Human Molecular Genetics, 1996, Vol. 5, No. 8 **1179–1186**

Identification of a novel sarcoglycan gene at 5q33 encoding a sarcolemmal 35 kDa glycoprotein

Vincenzo Nigro[1,*], Giulio Piluso[1], Angela Belsito[1], Luisa Politano[2], Annibale Alessandro Puca[3], Serenella Papparella[4], Elena Rossi[5], Giuseppe Viglietto[6], Maria Grazia Esposito[2], Ciro Abbondanza[1], Nicola Medici[1], Anna Maria Molinari[1], Giovanni Nigro[2] and Giovanni Alfredo Puca[1]

[1]Istituto di Patologia Generale ed Oncologia and [2]Dipartimento di Internistica Clinica e Sperimentale, Facoltà di Medicina, Seconda Università degli Studi di Napoli, 80138 Napoli, Italy, [3]Telethon Institute of Genetics and Medicine (TIGEM), San Raffaele Biomedical Science Park, Milano, Italy, [4]Istituto di Anatomia Patologica Veterinaria, Università degli Studi Federico II, Napoli, Italy, [5]Cattedra di Biologia Generale, Università degli Studi di Pavia and Servizio di Citogenetica, San Raffaele Biomedical Science Park, Milano, Italy and [6]Istituto dei Tumori Fondazione Sen. Pascale, Napoli, Italy

Received May 10, 1996; Revised and Accepted June 6, 1996

Mutations in any of the genes encoding the α, β or γ-sarcoglycan components of dystrophin-associated glycoproteins result in both sporadic and familial cases of either limb-girdle muscular dystrophy or severe childhood autosomal recessive muscular dystrophy. The collective name 'sarcoglycanopathies' has been proposed for these forms. We report the identification of a fourth member of the human sarcoglycan family. We named this novel cDNA δ-sarcoglycan. Its mRNA expression is abundant in striated and smooth muscles, with a main 8 kb transcript, encoding a predicted basic transmembrane glycoprotein of 290 amino acids. Antibodies specifically raised against this protein recognized a single band at 35 kDa on western blots of human and mouse muscle. Immunohistochemical staining revealed a unique sarcolemmal localization. FISH, radiation hybrid and YAC mapping concordantly linked the δ-sarcoglycan gene to 5q33, close to *D5S487* and *D5S1439*. The gene spans at least 100 kb and is composed of eight exons. The identification of a novel sarcoglycan component modifies the current model of the dystrophin–glycoprotein complex.

INTRODUCTION

In the last few years, the role of dystrophin, the product of the Duchenne muscular dystrophy gene at Xp21 (1,2 and references therein), has been thoroughly investigated and a group of previously unknown associated proteins was identified in skeletal muscle. Dystrophin was copurified with some integral membrane and membrane-associated glycoproteins collectively defined the dystrophin-glycoprotein complex (3–6). Dystrophin, like a large rod, is supposed to bind F-actin with its actinin-like N-terminus (7), while being anchored with its unique C-terminus at a sarcolemmal 43 kDa transmembrane glycoprotein (β-dystroglycan) (8). The β-dystroglycan binds in turn to an extracellular 156 kDa glycoprotein (α-dystroglycan). Both the α- and β-dystroglycan proteins are encoded by a single gene on chromosome 3p21, their differences resulting from a distinct proteolytic processing of a single precursor protein (8,9). The α-dystroglycan is bound to merosin (laminin-2), a component of the extracellular matrix and, in addition, is the receptor for agrin, a protein with a pivotal role in the clustering of acetylcholine receptors at the neuromuscular junction (10–12). This subcomplex, also defined as the dystrophin-dystroglycan complex, is represented as a bridge across the membrane between the cytoskeleton and the basal lamina, needed to confer stability to the sarcolemma and protecting muscle cells from long-term contraction-induced damage and necrosis. A second associated subcomplex is sarcoglycan, which is composed by at least three transmembrane glycoproteins, α-sarcoglycan (50DAG, A2, adhalin), β-sarcoglycan (43DAG, A3b) and γ-sarcoglycan (35DAG, A4). The role of sarcoglycan and its relationship to the dystroglycan complex remain obscure (13). A triplet of 59 kDa proteins, known as syntrophins, a 25DAP (A5) and the human homologues of the *Torpedo* 87 kDa (dystrobrevins) also seem to be involved in complexes with dystrophin (14–16).

Mutations in the genes encoding dystrophin-associated proteins may be predicted to be responsible for the clinically and genetically heterogeneous autosomal muscular dystrophies. In the last two years, three human members of the sarcoglycan complex were cloned and mutations in these genes were found to be associated to autosomal recessive muscular dystrophies. Both missense and nonsense mutations of the α-sarcoglycan gene at 17q21 (18,19), β-sarcoglycan at 4q12 (20,21) or γ-sarcoglycan at 13q12 (22) resulted in the disruption of the sarcoglycan complex

*To whom correspondence should be addressed

and were associated to phenotypes ranging from severe Duchenne-like muscular dystrophy, referred to as severe childhood autosomal recessive muscular dystrophy (SCARMD), to later onset and milder limb-girdle muscular dystrophy (LGMD). The latter is a progressive muscular dystrophy with weakness beginning in the hip and shoulder girdle and later progressing to the distal muscles. LGMD is genetically heterogeneous with dominant (*LGMD1*) and recessive (*LGMD2*) forms (23,24). Among the five recessive forms described, *LGMD2C*, *LGMD2D* and *LGMD2E* were conclusively co-mapped with the sarcoglycan genes (18–22). *LGMD2A* at 15q15 was associated with mutations of the calpain-3 gene (25), while the *LGMD2B* locus at 2p13 has been refined within the region covered by a YAC (26,27). The dominant *LGMD1A* was linked to 5q (28) with regional localization to a 7 cM interval spanned by *D5S178* and *IL9* at 5q31-q33 (29). Some LGMD1 families have been excluded from these intervals, indicating the presence of additional loci (30).

Isoelectric focusing analysis of the smaller sarcoglycan component has suggested that the 35 kDa DAG may be heterogeneous with the presence of a more basic protein component having the same molecular weight (31). Here, we report the identification, through expressed sequence tag (EST) database searching and cDNA library screening, of a novel human gene, belonging to the sarcoglycan family.

RESULTS

Isolation and sequence analysis of δ-sarcoglycan cDNAs

The sequence of the human and rabbit γ-sarcoglycan (U34976 and U36822) were compared using the TBLASTn algorithm with other sequences in the electronic databases, including the subset of expressed sequence tags (EST). No significant homologies were found, but the amino acids 27–47 of the γ-sarcoglycan were weakly similar to the translation of one open reading frame (ORF) of the EST with accession no. Z25253 (Fig. 1a). This sequence (clone A9H04l) was submitted since 1993 (two years before the cloning of the γ-sarcoglycan) by Genexpress and derived from the leg skeletal muscle. The short sequence was preceded and followed by a region without similarity and with stop codons in all reading frames. To ascertain whether this EST contained a rearranged fragment of a putative second γ-sarcoglycan gene, or this weak similarity was merely casual, we searched for this cDNA in a human skeletal muscle library. We synthesised two non overlapping oligonucleotides corresponding to the short similarity region and screened with each of them separately the duplicate filters of the library. Nine clones out of 500 000 were positive with both probes. They were purified and sequenced. Six of them contained the sequence of the probes inserted, however, in the context of a novel sequence. This sequence (Fig. 1b) encodes a single ORF of 870 nt with 58% nucleotide identity to both the human and rabbit γ-sarcoglycan sequences (GenBank/EMBL accession number X95191). The sequence flanking the first AUG conforms to the Kozak consensus for translational start sites. In the 5′ untranslated region (UTR), the cDNA has two in-frame stop codons. It encodes a protein of 290 amino acids with a predicted molecular mass of 32 174 daltons (the human γ-sarcoglycan weighs 32 352 daltons). The similarity is higher at the protein level (70%) (Fig. 2). This protein, here named δ-sarcoglycan, is a putative transmembrane glycoprotein with a small intracellular domain, a single transmembrane hydrophobic domain (amino acids 35 to 59) and a larger extracellular carboxyl terminus of 231 amino acids. This structure is identical to that seen for β-sarcoglycan (43DAG) and γ-sarcoglycan (35DAG) (20–22). The extracellular domain of δ-sarcoglycan contains four cysteine residues that are a common feature to all sarcoglycan family members (Fig. 2). These residues are believed to participate in intra- and intermolecular disulfide bond formation. Interestingly, the recessive mutation found in Ben Hamida 13q SCARMD causes the loss of the cysteine cluster, while preserving the proximal regions of the γ-sarcoglycan (22). A potential asparagine-linked glycosylation site is present at the same position of human and rabbit γ-sarcoglycan. The calculated isoelectric point is 9.15, a value identical to that of β-sarcoglycan (pI 9.2) and different from that of human γ-sarcoglycan (pI 5.6).

We then used this novel sequence as a hybridization probe for further cDNA library screening. Three additional cDNA clones were isolated from human foetal brain. The muscular sequence was confirmed without evidence of alternative splicings. To complete the cloning of the 5′ UTR, we performed 5′ rapid amplification of cDNA ends (RACE) using reverse nested oligonucleotide primers located at the 5′ end of the cDNA (32). Amplifications were carried out on muscle, heart and foetal brain cDNA templates. Several products were cloned and sequenced, until the longest extension in the 5′ UTR was determined.

Tissue distribution of δ-sarcoglycan expression

We employed Northern blot analyses to assay total RNA from various human adult tissues hybridized to δ-sarcoglycan cDNA probe, encompassing the coding region and part of the 3′UTR. We detected a major hybridizing transcript of approximately 8 kb and a less abundant form of about 3.6 kb. The strongest signal was observed in skeletal muscle and heart and a slightly weaker signal was present in smooth muscle (Fig. 3). Interestingly, much fainter bands of the same sizes were also visible in brain- and lung-derived RNA after longer exposures of the autoradiograms (not shown). The 8 kb mRNA is exceedingly larger than the coding region and the longest cDNA clone we obtained (3 kb) did not contain the complete 3′UTR. A very long 3′UTR with a minor alternative polyadenylation signal is predictable. Under our conditions, no cross hybridization signal from the 1.7 kb γ-sarcoglycan mRNA was observed.

Multiplex semiquantitative RT-PCR from random primed-cDNA from normal adult heart and skeletal muscle RNA was performed. We combined in the same assay primer pairs for δ-sarcoglycan together with primer pairs that amplify fragments of α-sarcoglycan, β-sarcoglycan, or γ-sarcoglycan. Using different combinations of primers we deduced that the quantitative expression of the δ-sarcoglycan may be comparable to that of the other sarcoglycans in skeletal muscle and slightly higher than the expression of α-sarcoglycan in heart (not shown).

The striated muscle-specific mRNA expression of the δ-sarcoglycan gene is similar to that of α- and γ-sarcoglycan, with a peculiar expression also in smooth muscle. β-sarcoglycan, which is widely expressed in nonmuscle tissues, diverges from all the others.

Genomic structure and chromosomal mapping

Genomic libraries screening, together with long-range (33) and vectorette PCR (34) were used to determine exon-intron boundaries of the gene. The δ-sarcoglycan is composed by eight exons. They are separated by introns, that are very long, with the

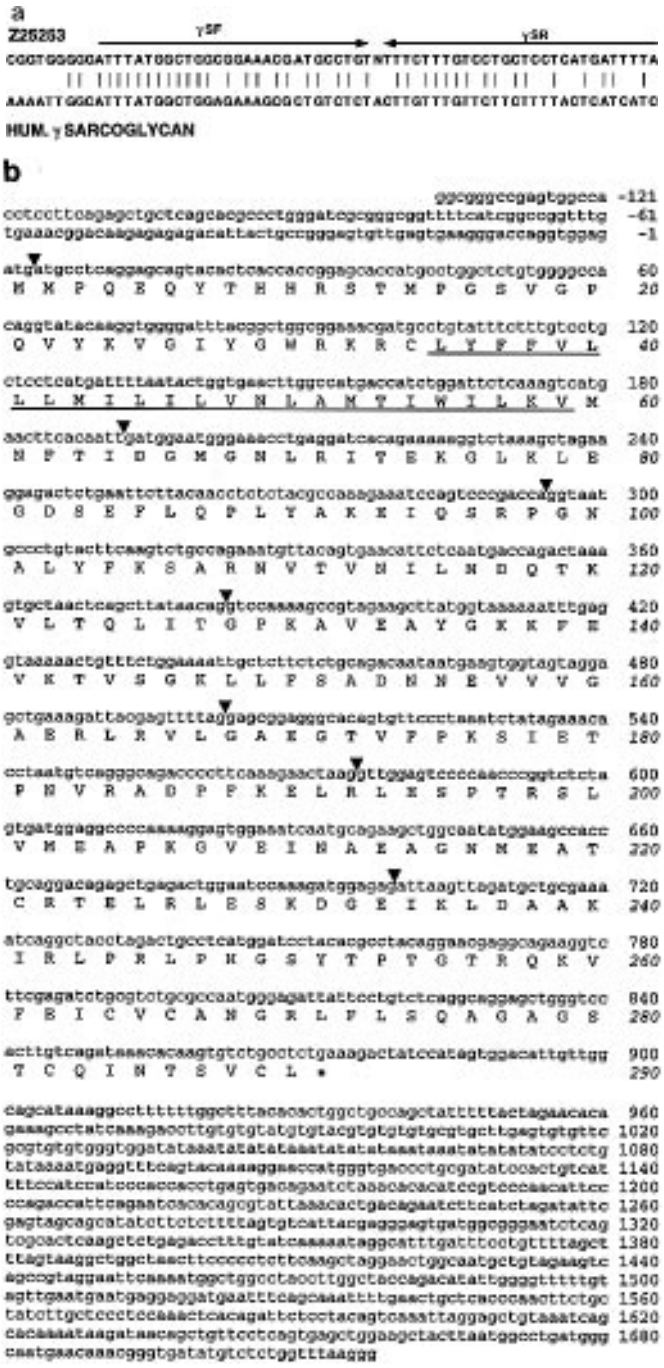

Figure 1. (**a**) Alignment of the EST Z25253 and human γ-sarcoglycan nucleotide sequences. Oligonucleotides γF and γR were used in the library screening. (**b**) Nucleotide and predicted amino acid sequence of the δ-sarcoglycan cDNA. Nucleotides and amino acids are numbered starting at the first methionine. The STOP codon is marked by an asterisk. Arrowheads indicate splice junctions. The putative transmembrane region is underlined. Two polymorphisms were found: C84T (Tyr28Tyr) (frequency: 0.20) and G290A (Arg97Gln) (frequency: 0.06). Accession number is X95191.

exception of the first (3 kb) and the last one (1.6 kb). The first exon contains only the first methionine and the termination signal is followed by a long 3′ UTR colinear with the genomic sequence. We calculated that the entire gene spans at least 100 kb.

To localize the δ-sarcoglycan gene we used the longest cDNA as a probe, after removal of a small repeat in the 3′UTR, for fluorescence *in situ* hybridization (FISH). A specific labelling on the long arm of chromosome 5, corresponding to band 5q33, was observed (not shown).

Since our data suggested that the genomic extension of the gene is very large, a duplex PCR was set up with specific primers located in the second exon together with primers located in the last exon. Under our stringent PCR conditions these primers uniquely amplified the human δ-sarcoglycan gene and not the γ-sarcoglycan or related rodent sequences. This duplex PCR was used on template DNAs isolated from a panel of 93 human-rodent radiation hybrids (Genebridge 4). A slightly different location between the second and the eighth exon of the gene was observed, thus confirming the large size of the gene. The more 3′ sequence was linked to *D5S820*, while the 5′ one was 3 cR more centromeric (*D5S487*).

A third independent method of mapping was employed. The duplex PCR assay was used to identify yeast artificial chromosomes (YAC) containing the gene. Among several primary pools that resulted positive, we isolated two YACs containing only the second exon, two YACs containing only the last exon and two CEPH YACs containing both sequences: YAC889c8 and 936d12. These YACs, belonging to the contig WC-393, contain several mapped markers (*D5S487, D5S662, D5S1439, D5S2112*). This confirmed both FISH and radiation hybrid results.

Localization of δ-sarcoglycan protein

A glutathione S-transferase (GST)-δ-sarcoglycan fusion protein was used to raise polyclonal antibodies in rabbit and monoclonal antibodies in mice. The rabbit immune serum was found to react on western blots with the bacterial GST-δ-sarcoglycan fusion protein, but neither with the GST alone, nor with a fusion protein containing the homologous region of γ-sarcoglycan. Fresh homogenates of mouse leg and human pectoral muscles were analyzed by western blotting and a single band of about 35–37 kDa was observed in both cases (Fig. 4). The immunoreactivity was totally competed out with the fusion protein used to raise the serum, but not with the control proteins. A nearly identical 35 kDa band was observed in COS-7 cell extracts, transiently transfected with an expression vector containing the full-length δ-sarcoglycan under the CMV promoter (Fig. 4).

Immunohistochemical staining of frozen sections from normal human biopsies showed that the protein is uniquely localized at the sarcolemma, according to the prediction based upon the amino acid sequence. No immunostaining of intracellular or interstitial components was observed (Fig. 5). The signal was specifically competed out by the fusion protein and identical results were obtained using an affinity-purified antibody. In sections from young DMD patients (3–5 yrs), no evident reduction of immunolabeling was detected. However, using sections from older DMD boys δ-sarcoglycan was found reduced. This reduction was confirmed by the immunoblot analysis (not shown).

```
δ-sarc.   VYKVGIYGWR KRCLYFFVLL LMILILVNLA MTIWILKVNN FTIDGMGNLR ITKEGLKLEG  81
γ-sarc.   ---I------ -----L---- -L-ILV---- L--------W -SPA---H-C V-KD--R---  82
β-sarc.   LS-T-LR-RK GNLAICVII- -F--AVI--I I-LV-WA-IR IGPN-CDSNE FH-S--LRFK  111
F07H5.2   I-R------- --F--T-I-I -A-G-VL--S L-F--MS-LD -SP--I-T-K -EDD-IRV--  129

δ-sarc.   DSEFLQPLYA KEIQSRPGNA LYFKSARNVT VNILNDQTKV LTQLITGPSA VEAYGKKFEV  141
γ-sarc.   E----F---- ---H--VDSS -LLQ-TQ--- --AR-SEGE- TGR-KV---N --VQNQQ-QI  142

          KTVSGKLLFS ADSGEVVVGA ERLRVLGAEG TVFFKSIETP NVRADPFKEL RLESPTRSLV  201
          NSND--P--T V-EK-----T DK---T-P-- AL-ES-V--- L------QD- ---------S  202

          NEAPKGVEIN AEAGNNEATC RTELRLESKD GEIKLDAAKI RLPRLPHGSY TPTGTRQKVF  261
          -D--R--H-Q -H--KI--LS QMDILFH-S- -MLV---ETV C--K-VQ-TW G-S-SS-SLY  262

          EICVCANGRL FLSQAGAGST CQINISVCL  290
          -----PD-K- Y--V--VST- --ENSHI--  291
```

Figure 2. Alignment of human δ-sarcoglycan and γ-sarcoglycan amino acid sequences. Numbers refer to the amino acid sequence starting at the first methionine. The amino acids 22–81 of human δ-sarcoglycan are also aligned to the human β-sarcoglycan and the *C. elegans* clone F07H5.2. Dashes indicate amino acid identities with the human δ-sarcoglycan.

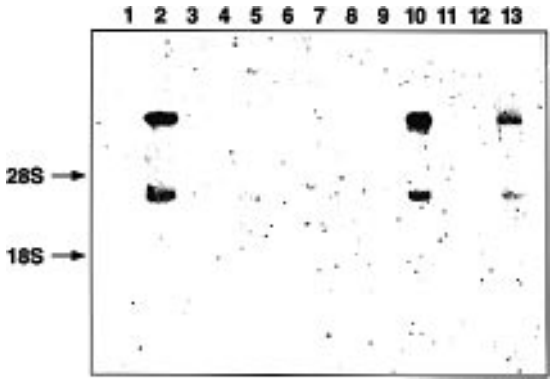

Figure 3. Northern blot analysis. The tissues represented in each lane are: 1 spleen; 2, heart; 3, stomach; 4, lung; 5, brain; 6, small intestine; 7, liver; 8, mammary gland; 9, pancreas; 10, skeletal muscle; 11, placenta; 12, kidney; 13, smooth muscle.

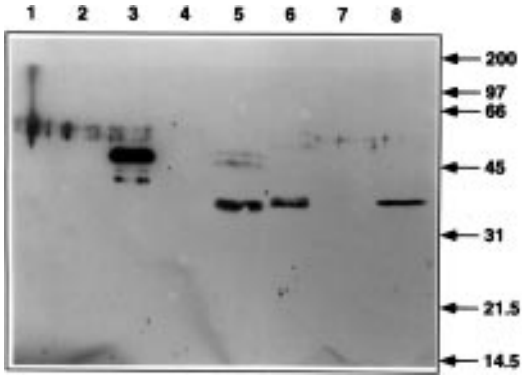

Figure 4. Western blot analysis. 1, GST alone; 2, GST-γ-sarcoglycan; 3, GST-δ-sarcoglycan; 4, bacterial extracts; 5, mouse skeletal muscle; 6, human skeletal muscle; 7, transfected antisense δ-sarcoglycan; 8, transfected δ-sarcoglycan.

DISCUSSION

In a widely accepted model, the dystrophin-glycoprotein complex is composed by two subcomplexes: the dystrophin-dystroglycan and the sarcoglycan (13). While the former serves as a bridge across the sarcolemma to the merosin, the latter is a joined complex of transmembrane glycoproteins with unclear function. The importance of sarcoglycan is underlined by the findings of mutations in any of the three sarcoglycan genes in patients with autosomal muscular dystrophies (18–22).

We have cloned and characterized a fourth sarcoglycan component of dystrophin-associated glycoproteins. The predicted protein product is similar in sequence, size and sarcolemmal localization to that of the recently cloned γ-sarcoglycan gene (22). Two differences are evident: a basic isoelectric point of 9.15 compared to that of 5.6 of the γ-sarcoglycan and the smooth muscle mRNA expression. Thus, there are at least two distinct, but homologous 35 kDa dystrophin-associated glycoproteins. We decided to name this novel gene δ-sarcoglycan.

The protein homology between γ-sarcoglycan and δ-sarcoglycan is very high (55% of amino acid identity, 70% similarity) (22). This is comparable with the homology among the three human syntrophins compared to each other (50 to 57% identity) (15). Two regions were also similar to the β-sarcoglycan sequence (20,21). Interestingly, the first 99 amino acids of δ-sarcoglycan were 66% similar to the *C. elegans* clone F07H5.2 (Fig. 2). Although this sequence was highly represented in the cDNA libraries, no useful EST has been found in the electronic databases. This may be due to the length of the 3′UTR which precludes the usage of poly(A^+)-tailed libraries. Such limitation could have affected the cDNA expression libraries used to clone the other 35DAG (22).

In view of the fact that there are two different genes encoding two overlapping 35DAGs, it is now possible to speculate why the disappearance of the 35DAG observed in patients with primary defects in its 13q12 gene was not so evident as the secondary loss of the 50DAG (35,36).

When the Campbell and Ozawa groups independently found that dystrophin was associated with several other proteins, they recognized by one-dimensional SDS-PAGE a 35–36 kDa band,

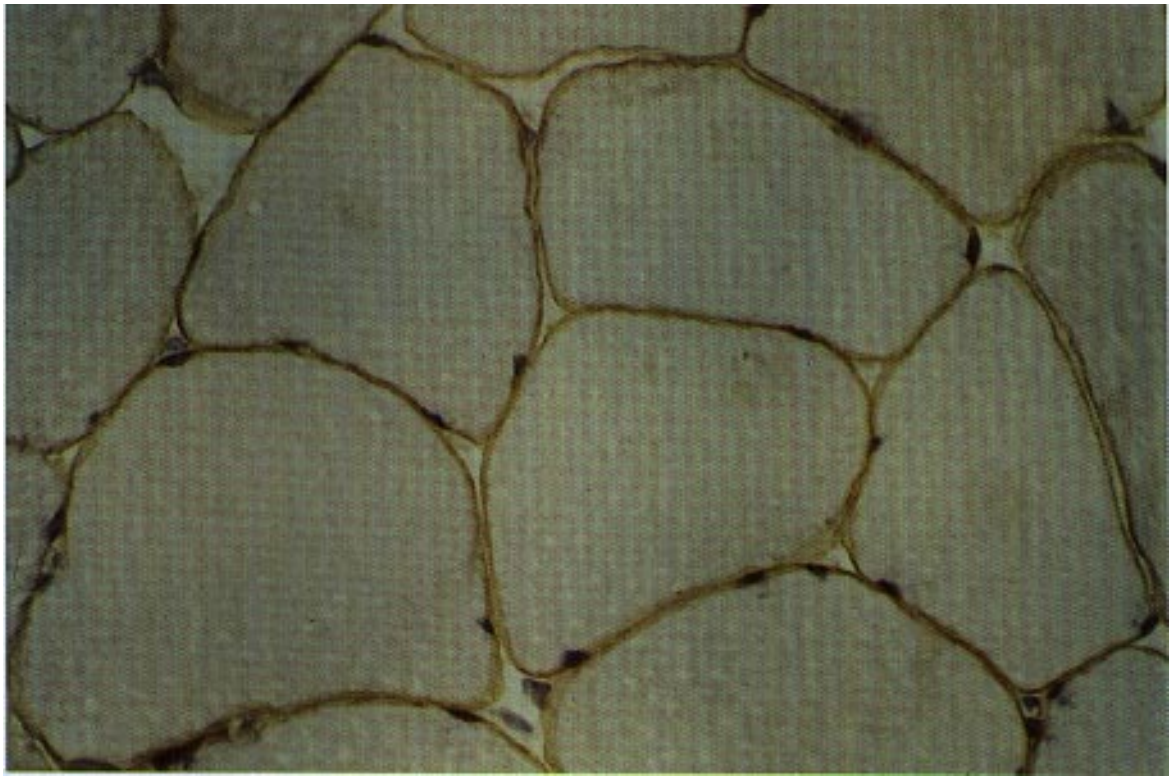

Figure 5. Immunohistochemical analysis of normal muscle.

originally named A4 or 35DAG (2–6). However, two-dimensional PAGE analysis of rabbit dystrophin-associated proteins (DAP) provided more insight about the complex (31). The 35DAG exhibited heterogeneity on isoelectric focusing, suggesting that there may be another component at this molecular weight. While many groups of protein spots have been recognized on the basis of their characteristics and/or by specific antibodies, some groups matched no known proteins. In particular, the basic spot series 10 (43 kDa) and 12 (35 kDa) remained unidentified. It has subsequently been demonstrated that the basic spot series 10 corresponds to the 43 kDa β-sarcoglycan (13). On the basis of our data, we suggest that the δ-sarcoglycan corresponds to spot series 12 detected in the dystrophin-DAP complex (31). Interestingly, accurate crosslinking experiments demonstrated in the sarcoglycan subcomplex two major crosslinked products of estimated molecular mass of about 110 kDa and 69 kDa, devoid of the 50 kDa α-sarcoglycan (13). It is attractive to speculate that δ-sarcoglycan could be linked together with γ-sarcoglycan in the 69 kDa product and with β- and γ-sarcoglycan in the 110 kDa product.

The finding that mutations affecting sarcoglycan genes can cause destabilization of the entire subcomplex, as observed in *LGMD* or *SCARMD* (18–22), suggests that the δ-sarcoglycan gene is another candidate for primary genetic cause of muscular dystrophy.

The mapping of human δ-sarcoglycan gene to chromosome 5q33 seemed, at first glance, particularly exciting, as the autosomal dominant *LGMD1A* was linked by Speer *et al.* to 5q31-33 (28,29). The large *LGMD1A* family, which was used in the mapping study, was characterized by elevated serum creatine kinase (CK), proximal leg weakness and a very slow progression. However, the *LGMD* gene was sublocalized within the interval between *D5S178* and *IL9*, a few cM centromeric to the assigned location of the δ-sarcoglycan gene. In addition, to our knowledge all the mutations so far detected in the other sarcoglycan genes were recessive. This seems reasonable, considering the structural role of their protein products. However, in principle one cannot exclude the existence of special dominant mutations, or manifesting heterozygous individuals.

We set up multiple SSCP protocols (37,38) to screen for mutations the entire coding region of δ-sarcoglycan. Preliminary data involved 140 available patients with autosomal muscular dystrophies (85 isolated cases, 55 familial cases with either dominant or recessive *LGMD*). We identified a candidate dominant missense mutation in one patient with 3-fold increase CK and mild proximal myopathy. This mutation (Asn211Tyr) was not found in 500 control chromosomes, or in the normal family members. Further studies are in progress to investigate the possible causative role of this mutation.

In spite of the heterogeneity on isoelectric focusing of the 35DAG, the cloning of another sarcoglycan gene had never been predicted. The homology between γ- and δ-sarcoglycan and their colocalization raise the issue of functional redundancies for these proteins. However, mutations in the γ-sarcoglycan may lead to lethal myopathies (22) and, therefore, it is clear that δ-sarcoglycan cannot compensate for γ-sarcoglycan loss of function. It is likely that a functional and/or structural specialization of each gene product precludes a vicarial role. In conclusion, the identification of a novel sarcoglycan component modifies the current model of the dystrophin-glycoprotein complex (Fig. 6). The possible consequences of either spontaneous or experimentally induced (i.e. gene knock-out) mutations in the δ-sarcoglycan gene will represent the subject of future investigations.

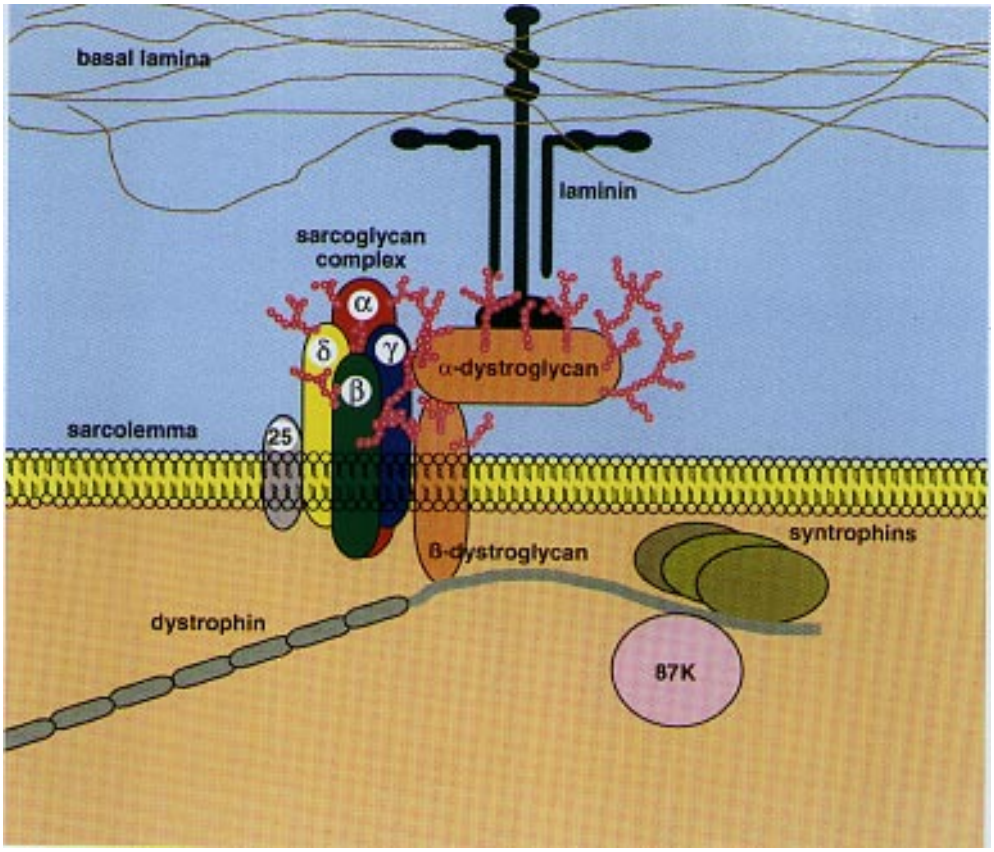

Figure 6. Schematic representation of the dystrophin-associated protein complex at the sarcolemma. This hypothetical model incorporates theδ-sarcoglycan (yellow) in the joined sarcoglycan subcomplex. The cytoplasmic domain of the sarcoglycan components is in-scale according to the predicted sequences, but the relative contacts are uncertain. α- and β-dystroglycan are of the same colour, because they are encoded by the same gene. The N-terminus of dystrophin (grey-green) is anchored to F-actin (not shown). The position of the 25DAP cannot be predicted. The small circles (pink) indicate sugar moieties.

MATERIALS AND METHODS

Isolation of human δ-sarcoglycan cDNAs

Screening of the phage libraries was performed according to standard procedures (39). We screened approximately 500 000 plaques of a human skeletal muscle random-primed cDNA library in λgt10 (Clontech). One set of filters was hybridized in 6× SSC at 37°C with the ^{32}P-labeled oligonucleotide γF (10^8 c.p.m.), while the duplicates were separately hybridized with the oligo γR. Washes were carried out at 45°C in 1× SSC. Nine clones, positive with both probes, were subcloned and manual sequence analysis was performed from both strands using the T7 DNA polymerase (Pharmacia). Each sequence was confirmed at least by two different phage clones. The other libraries were screened with an insert fragment of the cDNA cloned from the muscle library.

RACE

5′RACE was performed from 1 µg of total RNA from human tissues. Reverse transcriptase (BRL) was from the semi-random oligo 5′-NNNNNNSS-3′. After treatment with terminal transferase (BRL) in the presence of dATP, nested amplifications were performed starting from specific reverse primers located at positions 290–267, 159–126, 90–67 or 24–0 of the cDNA sequence together with standard RACE primers.

Northern blot

Human total RNA was isolated according to standard procedures (39) or was purchased from Clontech. RNA was separated by electrophoresis through a 1% agarose, 4% formaldehyde gel and capillary transferred to a nylon membrane (Schleicher und Schuell). Membranes were then hybridized with the 1.2 kb *Eco*RI fragment of δ-sarcoglycan in 40% formamide at 42°C.

PCR conditions

All PCR amplifications were performed in a thermal cycler PTC-100 (MJ Research) using a 10 µl reaction with 10–200 ng of template DNA, 0.5 µM of each primer, 3.5 mM dNTPs 1× multibuffer JD (composition available on request) and 0.5 U AmpliTaq polymerase (Perkin Elmer). Initial denaturation at 99°C (30 s) was followed by 30 cycles with denaturation at 96°C (15 s) and annealing-extension at 61°C for 3 min. Oligonucleotides corresponding to human δ-sarcoglycan cDNA 2–24 (exon 2, sense), 159–136 (exon 2, antisense), 740–764 (exon 8, sense) and 1157–1134 (exon 8, antisense) were used in the duplex assay.

FISH

FISH was performed on 46,XY chromosome preparations obtained from human blood cultures after 5-bromodeoxyuridine synchronization and thymidine release, with a 1 h colchicine block. The cDNA, labelled with biotin-16-dUTP by nick translation (Boehringer Mannheim), was hybridized to chromosome preparations previously denatured at 70°C in 70% formamide/2× SSC. Hybridization mixture (30 µl/slide) contained the biotinylated probe (200 ng/slide) plus salmon sperm. Detection was performed by incubation of the slides with fluorescein isothiocianate-avidin with two amplification steps. The slides screening was performed by taking into consideration only those chromosomes with signals present on both chromatids at the same band position. Two hundred metaphases were examined and in 42 chromosomes five signals were present at band 5q33. No other cluster of signals was present.

Induction and purification of fusion proteins

δ-sarcoglycan cDNA (amino acids 84–290) and γ-sarcoglycan cDNA (amino acids 60–291) were cloned in-frame with GST into pGEX-2TK and introduced in *E. coli* JM109 cells. Fifty ml overnight cultures were diluted 1:10 and induced with 0.1 mM IPTG to promote GST-fusion proteins [GST-δSarc(84–290) and GST-γSarc(60–291)] production. GST-fusion proteins were purified from preparative SDS-PAGE by cutting and electroelution of appropriate protein bands, followed by 20% TCA precipitation. Protein pellets were resuspended in PBS with 2% SDS and 100 mM 2-mercaptoethanol. Protein concentration was determined (Bio-Rad protein assay).

Antibody production

Anti δ-sarcoglycan antibodies were generated by subcutaneous injection of New Zealand white rabbits with 100 µg of purified GST-δ-sarcoglycan (amino acids 84–290) protein emulsified in Freund's complete adjuvant. Rabbits were boosted at two and three weeks with 100 µg of purified protein emulsified in Freund's incomplete adjuvant and bled the following week. The anti-serum titer (1:100 000) was determined by ELISA assay and specificity verified by western blotting. Any possible cross-reactivity was eliminated after overnight incubation at 4°C of anti-serum with purified GST-γSarc(60–291) protein (5 mg/ml) and *E. coli* total proteins (10 mg/ml).

Western blotting

Mouse and human frozen muscular tissue was crushed in a precooled mortar. Approximately 50 mg of sample were added to 200 µl of loading buffer (8 M Urea, 125 mM Tris–HCl pH 6.8, 4% SDS, 100 mM DTT and 0.001% Bromophenol blue), boiled for 3 min and centrifuged to remove SDS-insoluble proteins. Samples were run on 11% SDS-PAGE (40) and transferred to nitrocellulose sheets (41). Membranes were incubated 2 h at room temperature with rabbit antiserum diluted 1:4000 in PBS with 3% nonfat milk, 0.05% Tween20 and 0.05% NP40. After washing, membranes were incubated 1 h with peroxidase conjugated anti-rabbit IgG diluted 1:10 000 in PBS with 0.5% nonfat milk, 0.05% TWEEN20 and 0.05% NP40. Immunoreactive bands were visualized by ECL according to specifications of the manufacturer (Amersham).

Immunohistochemistry

Skeletal muscle biopsy specimens were from normal controls and DMD patients. Immunostaining was performed with the streptavidin-biotin peroxidase method (LSAB Kit- Dako, Denmark). Endogenous peroxidase was blocked with hydrogen peroxide 0.3% in absolute methanol for 30 min. After rinsing with PBS, sections were incubated for 30 min with normal serum. The primary antibody was diluted 1:2000–1:4000 in PBS containing 0.5% BSA. Biotinylated anti-mouse and anti-rabbit immunoglobulins (LSAB Dako kit, Dako) were used as secondary antibody. After washing in PBS, the sections were incubated in streptavidin conjugated to horseradish peroxidase. A monoclonal Ab anti-spectrin (Ylem) was used as a parallel control for muscle membrane integrity.

ACKNOWLEDGMENTS

The authors are indebted to Andrea Ballabio for many valuable comments on the manuscript. We wish to thank Cinzia Sala and Sandro Banfi for advice in mapping experiments. G. Pi. was supported by UILDM, Italy. The financial support of Telethon-Italy is gratefully acknowledged.

ABBREVIATIONS

cM, centimorgan; DAG, dystrophin-associated glycoprotein; EST, expressed sequence tag; FISH, fluorescence *in situ* hybridization; kb, kilobase pairs; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; UTR, untranslated region; YAC, yeast artificial chromosome.

REFERENCES

1. Hoffman, E.P., Brown, R.H.Jr., and Kunkel, L.M. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**, 919–928.
2. Emery,A.E.H. (ed.) (1993) *Duchenne Muscular Dystrophy.* Oxford: Oxford Medical Publications.
3. Campbell,K.P. and Kahl,S.D. (1989) Association of dystrophin and an integral membrane glycoprotein. *Nature* **338**, 259–262.
4. Yoshida,M. and Ozawa,E. (1990) Glycoprotein complex anchoring dystrophin to sarcolemma. *J. Biochem.* **108**, 748–752.
5. Ervasti,J.M., Ohlendieck,K., Kahl,S.D., Gaver,M.G. and Campbell,K.P. (1990) Deficiency of a glycoprotein component of the dystrophin complex in dystrophic muscle. *Nature* **345**, 315–319.
6. Ervasti,J.M., and Campbell,K.P. (1991) Membrane organization of the dystrophin-glycoprotein complex. *Cell* **66**, 1121–1131.
7. Koenig,M., Monaco,A.P. and Kunkel,L.M. (1988) The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. *Cell* **53**, 219–228.
8. Ibraghimov-Beskrovnaya,O., Ervasti,J.M., Leveille,C.J., Slaughter,C.A., Sernett,S.W. and Campbell,K.P. (1992) Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **355**, 696–702.
9. Ibraghimov-Beskrovnaya,O., Milatovich,A., Ozcelik,T., Yang,B., Koepnick,K., Francke,U. and Campbell,K.P. (1993) Human dystroglycan: skeletal muscle cDNA, genomic structure, origin of tissue specific isoforms and chromosomal localization. *Hum. Mol. Genet.* **2**, 1651–1657.
10. Campanelli,J.T., Roberds,S.L., Campbell,K.P., and Scheller,R.H. (1994) A role for dystrophin-associated glycoprotein and utrophin in agrin-induced AChR clustering. *Cell* **77**, 663–674.
11. Gee,S.H., Montanaro,F., Lindenbaum,M.H. and Carbonetto,S. (1994)α-dystroglycan, a dystrophin-associated glycoprotein, is a functional agrin receptor. *Cell* **77**, 675–686.
12. Hopf,C. and Hoch,W. (1996) Agrin binding to α-dystroglycan. *J. Biol. Chem.* **271**, 5231–5236.

13. Yoshida,M., Suzuki,A., Yamamoto,H., Noguchi,S., Mizuno,Y. and Ozawa,E. (1994) Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl β-D-glucoside. *Eur. J. Biochem.* **222**,1055–1061.
14. Adams,M.E., Butler,M.H., Dwyer,T.M., Peters,M.F., Murnane,A.A. and Froehner,S.C. (1993) Two forms of mouse syntrophin, a 58 kd dystrophin-associated protein, differ in primary structure and tissue distribution.*Neuron* **11**, 531–540.
15. Ahn,A.H., Freener,C.A., Gussoni,E., Yoshida,M., Ozawa,E. and Kunkel,L.M. (1996) The three human syntrophin genes are expressed in diverse tissues, have distinct chromosomal locations, and each bind to dystrophin and its relatives. *J. Biol. Chem.* **271**, 2724–2730.
16. Sadoulet-Puccio,H.M., Khurana,T.S., Cohen,J.B.and Kunkel,L.M. (1996) Cloning and characterization of the human homologue of a dystrophin related phosphoprotein found at the *Torpedo* electric organ post-synaptic membrane. *Hum. Mol. Genet.* **5**, 489–496.
17. McNally,E.M., Yoshida,M., Mizuno,Y., Ozawa,E. and Kunkel,L.M. (1994) Human adhalin is alternatively spliced and the gene is located on chromosome 17q21. *Proc. Natl. Acad. Sci. USA* **91**, 9690–9694.
18. Roberds,S.L., Leturcq,F., Allamand,V., Piccolo,F., Jeanpierre,M., Anderson,R.D., Lim,L.E., Lee,J.C., Tomé,F.M.S., Romero,N.B., Fardeau,M., Beckmann,J.S., Kaplan,J.-C. and Campbell,K.P. (1994) Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy. *Cell* **78**, 625–633.
19. Piccolo,F., Roberds,S.L., Jeanpierre,M., Leturcq,F., Azibi,K., Beldjord,C., Carrié,A., Récan,D., *et al.* (1995) Primary adhalinopathy: a common cause of autosomal recessive muscular dystrophy of variable severity. *Nature Genet.* **10**, 243–245.
20. Lim L.E., Duclos,F., Broux,O., Bourg,N., Sunada,Y., Allamand,V., Meyer,J., Richard,I., Moomaw,C., *et al.* (1995) β-sarcoglycan: characterization and role in limb-girdle muscular dystrophy linked to 4q12. *Nature Genet.* **11**, 257–265.
21. Bönnermann,C.G., Modi,R., Noguchi,S., Mizuno,Y., Yoshida,M., Gussoni,E., McNally,E., Duggan,D.J., Angelini,C., Hoffman,E.P., Ozawa,E., and Kunkel,L.M. (1995) β-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex. *Nature Genet.* **11**, 266–272.
22. Noguchi,S., Mc Nally,E.M., Ben Othmane,K., Hagiwara,Y., Mizuno,Y., Yoshida,M., *et al.* (1995) Mutations in the dystrophin-associated protein γ-sarcoglycan in chromosome 13 muscular dystrophy. *Science* **270**, 819–822.
23. Online Mendelian Inheritance in Man (OMIM). The Human Genome Data Base Project. John Hopkins University, Baltimore, MD (cited 1996). World Wide Web URL http:// www3.ncbi.nlm.nih.gov/Omim
24. Bushby,K.M.D. and Beckmann,J.S. (1995) The limb-girdle muscular dystrophies—proposal for a new nomenclature. *Neuromusc. Disord.* **4**, 337–343.
25. Richard,I., Broux,O., Allamand,V., Fougerousse,F., Chiannilkulchai,N., Bourg,N., Brenguier,L., Devaud,C., Pasturaud P., Roudaut,C., Hillaire,D., Passos-Bueno,M.-R., Zatz,M., Tischfield,J.A., Fardeau,M., Jackson,C.E., Cohen,D. and Beckmann,J.S. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* **81**, 27–40.
26. Bashir,R., Strachan,T., Keers,S., Stephenson,A., Mahjneh,I., Marconi,G., Nashef,L. and Bushby,K.M.D. (1994) A gene for autosomal recessive limb-girdle muscular dystrophy maps to chromosome 2p. *Hum. Mol. Genet.* **3**, 455–457.
27. Bushby,K.M.D., Bashir,R., Keers,S., Britton,S., Meisler,M., Del Mastro,R., Lovett,M., Strachan,T. (1996) Molecular genetics of chromosome 2-linked LGMD. *Neuromusc. Disord.* **6**, No. 2 S7.
28. Speer,M.C., Yamaoka,L.H., Gilchrist,J.H., Gaskell,C.P., Stajich,J.M., Vance,J.M., Kazantsev,A., Lastra,A.A., Haynes,C.S., Beckmann,J.S., Cohen,D., Weber,J.L., Roses,A.D. and Pericak-Vance,M.A. (1992) Confirmation of genetic heterogeneity in limb-girdle muscular dystrophy: linkage of an autosomal dominant form to chromosome 5q. *Am. J. Hum. Genet.* **50**, 1211–1217.
29. Yamaoka,L.H., Westbrook,C.A., Speer,M.C., Gilchrist,J.M., Jabs,E.W., Schweins,E.G., Stajich,J.M., Gaskell,P.C., Roses,A.D. and Pericak-Vance, M.A. (1994) Development of a microsatellite genetic map spanning 5q31-q33 and subsequent placement of the LGMD1A locus between D5S178 and IL9. *Neuromusc. Disord.* **4**,471–475.
30. Speer,M.C., Gilchrist,J.M., Chutkow,J.G., McMichael,R., Westbrook,C., Stajich,J.M., Jorgenson,E.M., Gaskell,C. P., Rosi,B.L., Ramesan,R., Vance,J.M., Yamaoka,L.H., Roses,A.D., and Pericak-Vance,M.A. (1995) Evidence for locus heterogeneity in autosomal dominant limb-girdle muscular dystrophy. *Am. J. Hum. Genet.* **57**, 1371–1376.
31. Yamamoto,H., Hagiwara,Y., Mizuno,Y., Yoshida,M., and Ozawa,E. (1993) Heterogeneity of dystrophin-associated proteins. *J. Biochem.* **114**, 132–139.
32. Frohman,M.A., Dush,M.K. and Martin G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
33. Cheng,S., Fockler,C., Barnes,W.M., and Higuchi,R. (1994) Effective amplification of long targets from cloned inserts and human genomic DNA. *Proc. Natl. Acad. Sci. USA* **91**, 5695–5699.
34. Riley,J., Butler,R., Ogilvie,D., Finniear,R., Jenner,D., Powell,S., Anand,R., Smith,J.C. and Markham,A.F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18**, 2887–2890.
35. Matsumura,K., Tomé,F.M.S., Collin,H., Azibi,K., Chaouch,M., Kaplan,J.-C., Fardeau,M. and Campbell,K.P. (1992) Deficiency of the 50K dystrophin-associated glycoprotein in severe childhood autosomal recessive muscular dystrophy. *Nature* **359**, 320–322.
36. Azibi,K, Bachner,L., Beckmann,J.S., Matsumura,K., Hamouda,E., Chaouch,M., Chaouch,A., Ait-Ouarab,R., Vignal,A., Weissenbach,J., Vinet,M.-C., Leturcq,F., Collin,H., Tomé,F.M.S., Reghis,A., Fardeau,M., Campbell,K.P., and Kaplan,J.-C. (1993) Severe childhood autosomal recessive muscular dystrophy with the deficiency of the 50 kDa dystrophin-associated glycoprotein maps to chromosome 13q12. *Hum. Mol. Genet.* **2**, 1423–1428.
37. Nigro,V., Politano,L., Nigro,G., Colonna Romano,S., Molinari A.M.,& Puca G.A. (1992) Detection of a nonsense mutation in the dystrophin gene by multiple SSCP. *Hum. Mol. Genet.* **1**, 517–520.
38. Nigro,V., Bruni,P., Ciccodicola,A., Politano,L., Nigro,G., Piluso,G., Cappa,V., Covone,A.E., Romeo,G., and D'Urso M. (1995) SSCP detection of novel mutations in patients with Emery-Dreifuss muscular dystrophy: definition of a small C-terminal region required for emerin function. *Hum. Mol. Genet.* **4**, 2003–2004.
39. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular cloning: A Laboratory Manual* 2nd. edn. (Cold Spring Harbor Laboratory Press, New York).
40. Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
41. Towbin,H., Staehelin,T. and Gordon,J. (1979) Electrophoretic transfer of proteins from polyacrilamide gels to nitrocellulose sheets: procedures and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.

NOTE ADDED IN PROOF

An article reporting the linkage analysis of a novel form of autosomal recessive limb-girdle muscular dystrophy has now been published [Passos-Bueno *et al.* (1996) *Hum. Mol. Genet.* **5**, 815–820]. The Authors have mapped the new locus to 5q33-34. All these patients have a severe DMD phenotype and, in addition, are α-sarcoglycan negative. After EMBL submission of δ-sarcoglycan cDNA sequence, a second sarcoglycan EST (N30492) was found in the database.