THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 271, No. 50, Issue of December 13, pp. 32321–32329, 1996
Printed in U.S.A.

Characterization of δ-Sarcoglycan, a Novel Component of the Oligomeric Sarcoglycan Complex Involved in Limb-Girdle Muscular Dystrophy*

(Received for publication, August 8, 1996)

Daniel Jung‡§, Franck Duclos‡§, Barbara Apostol¶, Volker Straub‡, Jane C. Lee‡, Valérie Allamand‡, David P. Venzke‡, Yoshihide Sunada‡, Carolyn R. Moomaw‖, Cynthia J. Leveille‡, Clive A. Slaughter‖, Thomas O. Crawford, John D. McPherson‡‡, and Kevin P. Campbell‡§§**

*From the ‡Howard Hughes Medical Institute and Department of Physiology and Biophysics, University of Iowa College of Medicine, Iowa City, Iowa 52242, the ¶Department of Biological Chemistry, University of California, Irvine, Irvine, California 92717, the ‖Howard Hughes Medical Institute, Biopolymer Facility, University of Texas Southwestern Medical Center, Dallas, Texas 75235, the **Department of Neurology, The Johns Hopkins University, Baltimore, Maryland 21287-8811, and the ‡‡Department of Genetics/Genome Sequencing Center, Washington University School of Medicine, St. Louis, Missouri 63108*

The sarcoglycan complex is known to be involved in limb-girdle muscular dystrophy (LGMD) and is composed of at least three proteins: α-, β-, and γ-sarcoglycan. δ-Sarcoglycan has now been identified as a second 35-kDa sarcolemmal transmembrane glycoprotein that shares high homology with γ-sarcoglycan and is expressed mainly in skeletal and cardiac muscle. Biochemical analysis has demonstrated that γ- and δ-sarcoglycan are separate entities within the sarcoglycan complex and that all four sarcoglycans exist in the complex on a stoichiometrically equal basis. Immunohistochemical analysis of skeletal muscle biopsies from patients with LGMD2C, LGMD2D, and LGMD2E demonstrated a reduction of the entire sarcoglycan complex in these muscular dystrophies. Furthermore, we have mapped the human δ-sarcoglycan gene to chromosome 5q33-q34 in a region overlapping the recently linked autosomal recessive LGMD2F locus.

The dystrophin-glycoprotein complex (DGC)[1] (1–4) in skeletal muscle is a complex of sarcolemmal proteins and glycoproteins. It is composed of dystrophin, a cytoskeletal actin-binding protein (5–7); the syntrophins, a 59-kDa triplet of intracellular proteins that bind the C-terminal domain of dystrophin (8–12); α-dystroglycan, a 156-kDa extracellular proteoglycan that binds the G domain of laminin (13–15); β-dystroglycan, a 43-kDa transmembrane glycoprotein (2, 3, 13) that binds the cysteine-rich region of dystrophin (16, 17); α-, β-, and γ-sarcoglycan, transmembrane glycoproteins of 50, 43, and 35 kDa, respectively (18–24); and a 25-kDa transmembrane protein (1–4). Recent experiments have demonstrated the existence of two complexes within the DGC (24, 25): the dystroglycan complex, composed of α- and β-dystroglycan, and the sarcoglycan complex, consisting of α-, β-, and γ-sarcoglycan.

Defects in DGC components lead to muscle fiber necrosis, the major pathological event in muscular dystrophies (26). In Duchenne muscular dystrophy (DMD), mutations in the dystrophin gene cause the loss of dystrophin and a reduction of the dystrophin-associated proteins (2, 5). One form of congenital muscular dystrophy has recently been characterized as being caused by mutations in the laminin α2-chain gene (27, 28). Limb-girdle muscular dystrophy (LGMD) represents a clinically and genetically heterogeneous class of disorders (29, 30). They are inherited as either autosomal dominant or recessive traits. An autosomal dominant form, LGMD1A, was mapped to 5q31-q33 (31, 32), while six genes involved in the autosomal recessive forms were mapped to 15q15.1 (LGMD2A) (33), 2p16-p13 (LGMD2B) (34), 13q12 (LGMD2C) (23, 35, 36), 17q12-q21.33 (LGMD2D) (19, 20), 4q12 (LGMD2E) (21, 22), and most recently 5q33-q34 (LGMD2F) (37). Patients with LGMD2C, -2D, and -2E have a deficiency of components of the sarcoglycan complex resulting from mutations in the genes encoding γ-, α-, and β-sarcoglycan, respectively (19, 21–23, 38–40). The gene responsible for LGMD2A has been identified as the muscle-specific calpain (41), whereas the genes responsible for LGMD1A, -2B, and -2F are still unknown.

Here, we first describe the cloning of a cDNA encoding a fourth sarcoglycan protein (δ-sarcoglycan), a novel 35-kDa component of the dystrophin-glycoprotein complex. Human δ-sarcoglycan is a transmembrane glycoprotein that is mainly expressed in skeletal and cardiac muscle and shares about 60% protein sequence identity with the recently cloned γ-sarcoglycan. We demonstrate that γ- and δ-sarcoglycan are separate entities within the DGC, despite their similar molecular weights and protein sequences. δ-Sarcoglycan is not an isoform of γ-sarcoglycan. We also show that α-, β-, γ-, and δ-sarcoglycan are equal in the sarcoglycan complex on a stoichiometric basis. We demonstrate that δ-sarcoglycan is reduced along with α-, β-, and γ-sarcoglycan in LGMD2C, -2D, and -2E. Therefore, a common feature of LGMD2C, -2D, and -2E is a specific loss of the entire sarcoglycan complex from the sarcolemma membrane. Finally, we map the human δ-sarcoglycan gene to chro-

* This work was supported in part by the Muscular Dystrophy Association. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ These authors contributed equally to this work.

§§ Investigator of the Howard Hughes Medical Institute. To whom correspondence should be addressed: Howard Hughes Medical Institute, Dept. of Physiology and Biophysics, University of Iowa College of Medicine, 400 Eckstein Medical Research Bldg., Iowa City, IA 52242-1101. Tel.: 319-335-7867; Fax: 319-335-6957; E-mail: kevin-campbell@uiowa.edu; WWW: http://www-camlab.physlog.uiowa.edu/home.htm.

[1] The abbreviations used are: DGC, dystrophin-glycoprotein complex; DMD, Duchenne muscular dystrophy; LGMD, limb-girdle muscular dystrophy; BSA, bovine serum albumin; PCR, polymerase chain reaction; PBS, phosphate-buffered saline; FITC, fluoroisothiocyanate; PAGE, polyacrylamide gel electrophoresis; HPLC, high pressure liquid chromatography.

mosome 5q33–34 within the recently reported disease interval for LGMD2F (5q33–34) (37).

EXPERIMENTAL PROCEDURES

Peptide Sequencing and Isolation of Human δ-Sarcoglycan cDNA—Peptides of the 35-kDa component of purified rabbit skeletal muscle DGC were obtained as described previously (24). These peptide sequences were used to search the data base of expressed sequence tags (dbEST) using the TBLASTN search program at the National Center for Biotechnology Information. An EST isolated from human placenta cDNA library (42) was identified. Clone 259220, from which the EST was generated, was obtained from Research Genetics, Inc. The clone was fully sequenced on both strands using an Applied Biosystems, Inc. automated sequencer and determined to contain the full coding region of the human δ-sarcoglycan cDNA. Primary structure and site detection analysis were performed using PC/GENE software (Oxford Molecular).

Northern Blot Analysis—Clontech adult human multiple tissue Northern blots containing 2 μg of poly(A)+ RNA per lane were probed with the full-length δ-sarcoglycan cDNA clone.

Antibodies to Components of the Dystrophin Glycoprotein Complex—Antibodies against the different components of the dystrophin-glycoprotein complex were produced in a goat using the purified rabbit skeletal muscle dystrophin-glycoprotein complex (1) as previously described with a sheep (43) (Elmira Biologicals). Goat 20 antiserum recognizes all components of the DGC, and specific antibodies were affinity purified using Immobilon P strips of DGC components or fusion proteins of DGC components. Monoclonal antibodies $VIA4_2$ against dystrophin and $IVD3_1$ against α-sarcoglycan were previously characterized (2). Affinity-purified rabbit 172 antibody against β-sarcoglycan and affinity-purified sheep 0025 antibody against γ-sarcoglycan were previously characterized (21, 24). Monoclonal antibody against the human laminin (MAB1922) α2 chain was purchased from Chemicon.

δ-Sarcoglycan Antibodies—A peptide corresponding to the rabbit skeletal muscle δ-sarcoglycan sequence (amino acids 166–176) was chemically synthesized with an additional cysteine at the N terminus. This synthetic peptide was conjugated with the N-terminal cysteine of keyhole limpet hemocyanin (Pierce) using *m*-maleimidobenzoic acid-*N*-hydroxysuccinimide ester (Pierce), mixed with Freund's adjuvant (Sigma) and injected into sheep 0074. Polyclonal antibodies against the peptide were affinity purified from crude sera using BSA-conjugated peptide as described previously (24).

Fusion Protein Constructs—The extracellular domains of γ-sarcoglycan (amino acids 62–291) and δ-sarcoglycan (amino acids 93–256) were amplified by PCR in Perkin Elmer Thermocycler using specific primers containing *Bam*HI or *Eco*RI restriction sites. Following restriction digest with *Bam*HI and *Eco*RI, the products were subcloned into pGEX2TK (44) and introduced into DH5α cells. Overnight cultures (50 ml) were diluted 1:10 and induced with isopropyl thiogalactopyranoside (United States Biochemicals) to promote fusion protein production. Fusion proteins were purified on a glutathione-agarose column (45).

Immunoblot and Immunofluorescence Analysis—Crude rabbit skeletal muscle sarcolemma, purified DGC (46, 47), and isolated sarcoglycan complex (47) were prepared as described previously. Proteins were resolved on a 3–12%, 3–20%, 5–12%, or 10% SDS-polyacrylamide gel (24) stained with Coomassie Blue or transferred to nitrocellulose by electroblotting (49). Blots were incubated overnight in a 1:15 dilution of affinity-purified anti-δ-sarcoglycan antibody, 1:15 dilution of affinity-purified anti-γ-sarcoglycan antibody, or 1:15 dilution of affinity-purified anti-fusion protein M from goat 20 antibody in Blotto (5% nonfat dried milk in Tris-buffered saline (20 mM Tris-HCl, 200 mM NaCl, pH 7.4)). Blots were then incubated with a horseradish peroxidase-conjugated anti-sheep IgG or anti-goat IgG secondary antibody (Boehringer Mannheim) for 1 h. Antibody staining was detected with H_2O_2 in Tris-buffered saline with 4-chloro-1-naphthol as a substrate or by ECL (Amersham). Coomassie Blue-stained gels were analyzed densitometrically using a Molecular Dynamics model 300A scanning densitometer.

For immunofluorescence, 7-μm transverse cryosections were prepared from control, DMD, and LGMD muscle biopsies. The following procedures were performed at room temperature. Sections were treated with AB blocking solutions (Vector), blocked with 5% BSA in PBS for 30 min, and then incubated with a 1:20 dilution of affinity-purified anti-δ-sarcoglycan antibody in 5% BSA/PBS for 90 min. Antibodies against the following components of the DGC were also tested: dystrophin, laminin α2 chain, α-sarcoglycan, β-sarcoglycan, and γ-sarcoglycan. After extensive washing with PBS/1% BSA, sections were incubated with biotinylated secondary antibodies (1:500) for 30 min, washed with PBS/1% BSA, and then incubated with FITC-conjugated streptavidin (1:1000) for 30 min. After rinsing with PBS, sections were mounted with FITC-guard (Testog) and observed under a Zeiss Axioplan fluorescence microscope.

TABLE I
Peptide sequences from 35-kDa component of the dystrophin-glycoprotein complex

Tryptic peptides were prepared for acquisition of internal sequence information by SDS-PAGE of purified DGC through 10% acrylamide gels and electrotransferred to polyvinylidene difluoride membrane, and the protein band corresponding to the 35-kDa protein was excised. The immobilized protein was then digested with trypsin. Peptides were purified by reverse phase HPLC using Applied Biosystems model 130A HPLC system equipped with 2.1 × 100-mm RP-300 column. The purified tryptic peptides were sequenced by Automated Edman degradation using an Applied Biosystems model 470A sequencer equipped with an on-line model 120A phenylthiohydantoin derivative analyzer using the manufacturer's standard programming and chemicals.

Amino acid sequence	Position in γ-sarcoglycan	Position in δ-sarcoglycan
EQYLTATE	3–11	
DGLILEGES	76–84	
LKVGPQXVEVQSQQXQINS	126–144	
SLFTVDEEEVVVGTDRL	148–165	
VTGPEGALFEHSVETPLV	167–184	
SLSMDAP	200–207	
LVQGTQAASG	248–57	
STMPGSVGPQVYK		12–24
GLKLEGDSEFLQPLYAK		76–92
EIQSLPGNALYFK		93–105
NVTVNILNDQTK		109–120
VLTQLITGP		121–129
KFEVK		138–142
LFSADMNNXVVGAFR		149–163
VLGAEGTVFPK		166–176
IETPNVR		178–184
ADPFKEL		185–191
ALVMEAP		199–205
VEINAEAGNMEA		208–219

Enzymatic Deglycosylation—DGC (0.5 mg/ml) in 20 mM $NaHPO_4$, pH 7.5, 50 mM EDTA, and 0.2% SDS was boiled for 2 min and incubated 16 h at 37 °C after addition of 1% Triton X-100 and 5 units of endoglycosidase/PNGase F (Oxford Glycosystems). After incubation, the samples were resolved on 10% or 5–15% SDS-PAGE and analyzed either on Western blot with the affinity-purified anti-δ-sarcoglycan and affinity-purified anti-γ-sarcoglycan antibodies or by Coomassie Blue staining.

Localization of the δ-Sarcoglycan Gene to Chromosome 5q33-q34—Primers corresponding to human δ-sarcoglycan cDNA nucleotides 781–800 (sense) and 1040–1060 (antisense) were used to amplify a 280-bp product. The following cycling parameters were used: 94 °C, 3 min (94 °C, 1 min; 58 °C, 1 min; 72 °C, 1 min) × 35; 72 °C, 7 min. DNA from a panel of 25 human rodent somatic cell hybrids (BIOS Corp.), containing various combinations of human chromosomes or different regions of chromosome 5q, and DNA isolated from 180 chromosome 5-specific radiation hybrids were analyzed by PCR. PCR products were separated using 3% agarose, and the hybrids were scored for the presence or absence of the specific PCR product using ethidium bromide. The position of the δ-sarcoglycan locus relative to previously analyzed markers was determined using the rh2pt and rhmaxlik programs (50).

Limb-Girdle Muscular Dystrophy Patients—Patients were initially identified with LGMD clinical features. Mutation analysis in the sarcoglycan genes allowed the confirmation of the LGMD phenotype and the classification of the disease as LGMD2C, -2D, and -2E. The LGMD2C patient was found to have deletions in each allele of the γ-sarcoglycan gene, a 124-bp deletion and a 208-bp deletion. The LGMD2D patient had missense mutation in each allele of the α-sarcoglycan gene, a substitution of arginine in position 77 to a cysteine, and isoleucine in position 124 to a threonine. The LGMD2E patient has been described previously (21).

RESULTS

δ-Sarcoglycan cDNA and Primary Structure Analysis—To characterize the 35-kDa component of the DGC, purified rabbit skeletal muscle DGC was separated by SDS-PAGE and transferred to polyvinylidene difluoride membrane. The region of the membrane containing the 35-kDa protein was excised, and the immobilized protein was digested with trypsin. Peptides were

a

```
TTACTGCCGGGAGTGTTGAGTGAAGGGACCAGGTGGAGATGATGCCTCAGGAGCAGTAC   59
                                          M  M  P  Q  E  Q  Y    7

ACTCACCACCGGAGCACCATGCCTGGCTCTGTGGGGCCACAGGTATACAAGGTGGGGATT 119
 T  H  H  R  S  T  M  P  G  S  V  G  P  Q  V  Y  K  V  G  I   27

TATGGCTGGCGGAAACGATGCCTGTATTTCTTTGTCCTGCTCCTCATGATTTTAATACTG 179
 Y  G  W  R  K  R  G  L  Y  F  F  V  L  L  L  M  I  I  I  L   47

GTGAACTTGGCCATGACCATCTGGATTCTCAAAGTCATGAACTTCACAATTGATGGAATG 239
 V  N  L  A  M  T  I  W  I  L  K  V  M  N*  F  T  I  D  G  M   67

GGAAACCTGAGGATCACAGAAAAAGGTCTAAAGCTAGAAGGAGACTCTGAATTCTTACAA 299
 G  N  L  R  I  T  E  K  G  L  K  L  E  G  D  S  E  F  L  Q   87

CCTCTCTACGCCAAAGAAATCCAGTCCCGACCAGGTAATGCCCTGTACTTCAAGTCTGCC 359
 P  L  Y  A  K  E  I  Q  S  R  P  G  N  A  L  Y  F  K  S  A  107

AGAAATGTTACAGTGAACATTCTCAATGACCAGACTAAAGTGCTAACTCAGCTTATAACA 419
 R  N*  V  T  V  N  I  L  N  D  Q  T  K  V  L  T  Q  L  I  T  127

GGTCCAAAAGCCGTAGAAGCTTATGGTAAAAAATTTGAGGTAAAAACTGTTTCTGGAAAA 479
 G  P  K  A  V  E  A  Y  G  K  K  F  E  V  K  T  V  S  G  K  147

TTGCTCTTCTCTGCAGACAATAATGAAGTGGTAGTAGGAGCTGAAAGATTACGAGTTTTA 539
 L  L  F  S  A  D  N  N  E  V  V  V  G  A  E  R  L  R  V  L  167

GGAGCGGAGGGCACAGTGTTCCCTAAATCTATAGAAACACCTAATGTCAGGGCAGACCCC 599
 G  A  E  G  T  V  F  P  K  S  I  E  T  P  N  V  R  A  D  P  187

TTCAAAGAACTAAGGTTGGAGTCCCCAACCCGGTCTCTAGTGATGGAGGCCCCAAAAGGA 659
 F  K  E  L  R  L  E  S  P  T  R  S  L  V  M  E  A  P  K  G  207

GTGGAAATCAATGCAGAAGCTGGCAATATGGAAGCCACCTGCAGGACAGAGCTGAGACTG 719
 V  E  I  N  A  E  A  G  N  M  E  A  T  C  R  T  E  L  R  L  227

GAATCCAAAGATGGAGAGGTGAGGGATGAGAAGGACAGAAGTTCAAAGAGCTACAGCTTC 779
 E  S  K  D  G  E  V  R  D  E  K  D  R  S  S  K  S  Y  S  F  247

AACAGGCCAACCCTTCCCATAACTGGTTGACCTCGGAGTTGGATCCTACAGTGTATCAAC 839
 N  R  P  T  L  P  I  T  G  stop                               256

AAAAGGAGCCAAGCAGGTTTTATTTCTGAAACAATTAATTGAGCAGCATGATTATAAGCC 899
AAACCCACAATCCATCAAAGTGATGATTTCTTATTTGTAAAATGCGGAGATAATGGCATG 959
TATTCCAAGTACAGAATTATATGACCATGAAAATGAATGCTATTTTCAAATTCTCTCTTG 1019
TCACCTTAAAATAAGATTTTGTTAGCCAACATAATTAAGCTGTATATATTATACACATCT 1079
GGCTCAAGAAAAAAAAAAAAAAAAAAAAAAA                              1110
```

b

```
Gamma MVREQYTTATE GICIERPENQ YVYKIGIYGW RKRCLYLFVL LLLIILVVNL  51
Delta  MMPQEQYTHH RSTMPGSVGP QVYKVGIYGW RKRCLYFFVL LLMILIIVNL  50

Gamma  ALTIWILKVM WFSPAGMGHL CVTKDGLRLE GESEFLFPLY AKEIHSRVDS 101
Delta  AMTIWILKVM NFTIDGMGNL RITEKGLKLE GDSEFLQPLY AKEIQSRPGN 100

Gamma  SLLLQSTQNV TVNARNSEGE VTGRLKVGPK MVEVQNQQFQ INSNDGKPLF 151
Delta  ALYFKSARNV TVNILNDQTK VLTQLITGPK AVEAYGKKFE VKTVSGKLLF 150

Gamma  TVDEKEVVVG TDKLRVIGPE GALFEHSVET PLVRADPFQD LRLESPTRSL 201
Delta  SADNNEVVVG AERLRVLGAE GTVFPKSIET PNVRADPFKE LRLESPTRSL 200

Gamma  SMDAPRGVHI QAHAGKIEAL SQMDILFHSS DGMLVLDAET VCLPKLVQGT 251
Delta  VMEAPKGVEI NAEAGNMEAT CRTELRLESK DGEVRDEKDR SSKSYSFNRP 250

Gamma  WGPSGSSQSL YEICVCPDGK LYLSVAGVST TCQEHSHICL           291
Delta  TLPITG                                               256
```

FIG. 1. **Human δ-sarcoglycan cDNA and translated amino acid sequences.** *a*, human δ-sarcoglycan. The predicted transmembrane domain is indicated by the *shadowed box*. Two consensus sites for *N*-linked glycosylation are indicated by *asterisks*. Peptide sequence fragments from the 35-kDa band of the purified DGC are *underlined*. *b*, amino acid sequence comparison of δ- and γ-sarcoglycan. The δ-sarcoglycan sequence is available from GenBank under accession number U58331.

purified by reverse-phase HPLC and subjected to amino acid sequencing by Edman degradation. Different peptide sequences were obtained (Table I). We compared these peptide sequence fragments with the protein sequence of the recently reported γ-sarcoglycan primary structure, a 35-kDa component of the DGC (23). Several of these peptide sequences were not

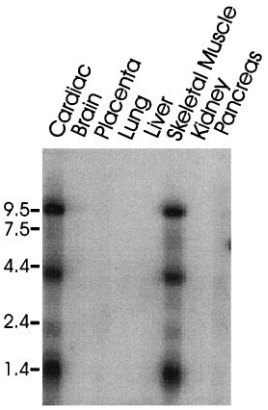

FIG. 2. **Expression of δ-sarcoglycan in adult human tissues.** δ-sarcoglycan cDNA was used to probe a blot of poly(A)$^+$ selected RNA (Clontech) from various human tissues.

found in the reported γ-sarcoglycan sequence, indicating the existence of at least a second protein of similar molecular weight within the DGC. To characterize this second 35-kDa protein, we searched the GenBank data base of expressed sequence tags (dbEST) with the unidentified peptide sequences. An EST encoding one of the peptide fragments was identified and isolated from a normalized human placenta cDNA library constructed by Bento Soares *et al*. (42). Clone 259220, from which the EST was generated, was obtained from the IMAGE Consortium and fully sequenced on both strands. The 1.1-kilobase cDNA sequence contains a single 768-bp open reading frame with a presumptive initiator methionine at nucleotide 38 preceded by an in-frame stop codon, a stop codon at nucleotide 806, and a poly(A)$^+$ tail. The open reading frame encodes a protein of 256 amino acids with a predicted molecular mass of 29 kDa and a predicted isoelectric point of 9.48 and contains most of the 35-kDa peptides that were not present in the primary structure of γ-sarcoglycan (Table I, Fig. 1*a*). Amino acid sequence analysis revealed a single transmembrane domain and no N-terminal signal sequence. Therefore, the protein is likely to be a type II transmembrane protein with an intracellular N terminus. This membrane topology is consistent with the location of two putative *N*-linked glycosylation sites, which are C-terminal to the transmembrane domain. The predicted membrane organization is similar to that of α-, β-, and γ-sarcoglycan, all of them having single transmembrane domains and large extracellular and small intracellular domains (18–23). Furthermore, the amino acid sequence is about 60% identical to γ-sarcoglycan (Fig. 1*b*), suggesting a similarity in function for these proteins. No significant homologies to other previously characterized proteins or domains were found in the data bases.

Tissue Distribution of δ-Sarcoglycan mRNA—To determine the tissue distribution of δ-sarcoglycan, we performed Northern blot analysis of RNA from multiple human adult tissues with the full-length coding region of δ-sarcoglycan. A predominant transcript of approximately 1.4 kilobases is detected in skeletal and cardiac muscle (Fig. 2). Signals of 2.3, 4.3 and 9.5 kilobases were also detected in these tissues when probing with the full-length coding region of δ-sarcoglycan. This result suggests the existence of alternatively spliced and/or polyadenylated forms of the message. The expression pattern is similar to that of α- and γ-sarcoglycan transcripts, which are only expressed in skeletal and cardiac muscle (18–20, 23), but differs from the β-sarcoglycan transcript, which is ubiquitously expressed (21, 22).

Molecular Organization of the Sarcoglycan Complex—To characterize the δ-sarcoglycan protein, polyclonal antibodies were produced against a specific synthetic peptide (amino acids 166–176). This peptide sequence is not conserved entirely in γ-sarcoglycan and likely represents a unique epitope for δ-sarcoglycan in the extracellular domain. To determine the specificity of the affinity-purified anti-δ-sarcoglycan peptide antibodies, glutathione *S*-transferase fusion proteins, containing either the extracellular domain of γ-sarcoglycan (FP-N) or the extracellular domain of δ-sarcoglycan (FP-M) were constructed. These fusion proteins were then tested for reactivity with the affinity-purified anti-δ-sarcoglycan peptide antibodies. As shown in Fig. 3*a*, affinity-purified anti-γ-sarcoglycan peptide antibodies (24) stained fusion protein N but not fusion protein M, whereas affinity-purified anti-δ-sarcoglycan peptide antibodies stained fusion protein M but not fusion protein N. In addition, affinity-purified anti-FP-M antibodies from goat serum immunized with the entire DGC stained both fusion protein N and M. These results confirmed the specificity of the antibodies for γ- and δ-sarcoglycan. Furthermore, the fact that affinity-purified anti-FP-M antibodies reacted with both fusion protein N and M demonstrated the sequence homology between γ- and δ-sarcoglycan. The δ-sarcoglycan affinity-purified antibodies reacted on immunoblots with a single protein band at 35 kDa in rabbit skeletal crude sarcolemmal membranes and purified DGC (Fig. 3*b*). The sarcolemmal staining was much weaker than in the purified DGC, indicating an enrichment of δ-sarcoglycan in the DGC. In immunofluorescence studies, affinity-purified anti-peptide antibodies selectively stained the sarcolemma of human skeletal muscle (Fig. 3*c*). Skeletal muscle from DMD patients showing a reduction of DGC components also showed reduced sarcolemmal expression of δ-sarcoglycan by immunofluorescence analysis (Fig. 3*c* and data not shown). Thus, enrichment of δ-sarcoglycan in the DGC and its reduction in muscle from DMD patients demonstrated that it is an integral component of the DGC.

We have previously demonstrated that α-, β-, and γ-sarcoglycan are tightly associated in the sarcolemma and constitute the sarcoglycan complex within the DGC (24). The high degree of sequence similarity between γ- and δ-sarcoglycan suggested that the latter may also be a member of the sarcoglycan complex. To address this question, we have probed the sarcoglycan complex, prepared from a fraction enriched in microsomes (24), with affinity-purified anti-δ-sarcoglycan peptide antibodies. δ-Sarcoglycan was specifically detected in the purified sarcoglycan complex (Fig. 3*d*), demonstrating its tight association with α-, β-, and γ-sarcoglycan.

Similarity in molecular weight between γ- and δ-sarcoglycan has so far rendered these proteins indistinguishable from each other by SDS-PAGE. However, as determined by Western blot analysis, treatment of purified DGC with endoglycosidase F/PNGase F resulted in a decrease of the apparent molecular mass of γ-sarcoglycan by approximately 2 kDa, whereas the apparent molecular mass of δ-sarcoglycan decreased by approximately 3 kDa (Fig. 3, *e* and *f*). Therefore, deglycosylation of γ- and δ-sarcoglycan allows the separation of these two proteins by SDS-PAGE. This finding also confirms the *N*-linked glycosylation of δ-sarcoglycan predicted by sequence analysis and suggests that the discrepancy between the calculated and apparent molecular weights is primarily due to glycosylation. These results demonstrate again the specificity of both γ- and δ-sarcoglycan antibodies. Furthermore, densitometric analysis of Coomassie Blue-stained SDS-PAGE of endoglycosidase F/PNGase F-treated DGC showed the four sarcoglycan proteins to be equal on a stoichiometric basis (Fig. 4).

Chromosomal Localization—To determine the chromosomal localization of the δ-sarcoglycan gene, we designed primers to

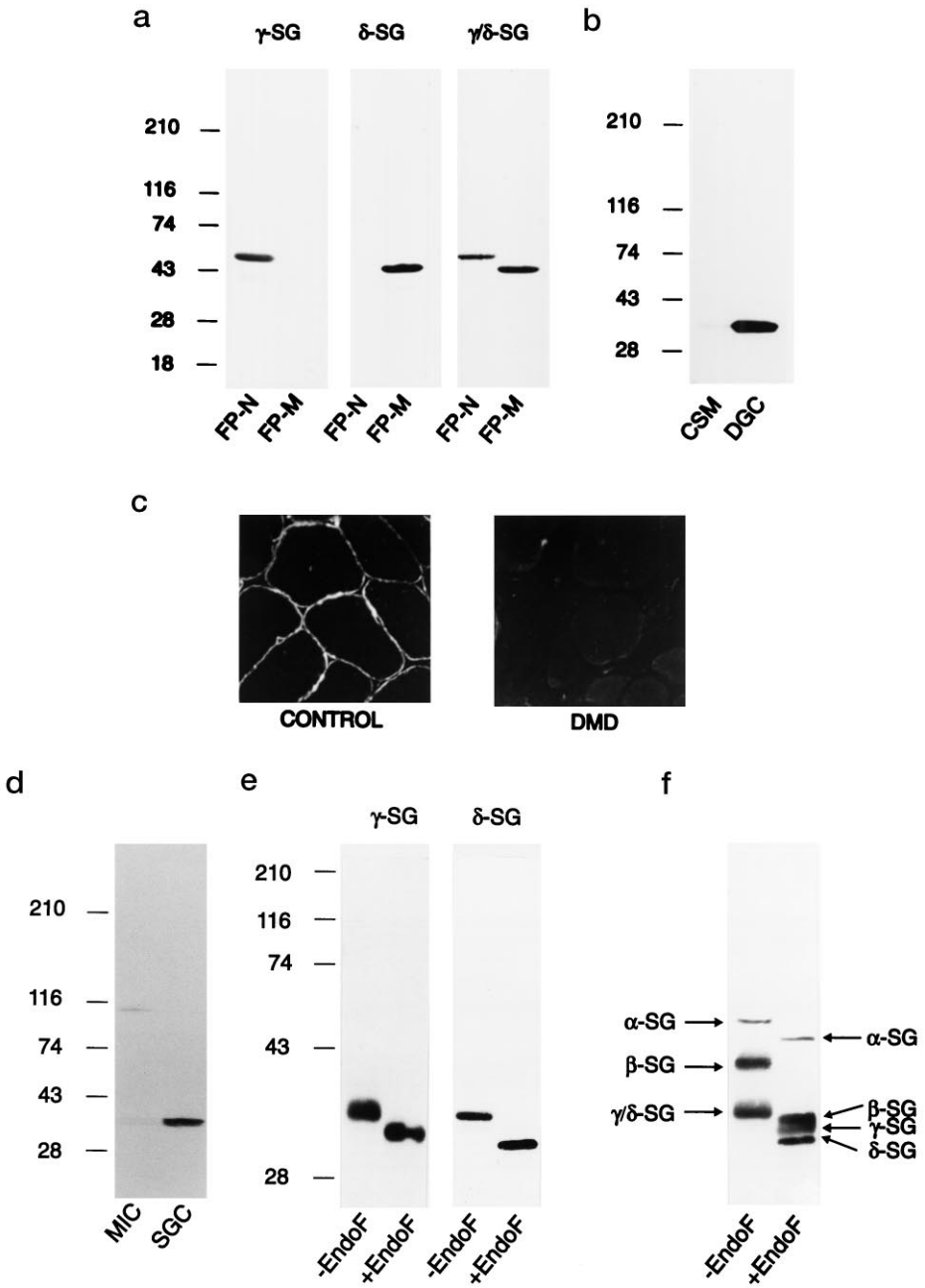

FIG. 3. **Characterization of δ-sarcoglycan protein.** *a*, γ-sarcoglycan extracellular domain fusion protein (FP-N) and δ-sarcoglycan extracellular domain fusion protein (FP-M) were resolved on 3–20% SDS-PAGE and analyzed by Western blotting with affinity-purified anti γ-sarcoglycan peptide antibodies (γ-*SG*), affinity-purified anti δ-sarcoglycan peptide antibodies (δ-*SG*), or affinity-purified anti-fusion protein M antibodies from goat 20 (γ/δ-*SG*). *b*, rabbit crude sarcolemmal membranes (*CSM*) and purified dystrophin-glycoprotein complex (DGC) were resolved on 3–12% SDS-PAGE and analyzed by Western blotting with affinity-purified anti δ-sarcoglycan peptide antibodies. *c*, immunofluorescence analysis of transverse cryosections of normal and DMD skeletal muscle stained with affinity-purified anti δ-sarcoglycan antibodies. *d*, rabbit skeletal muscle solubilized microsomes (*MIC*) and isolated sarcoglycan complex (SGC) were resolved on 3–12% SDS-PAGE and analyzed by Western blotting with affinity-purified anti δ-sarcoglycan peptide antibodies. *e*, effect of endoglycosidase F/PNGase F on δ-sarcoglycan electrophoretic mobility. Purified DGC was treated (+*endoF*) or not (−*endoF*) with endoglycosidase/PNGase F, resolved on 10% SDS-PAGE, and analyzed by Western blotting with affinity-purified anti γ-sarcoglycan peptide antibodies (γ-*SG*) or anti δ-sarcoglycan peptide antibodies (δ-*SG*). *f*, effect of endoglycosidase F/PNGase F on sarcoglycan protein electrophoretic mobility. Purified DGC was treated (+*endoF*) or not (−*endoF*) with endoglycosidase/PNGase F, resolved on 10% SDS-PAGE, and analyzed by Western blotting with antibodies to α-, β-, γ-, and δ-sarcoglycan. The positions of the sarcoglycan proteins are indicated by *arrows*.

amplify a fragment of the δ-sarcoglycan gene from a panel of human rodent somatic cell hybrids containing various combinations of human chromosomes. Using this method, the human δ-sarcoglycan gene was assigned to chromosome 5q (data not shown). To narrow down the chromosomal region, we used the same process to analyze DNA isolated from human rodent somatic cell hybrids containing various fragments of chromosome 5q. The δ-sarcoglycan gene was mapped to 5q31.3-q33.2 (data not shown), a region that overlaps with the candidate interval for autosomal dominant LGMD1A (31, 32). We then analyzed a radiation hybrid panel, which allowed us to precisely map the gene between markers D5S497 and D5S378 (Fig. 5 and Table II). Passos-Bueno *et al.* (37) recently reported a new form of LGMD linked to 5q33-q34 (LGMD2F) and mapped the gene between markers D5S470 and D5S820, a region that overlaps the interval defined by markers D5S497

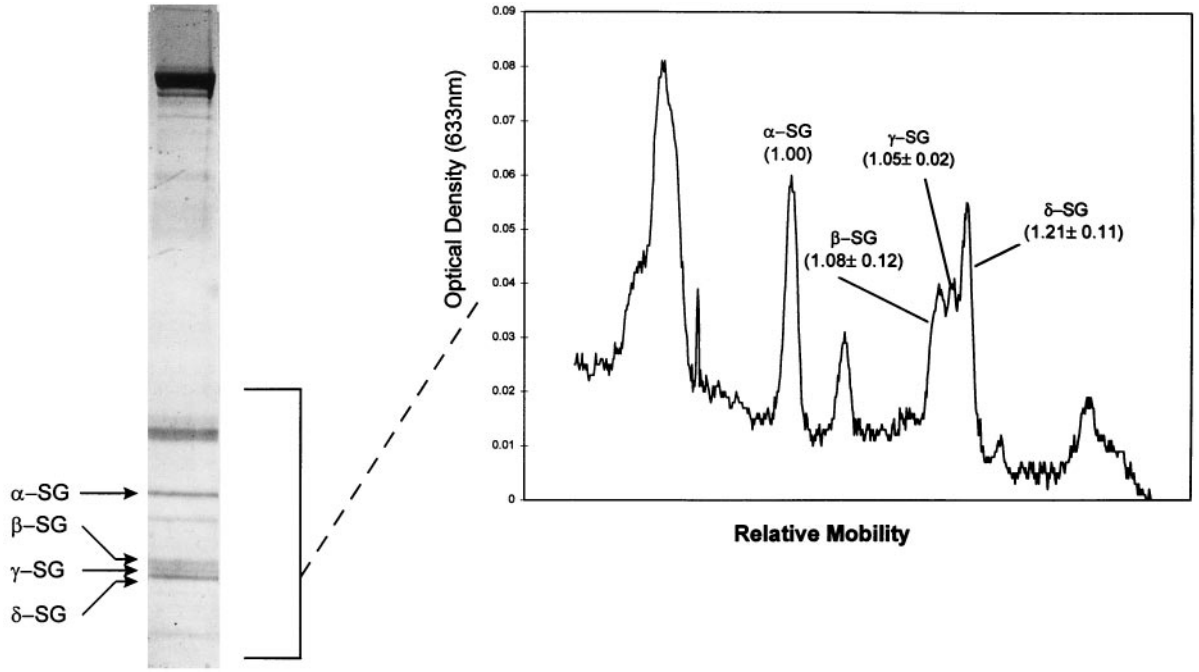

FIG. 4. **Stoichiometric ratios of the sarcoglycan complex.** Purified DGC was treated with endoglycosidase/PNGase F, resolved on 5–12% SDS-PAGE, and stained with Coomassie Blue. The position of α-, β-, γ-, and δ-sarcoglycan (α-*SG*, β-*SG*, γ-*SG*, and δ-*SG*) are indicated by *arrows*. Densitometric scan of the Coomassie Blue stained gel was performed, and the volume of the peaks shown in *parentheses* were converted to molar ratio with α-sarcoglycan set at 1 (*right panel*). Values represent the means of three independent measurement ± S.D. The peak between α- and β-sarcoglycan corresponds to β-dystroglycan.

and D5S378 (Fig. 5). The chromosomal localization of the δ-sarcoglycan gene to 5q33-q34, together with the involvement of the other sarcoglycans in LGMDs, makes δ-sarcoglycan a strong candidate for LGMD2F.

Sarcoglycan Complex in LGMD—It has been demonstrated that mutations in α-, β-, or γ-sarcoglycan lead to LGMD, characterized by a marked reduction of the sarcoglycan complex (19, 21–23, 38, 39). Since δ-sarcoglycan is a novel member of this complex, we examined the status of this protein in skeletal muscle biopsy specimens from patients with characterized LGMD2C, -2D, and -2E. Immunofluorescence analysis revealed an altered expression of δ-sarcoglycan compared to normal human skeletal muscle. Some patients showed absence of the protein, whereas in others some sarcolemmal staining was still observed. In addition to a reduced expression of δ-sarcoglycan, those patients also showed a reduction or total loss of the other sarcoglycan proteins (Fig. 6). Therefore, a common feature of LGMD2C, -2D, and -2E is a specific absence of the sarcoglycan complex at the sarcolemma.

DISCUSSION

The different components of the DGC were initially characterized based on their electrophoretic mobilities (1–4). One component, identified as a 35-kDa protein band has recently been cloned and named γ-sarcoglycan (23). However, we demonstrated here by direct protein sequencing and molecular cloning that the 35-kDa protein band contained a second protein distinct from γ-sarcoglycan and was encoded by a different gene. The identity of this protein, named δ-sarcoglycan, was characterized using a specific antibody. In particular, δ-sarcoglycan copurified with the DGC and sarcoglycan complexes and colocalized with the other components of the DGC to the sarcolemma. The identification of this novel component of the sarcoglycan complex allowed us to update the molecular organization of the dystrophin-glycoprotein complex (Fig. 7).

The predicted molecular mass of δ-sarcoglycan is 29 kDa, whereas its apparent molecular mass is 35 kDa. This discrepancy is mainly due to *N*-glycosylation of δ-sarcoglycan. Interestingly, the structure of δ-sarcoglycan is similar to β- and γ-sarcoglycan, with a short N-terminal intracellular domain and a large glycosylated C-terminal extracellular domain. Furthermore, δ- and γ-sarcoglycan share about 60% identity in primary structure, suggesting that these two proteins may have a similar function within the sarcoglycan complex. The structural homology as well as the similarity in molecular weight between γ- and δ-sarcoglycan have rendered these proteins indistinguishable by Western blot analysis. However, we demonstrated here that deglycosylation of purified DGC allows the separation of γ- and δ-sarcoglycan. Furthermore, the predicted isoelectric point (pI) of δ-sarcoglycan is 9.48, whereas γ-sarcoglycan displays a pI of 5.0 (23). Yamamoto *et al.* (51) have shown that the 35-kDa band from purified DGC is resolved on two-dimensional SDS-PAGE into two distinct spots, one being identified as γ-sarcoglycan with an acid pI and the second as an uncharacterized more basic protein. We observed that this more basic protein is specifically detected with the anti-δ-sarcoglycan antibodies.[2] Therefore, two-dimensional SDS-PAGE gel analysis of purified DGC may also allow the separation of γ-sarcoglycan from δ-sarcoglycan.

δ-Sarcoglycan, similarly to α- and γ-sarcoglycan, is strongly expressed in skeletal and cardiac muscle, whereas the β-sarcoglycan transcript is ubiquitously expressed. Furthermore, several transcripts were detected for δ-sarcoglycan, suggesting the existence of alternatively spliced forms. However, two peptides from the 35-kDa protein band were not found in γ- or δ-sarcoglycan. Using the TBLASTN search program, we identified an EST from mouse embryonic skeletal muscle containing one of these peptides. This EST encodes a C-terminal isoform of δ-sarcoglycan. Further work is currently in progress to characterize this isoform in human skeletal muscle.

Immunostaining of skeletal muscle biopsies from patients

[2] D. Jung, unpublished results.

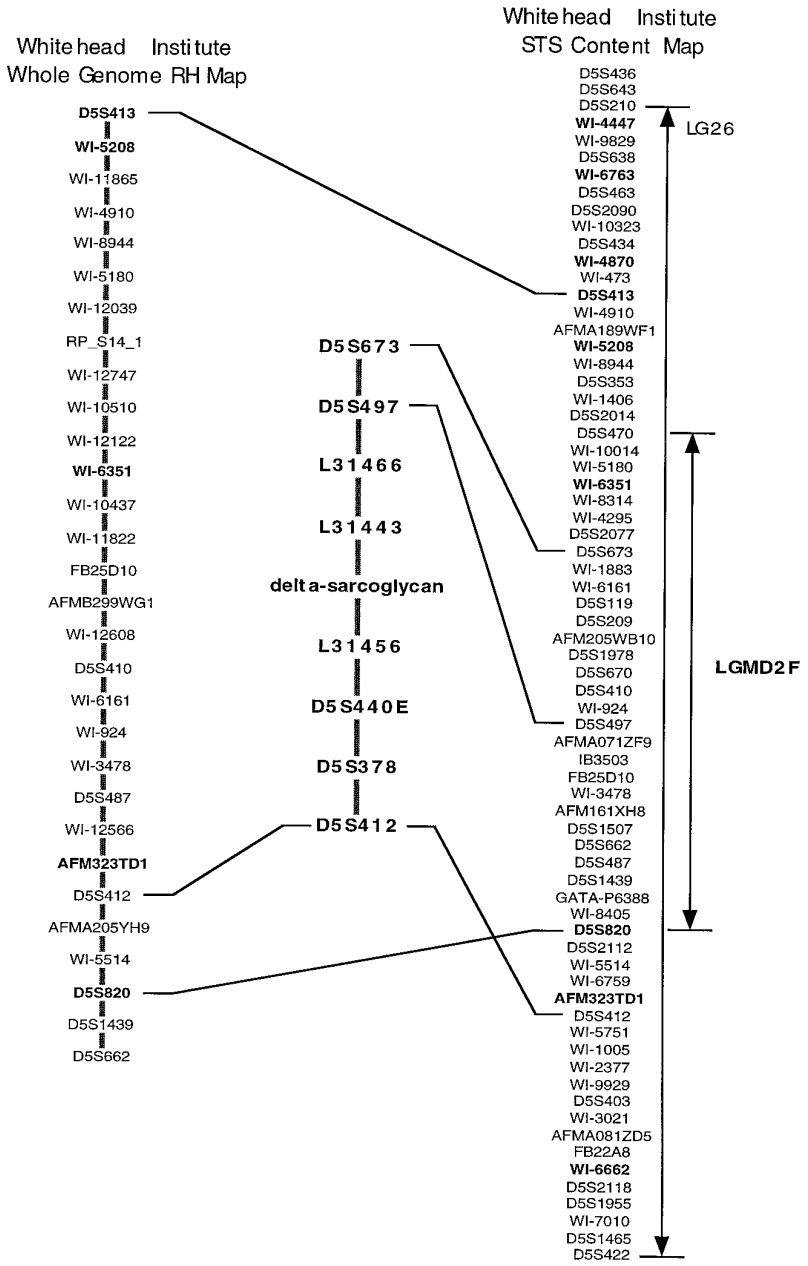

FIG. 5. **Chromosomal localization of the δ-sarcoglycan gene.** The radiation hybrid map generated for this study (*center*) is shown aligned with the relevant portions of the December 1995 release of the whole genome maps from the Whitehead Institute. The regions linked to LGMD2F in family LG26 (the minimal interval for LGMD2F) are indicated by the *vertical lines* on the *right* (37). The positioning of delta-sarcoglycan is described in the text. Selected markers that are on multiple maps are connected with *black lines*. Only markers in *bold text* on the Whitehead Institute RH map are ordered with an LOD score >2.5, which may account for the discrepancy between the two Whitehead Institute maps. All maps are drawn with the distal end at the bottom. Relative distances between markers are not shown, and the scale of each map is different.

TABLE II

Markers linked to the δ-sarcoglycan locus with LOD score >6

The order of the chromosome 5 loci was previously determined using the rhmaxlik program (50) as part of a chromosome 5 radiation hybrid map. Only markers that were linked to the δ-sarcoglycan locus with a LOD score >6 are shown. The order of these markers relative to the δ-sarcoglycan locus was then determined using the rhmaxlik program. Only a single maximum likelihood locus order was predicted with a log-likelihood of log(10) −153. The predicted order was: D5S673–D5S497–L31466–L31443–delta-sarcoglycan–L31456–D5S440E–D5S378–D5S412. Loci are listed in order from centromere to telomere. Information for loci beginning with "D" can be found in the GDB and those starting with "L" in the GSDB.

Locus	cR6000	LOD
D5S673	73.7	6.89
D5S497	48	11.27
L31466	43	11.19
L31443	37.8	13.29
L31456	20.8	19.90
D5S440E	26.1	17.70
D5S378	34.0	14.52
D5S412	51.3	10.82

with LGMD2C, -2D, or -2E demonstrated a reduction of all the sarcoglycan proteins (α, β, γ, and δ) at the sarcolemma, while the other components of the DGC were preserved. Therefore, a common feature of LGMD2C, -2D, and -2E is a specific absence of the sarcoglycan complex at the sarcolemma. This finding demonstrated that all sarcoglycans are required to maintain the integrity of the complex. We have previously demonstrated that in normal muscle the sarcoglycan proteins are tightly associated (24). This tight association may confer a high stability to the complex. The loss of only one sarcoglycan may render the complex more vulnerable to degradation and result in a destabilization of the complex during the process of its translocation to the sarcolemma. Such a mechanism has already been described for the cystic fibrosis transmembrane conductance regulator (52). The mechanism by which the absence of the sarcoglycan complex from the membrane leads to a dystrophic phenotype is currently unknown. It is possible that the sarcoglycan complex contributes to the DGC function, which is believed to link the extracellular matrix and the cy-

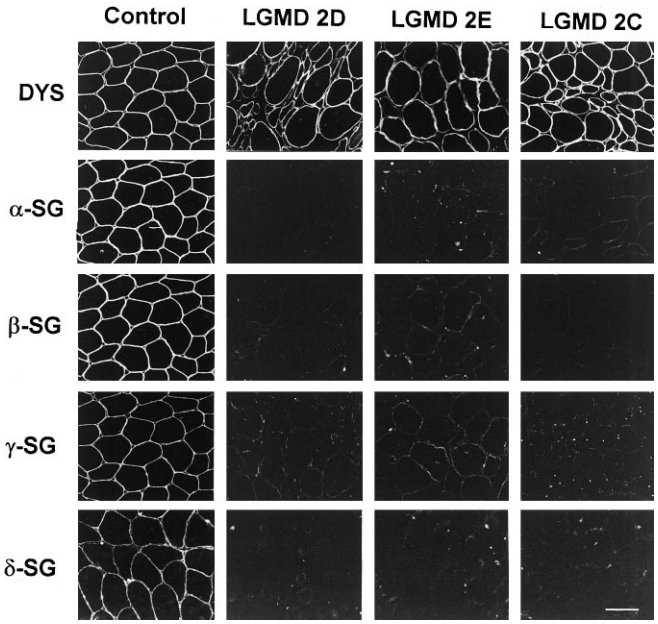

FIG. 6. **Specific deficiency of δ-sarcoglycan in LGMD2C, -2D, and -2E.** Immunofluorescence of human control (control), LGMD2C patient (LGMD2C), LGMD2D patient (LGMD2D), and LGMD2E patient (LGMD2E) deltoid muscles stained with affinity-purified anti-dystrophin (*DYS*), α-sarcoglycan (*α-SG*), β-sarcoglycan (*β-SG*), γ-sarcoglycan (*γ-SG*), and δ-sarcoglycan (*δ-SG*) antibodies are shown. *Bar*, 50 μM.

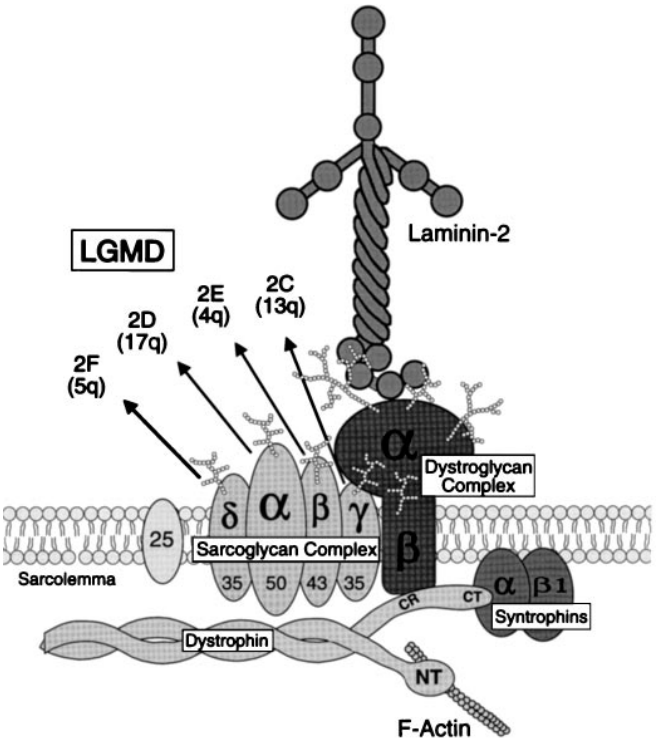

FIG. 7. **Schematic model of the dystrophin-glycoprotein complex.** The involvement of each sarcoglycan protein in LGMD is indicated by the *arrows*.

toskeleton. Absence of this functional link may result in sarcolemmal instability and greater susceptibility to stress-induced damage, as it has been suggested for Duchenne muscular dystrophy (26).

So far, LGMD2C, -2D, and -2E have been characterized by primary defects in γ-, α-, and β-sarcoglycan, respectively (19, 21–23, 38, 39), leading to the absence of the entire sarcoglycan complex. Since, the four sarcoglycan proteins are equal on a stoichiometric basis within the complex and that each protein is required to maintain the integrity of the complex, one may expect that a defect in δ-sarcoglycan also leads to the disruption of the sarcoglycan complex resulting in a LGMD phenotype. Interestingly, the δ-sarcoglycan gene has been mapped to chromosome 5q33–34, which overlaps with the region recently linked to a new form of autosomal recessive LGMD, LGMD2F (37).

Acknowledgments—We are grateful to Dr. Rita Shiang for helpful discussion and technical assistance. We also thank Drs. Rachelle Crosbie, Michael Henry, Leland E. Lim, and Steve Roberds for helpful comments on the manuscript and the University of Iowa DNA Core Facility for DNA sequencing.

Note Added in Proof—Recently, two papers have been published which report the cloning of δ-sarcoglycan (Nigro, V., Piluso, G., Belsito, A., Politano, L., Puca, A. A., Papparella, S., Rossi, E., Viglietto, G., Esposito, M. G., Abbondanza, C., Medici, N., Molinari, A. M., Nigro, G., and Puca, G. A. (1996) *Hum. Mol. Genet.* **5,** 1179–1186) and the demonstration that a mutation in δ-sarcoglycan causes LGMD2F (Nigro, V., de Sá Moreira, E., Piluso, G., Vainzof, M., Belsito, A., Politano, L., Puca, A. A., Passos-Bueno, M. R., and Zatz, M. (1996) *Nat. Genet.* **14,** 195–198). Interestingly, the reported sequence of δ-sarcoglycan differs in the C-terminal region from our δ-sarcoglycan sequence suggesting the existence of two isoforms of δ-sarcoglycan (δ_1 and δ_2) in skeletal muscle.

REFERENCES

1. Campbell, K. P. & Kahl, S. D. (1989) *Nature* **338,** 259–262
2. Ervasti, J. M., Ohlendieck, K., Kahl, S. D., Gaver, M. G. & Campbell, K. P. (1990) *Nature* **345,** 315–331
3. Yoshida, M. & Ozawa, E. (1990) *J. Biochem. (Tokyo)* **108,** 748–752
4. Ervasti, J. M. & Campbell, K. P. (1991) *Cell* **66,** 1121–1131
5. Hoffman, E. P., Brown, R. H., Jr. & Kunkel, L. M. (1987) *Cell* **51,** 919–928
6. Koenig, M., Monaco, A. P. & Kunkel, L. M. (1988) *Cell* **53,** 219–228
7. Hemmings, L., Kuhlman, P. A. & Critchley, D. R. (1992) *J. Cell Biol.* **116,** 1369–1380
8. Adams, M. E., Butler, M. H., Dwyer, T. M., Peters, M. F., Murnane, A. A. & Froehner, S. C. (1993) *Neuron* **11,** 531–540
9. Yang, B., Ibraghimov-Beskrovnaya, O., Moomaw, C. R., Slaughter, C. A. & Campbell, K. P. (1994) *J. Biol. Chem.* **269,** 6040–6044
10. Yang, B., Jung, D., Rafael, J. A., Chamberlain, J. S. & Campbell, K. P. (1995) *J. Biol. Chem.* **270,** 4975–4978
11. Ahn, A. H. & Kunkel, L. M. (1995) *J. Cell. Biol.* **128,** 363–371
12. Suzuki, A., Yoshida, M. & Ozawa, E. (1995) *J. Cell. Biol.* **128,** 373–381
13. Ibraghimov-Beskrovnaya, O., Ervasti, J. M., Leveille, C. J., Slaughter, C. A., Sernett, S. W. & Campbell, K. P. (1992) *Nature* **355,** 696–702
14. Ervasti, J. M. & Campbell, K. P. (1993) *J. Cell Biol.* **122,** 809–823
15. Gee, S. H., Blacher, R. W., Douville, P. J., Provost, P. R., Yurchenco, P. D. & Carbonetto, S. (1993) *J. Biol. Chem.* **268,** 14972–14980
16. Jung, D., Yang, B., Meyer, J., Chamberlain, J. A. & Campbell, K. P. (1995) *J. Biol. Chem.* **270,** 27305–27310
17. Suzuki, A., Yoshida, M., Hayashi, K., Mizuno. Y., Hagiwara, Y. & Ozawa, E. (1994) *Eur. J. Biochem.* **220,** 283–292
18. Roberds, S. L., Anderson, R. D., Ibraghimov-Beskrovnaya, O. & Campbell, K. P. (1993) *J. Biol. Chem.* **268,** 23739–23742
19. Roberds, S. L. *et al.* (1994) *Cell* **78,** 625–633
20. McNally, E. M., Yoshida, M., Mizuno, Y., Ozawa, E. & Kunkel, L. M. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 9690–9694
21. Lim, L. E., Duclos, F., & Broux, O. *et al.* (1995) *Nat. Genet.* **11,** 257–265
22. Bönnemann, C. G. *et al.* (1995) *Nature Genet.* **11,** 266–273
23. Noguchi, S. *et al.* (1995) *Science* **270,** 819–822
24. Jung, D. *et al.* (1996) *FEBS Lett.* **381,** 15–20
25. Yoshida, M., Suzuki, A., Yamamoto, H., Mizuno, Y. & Ozawa, E. (1994) *Eur. J. Biochem.* **222,** 1055–1061
26. Campbell, K. P. (1995) *Cell* **80,** 675–679
27. Helbling-Leclerc, A. *et al.* (1995) *Nat. Genet.* **11,** 216–218
28. Nissinen, M. *et al.* (1996) *Am. J. Hum. Genet.* **58,** 1177–1184
29. Stevenson, A. C. (1953) *Ann. Eugen.* **18,** 50–91
30. Walton, J. N. & Nattrass, F. J. (1954) *Brain* **77,** 169–231
31. Speer, M. C. *et al.* (1992) *Am. J. Hum. Genet.* **50,** 1211–1217
32. Yamaoka, L. Y. *et al.* (1994) *Neuromusc. Disord.* **4,** 471–475
33. Beckmann, J. S. *et al.* (1991) *C. R. Acad. Sci. (Paris)* **312,** 141–148
34. Bashir, R., Strachan, T., Keers, S., Stephenson, A., Mahjneh, I., Marconi, G., Nashef, L. & Bushby, K. M. (1994) *Hum. Mol. Genet.* **3,** 455–457
35. Ben Othmane, K. *et al.* (1992) *Nat. Genet.* **2,** 315–317
36. Azibi, K. *et al.* (1993) *Hum. Mol. Genet.* **2,** 1423–1428
37. Passos-Bueno, M. R., Moreira, E. S., Vainzof, M., Marie, S. K. & Zatz, M. (1996) *Hum. Mol. Genet.* **5,** 815–820
38. Matsumura, K., Tomé, F. M., Collin, H., Azibi, K., Chaouch, M., Kaplan, J.-C., Fardeau, M. & Campbell, K. P. (1992) *Nature* **359,** 320–322
39. Piccolo, F. *et al.* (1995) *Nat. Genet.* **10,** 243–245
40. Passos-Bueno, M. R. *et al.* (1995) *Hum. Mol. Genet.* **4,** 1163–1167

41. Richard, I. *et al.* (1995) *Cell* **81,** 27–40
42. Bento Soares, M., Bonaldo, M. De F., Jelene, P., Su, L., Lawton, L. & Efstratiadis, A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 9228–9232
43. Ohlendieck, K. & Campbell, K. P. (1991) *J. Cell. Biol.* **115,** 1685–1694
44. Smith, D. B. & Johnson, K. S. (1988) *Gene* (*Amst.*) **67,** 31–40
45. Smith, D. B. & Corcoran, L. M. (1987) in *Current Protocols in Molecular Biology* (Ausubel, F. M. *et al.*, eds) 16.7.1–16.7.8
46. Ervasti, J. M., Kahl, S. D. & Campbell, K. P. (1991) *J. Biol. Chem.* **266,** 9161–9165
47. Ohlendieck, K., Ervasti, J. M., Snook, J., B. & Campbell, K. P. (1991) *J. Cell. Biol.* **112,** 135–148
48. Laemmli, U. K. (1970) *Nature* **227,** 680–685
49. Towbin, H., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76,** 4350–4354
50. Boehnke, M., Lange, K. & Cox, D. R. (1991) *Am. J. Hum. Genet.* **49,** 1174–1188
51. Yamamoto, H., Hagiwara, Y., Mizuno, Y., Yoshida, M. & Ozawa, E. (1993) *J. Biochem.* (*Tokyo*) **114,** 132–139
52. Welsh, M. J. & Smith, A. E. (1993) *Cell* **73,** 1251–1254