-way ligation. To this end, three fragments were cloned in the correct orientation into the *Bgl*II site of pSG8·Eco: (i) a 1.1 kb *Bgl*II–*Nco*I fragment from pSGmDMPK C; (ii) a 0.5 kb *Nco*I–*Xba* fragment obtained through PCR amplification on pSGm-DMPK C- (for pSGmDMPK A) or pSGmDMPK D- (for pSGm-DMPK B) derived templates with forward (5′-GGACC-GGCTCACTGCCATGGWGA-3′) and reverse (5′-GGCCA-TCTAGATGGGAAGGTGGATCCGTGGCCC-3′) primers; and (iii) a 1.0 kb *Xba*I–*Bgl*II fragment isolated via PCR amplification on a pBlmDMPK C-derived template with a (5′-ACACTCTA-GATGGCCCCCCGGCCGTGGCTGT-3′) forward and a T7 reverse primer. In this way, the four nucleotide CTAG sequence of region VII (absent in pSGmDMPK C and D) was introduced by fusion of the primer-derived *Xba*I sites flanking fragments 2 and 3 (underlined).

Finally, pSGmDMPK E (+I, +II, +III, –IV, –V, –VI, –VII) and pSGmDMPK F (+I, –II, +III, –IV, –V, –VI, –VII) were created by replacing the 592 bp *Bsp*EI–*Eag*I fragment (including region VI)

in either pSGmDMPK C or D, for a 459 bp *Bsp*EI–*Eag*I fragment (lacking region VI). The 459 bp *Bsp*EI–*Eag*I fragment was generated from mouse stomach *DMPK* mRNA by RT–PCR. The reverse transcriptase reaction was carried out with random hexamer primers and, for PCR, a combination of forward (5′-CTCACCCGGCAGAGCCTGA-3′) and reverse (5′-TGTGCTGGCAGAGGTCTT-3′) primers was used.

All six mDMPK A–F constructs were sequenced, and constructs that had an exact sequence match with the cognate exon segments in the published genomic sequence of the mouse myotonic dystrophy kinase gene (27) were chosen for further use. Only in the 3′-UTRs of certain cDNA isoforms were mouse strain-dependent sequence polymorphisms identified (to be reported elsewhere).

Transfection of mDMPK cDNAs in COS-1 cells

COS-1 cells were grown at 37°C under a 5% CO_2 atmosphere in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS). For transfection of the six mouse *DMPK* cDNAs, cells were cultured in 6-well plates and transfected using DEAE–dextran. In brief, COS-1 cells were washed once with Optimem (Life Technologies, Breda, The Netherlands) and incubated for 2 h in 1 ml of Optimem containing 2 μg of plasmid DNA, 50 μg/ml DEAE–dextran and 0.2 mM chloroquine. The transfection medium was removed and cells were incubated for 2 min in 1 ml of 10% dimethylsulfoxide (DMSO) in PBS. The DMSO solution was removed, standard DMEM supplemented with 10% FCS was added and cells were grown as described above. After 22–24 h, cells were washed twice in PBS and lysed in hot 2× SDS sample buffer. These whole-cell lysates were analysed using an 8% SDS–polyacrylamide gel and western blotting (see above). A construct unrelated to *DMPK*, encoding the FERM domain of PTP-BL, pSG8BL-FERM-VSV (43), was included as a negative control.

ACKNOWLEDGEMENTS

This study was supported by grants to B.W. from the Dutch Beatrixfonds, the American Muscular Dystrophy Association (MDA), the Associaton Française contre les Myopathies (AFM) and the Netherlands Organization of Scientific Research (NWO).

REFERENCES

1. Harper, P.S. (1989) *Myotonic Dystrophy*. 2nd edn. W.B. Saunders, London, UK.
2. Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.-P., Hudson, T. *et al.* (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell*, **68**, 799–808.
3. Fu, Y.-J., Pizutti, A., Fenwick, R.G.F., King, J., Rajnarayan, S., Dunne, P.W., Dubel, J., Nasser, G.A., Ashizawa, T., de Jong, P. *et al.* (1992) An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science*, **255**, 1256–1258.
4. Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendriks, W., O'Hoy, K., Baird, S., Sabourin, L., Lennon, G. *et al.* (1992) Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nature Genet.*, **1**, 261–266.
5. Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barceló, J., O'Hoy, K. *et al.* (1992) Myotonic dystrophy mutation: an unstable CTG repeat in the 3′ untranslated region of the gene. *Science*, **255**, 1253–1255.
6. Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Readon, W., Buckler, A.J., Harper, P.S., Housman, D.E. and Shaw, D.J. (1992) Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature*, **355**, 545–546.
7. Hunter, A.G.W., Tsilfidis, C., Mettler, G., Jacob, P., Mahadevan, M., Surh, L.C. and Korneluk, R.G. (1992) The correlation of age at onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *J. Med. Genet.*, **29**, 774–779.
8. Reddy, P.S. and Housman, D.E. (1997) The complex pathology of trinucleotide repeats. *Curr. Opin. Cell Biol.*, **9**, 364–372.
9. Brewster, B.S., Groenen, P. and Wieringa, B. (1998) Myotonic dystrophy: clinical and molecular analysis. In Emery, A.E.H. (ed.), *Neuromuscular Disorders: Clinical and Molecular Genetics.* John Wiley & Sons Ltd, pp. 323–364.
10. Jansen, G., Groenen, P.J.T.A., Bächner, D., Jap, P.H.K., Coerwinkel, M., Oerlemans, F., Van den Broek, W., Golsch, B., Pette, D., Plomp, J.J. *et al.* (1996) Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nature Genet.*, **13**, 316–324.
11. Reddy, S., Smith, D.B.J., Rich, M.M., Leferovich, J.M., Reilley, P., David, B.M., Tran, K., Rayburn, H., Brondon, R., Cros, D. *et al.* (1996) Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nature Genet.*, **13**, 325–335.
12. Philips, A.V., Timchencko, L.T. and Cooper, T.A. (1998) Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science*, **280**, 737–741.
13. Timchenko, L.T., Miller, J.W., Timchenko, N.A., DeVore, D.R., Datar, K.V., Lin, L., Roberts, R., Caskey, C.T. and Swanson, M.S. (1996) Identification of a $(CUG)_n$ triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res.*, **24**, 4407–4414.
14. Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science*, **241**, 42–52.
15. Waring, J.D., Haq, R., Tamai, K., Sabourin, L.A., Ikeda, J.-E. and Korneluk, R.G. (1996) Investigation of myotonic dystrophy kinase isoform translocation and membrane association. *J. Biol. Chem.*, **271**, 15187–15193.
16. Ishizaki, T., Maekawa, M., Fujisawa, K., Okawa, K., Iwamatsu, A., Fujita, A., Watanabe, N., Saito, Y., Kakizuka, A., Morii, N. and Narumiya, S. (1996) The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr protein kinase homologous to myotonic dystrophy kinase. *EMBO J.*, **15**, 1885–1893.
17. Groenen, P. and Wieringa, B. (1998) Expanding complexity in myotonic dystrophy. *Bioessays*, **20**, 901–912.
18. Benders, A.G.M., Groenen, P.J.T.A., Oerlemans, F.T.J.J., Veerkamp, J.H. and Wieringa, B. (1997) Myotonic dystrophy protein kinase is involved in the modulation of the Ca^{2+} homeostasis in skeletal muscle cells. *J. Clin. Invest.*, **100**, 1440–1447.
19. Balasubramanyam, A., Iyer, D., Stringer, J.L., Beaulieu, C., Potvin, A., Nuemeyer, A.M., Avruch, J. and Epstein, H.F. (1998) Developmental changes in expression of myotonic dystrophy protein kinase in the rat central nervous system. *J. Comp. Neurol.*, **394**, 309–325.
20. van der Ven, P.F.M., Jansen, G., van Kuppevelt, T.H.M.S.M., Perryman, M.B., Lupa, M., Dunne, P.W., ter Laak, H.J., Jap, P.H.K., Veerkamp, J.H., Epstein, H.F. and Wieringa, B. (1993) Myotonic dystrophy kinase is a component of neuromuscular junctions. *Hum. Mol. Genet.*, **2**, 1889–1894.
21. Whiting, E.J., Waring, J.D., Tamai, K., Somerville, M.J., Hincke, M., Staines, W.A., Ikeda, J.-E. and Korneluk, R.G. (1995) Characterization of myotonic dystrophy kinase (DMK) protein in human and rodent muscle and central nervous tissue. *Hum. Mol. Genet.*, **4**, 1063–1072.
22. Dunne, P.W., Ma, L., Casey, D.L., Harati, Y. and Epstein, H.F. (1996) Localisation of myotonic dystrophy protein kinase in skeletal muscle and its alteration with disease. *Cell Motil. Cytoskel.*, **33**, 52–63.
23. Salvatori, S., Biral, D., Furlan, S. and Marin, O. (1997) Evidence for localization of the myotonic dystrophy protein kinase to the terminal cisternae of the sarcoplasmic reticulum. *J. Muscle Res. Cell Motil.*, **18**, 429–440.
24. Shimokawa, M., Ishiura, S., Kameda, N., Yamamoto, M., Sasagawa, N., Saitoh, N., Sorimachi, H., Ueda, H., Ohno, S., Suzuki, K. and Kobayashi, T. (1997) Novel isoform of myotonin protein kinase. Gene product of myotonic dystrophy is localized in the sarcoplasmic reticulum of skeletal muscle. *Am. J. Pathol.*, **150**, 1285–1295.
25. Ueda, H., Kameda, N., Baba, T., Terada, N., Shimokawa, M., Yamamoto, M., Ishiura, S., Kobayashi, T. and Ohno, S. (1998) Immunolocalization of myotonic dystrophy protein kinase in corbular and junctional sarcoplasmic reticulum of human cardiac muscle. *Histochem. J.*, **30**, 245–251.

26. Fu, Y.-H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, R.A., Pizutti, A., Ashizawa, T., Perryman, M.B., Scarlato, G., Fenwick Jr, R.G. and Caskey, C.T. (1993) Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science*, **260**, 235–238.
27. Mahadevan, M.S., Amemiya, C., Jansen, G., Sabourin, L., Baird, S., Neville, C.E., Wormskamp, N., Segers, B., Batzer, M., Lamerdin, J. *et al.* (1993) Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. *Hum. Mol. Genet.*, **2**, 299–304.
28. Shaw, D.J., McCurrach, M., Rundle, S.A., Harley, H.G., Crow, S.R., Sohn, R., Thirion, J.-P., Hamshere, M.G., Buckler, A.J., Harper, P.S. *et al*. (1993) Genomic organization and transcriptional units at the myotonic dystrophy locus. *Genomics*, **18**, 673–679.
29. Maeda, M., Taft, C.S., Bush, E.W., Holder, E., Bailey, W.M., Neville, H., Perryman, M.B. and Bies, R.D. (1995) Identification, tissue-specific expression and subcellular localisation of the 80- and 71-kDa forms of myotonic dystrophy kinase protein. *J. Biol. Chem.*, **270**, 20246–20249.
30. Pham, Y.C.N., Nguyen thi Man, Le Thanh Lam and Morris, G.E. (1998) Localization of myotonic dystrophy protein kinase in human and rabbit tissues using a new panel of monoclonal antibodies. *Hum. Mol. Genet.*, **7**, 1957–1965.
31. Oshima, Y. and Gotoh, Y. (1987) Signals for the selection of a splice site in pre-mRNA. Computer analysis of splice junction sequences and like sequences. *J. Mol. Biol.*, **195**, 247–259.
32. Gooding, C., Roberts, G., Moreau, G., Nadal-Ginard, B. and Smith, C.W.J. (1994) Smooth muscle-specific switching of α-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon. *EMBO J.*, **13**, 3861–3872.
33. Roberts, G.C., Gooding, C. and Smith, C.W.J. (1996) Smooth muscle alternative splicing induced in fibroblasts by heterologous expression of a regulatory gene. *EMBO J.*, **15**, 6301–6310.
34. Singh, R., Valcárcel, J. and Green, M.R. (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*, **268**, 1173–1176.
35. Gallego, M.E., Gattoni, R., Stévenin, Marie, J. and Expert-Bezançon, A. (1997) The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the β-tropomyosin alternative exon 6A. *EMBO J.*, **16**, 1772–1784.
36. Mount, S.M. (1998) Genetic depletion reveals an essential role for an SR protein splicing factor in vertebrate cells. *Bioessays*, **19**, 189–192.
37. Koike, H., Saitoh, N., Sasagawa, N., Watanabe, T., Shimokawa, M., Sorimachi, H., Arahata, K., Ishiura, S. and Suzuki, K. (1998) Identification and purification of myotonin protein kinase (MtPK) from rat skeletal muscle sarcoplasmic reticulum. *Biomed. Res.*, **19**, 93–99.
38. Timchenko, L., Nastainczyk, W., Schneider, T., Patel, B., Hofmann, F. and Caskey, C.T. (1995) Full-length myotonin protein kinase (72 kDa) displays serine kinase activity. *Proc. Natl Acad. Sci. USA*, **92**, 5366–5370.
39. Matsui, T., Amano, M., Yamamoto, T., Chihara, K., Nakafuku, M., Ito, M., Nakano, T., Okawa, K., Iwamatsu, A. and Kaibuchi, K. (1996) Rho-associated kinase, a novel serine/threonine kinase, as a putative target for the small GTP binding protein Rho. *EMBO J.*, **15**, 2208–2216.
40. Leung, T., Chen, X.Q., Tan, I., Manser, E. and Lim. L. (1998) Myotonic dystrophy kinase-related Cdc42 binding kinase acts as a Cdc42 effector in promoting cytoskeletal reorganization. *Mol. Cell. Biol.*, **18**, 130–140.
41. Leung, T., Chen, X.Q., Manser, E. and Lim, L. (1996) The p160 RhoA-binding kinase ROK (is a member of a kinase family and is involved in the reorganization of the cytoskeleton. *Mol. Cell. Biol.*, **16**, 5313–5327.
42. Green, S., Issemann, I. and Sheer, E. (1988) A versatile *in vivo* and *in vitro* eukaryotic expression vector for protein engineering. *Nucleic Acids Res.*, **16**, 369.
43. Cuppen, E., Wijers, M., Schepens, J., Fransen, J., Wieringa, B. and Hendriks, W. (1999) A FERM domain governs apical confinement of PTP-BL in epithelial cells. *J. Cell Sci.*, **112**, 3299–3308.