Sarcospan, the 25-kDa Transmembrane Component of the Dystrophin-Glycoprotein Complex*

(Received for publication, August 7, 1997, and in revised form, September 24, 1997)

Rachelle H. Crosbie‡, Jim Heighway¶†, David P. Venzke‡, Jane C. Lee‡, and Kevin P. Campbell***

From the §Howard Hughes Medical Institute, Department of Physiology and Biophysics and the Department of Neurology, University of Iowa College of Medicine, Iowa City, Iowa 52242 and the ¶Paterson Institute for Cancer Research, Christie Hospital (NHS) Trust, Wilmslow Road, Manchester, M20 9BX, United Kingdom

The dystrophin-glycoprotein complex is a multisubunit protein complex that spans the sarcolemma and forms a link between the subsarcolemmal cytoskeleton and the extracellular matrix. Primary mutations in the genes encoding the proteins of this complex are associated with several forms of muscular dystrophy. Here we report the cloning and characterization of sarcospan, a unique 25-kDa member of this complex. Topology algorithms predict that sarcospan contains four transmembrane spanning helices with both N- and C-terminal domains located intracellularly. Phylogenetic analysis reveals that sarcospan's arrangement in the membrane as well as its primary sequence are similar to that of the tetraspan superfamily of proteins. Sarcospan co-localizes and co-purifies with the dystrophin-glycoprotein complex, demonstrating that it is an integral component of the complex. We also show that sarcospan expression is dramatically reduced in muscle from patients with Duchenne muscular dystrophy. This suggests that localization of sarcospan to the membrane is dependent on proper dystrophin expression. The gene encoding sarcospan maps to human chromosome 12p11.2, which falls within the genetic locus for congenital fibrosis of the extraocular muscle, an autosomal dominant muscular dystrophy.

In skeletal muscle fibers, the dystrophin-glycoprotein complex (DGC) is located at the sarcolemma and is composed of both peripheral and integral membrane proteins. Collectively, these proteins provide a physical connection between the extracellular matrix and the intracellular cytoskeleton of muscle cells. Disruption of this linkage eventually progresses to muscle cell necrosis, as evidenced by the dystrophic muscle phenotypes that result from defects in several of the DGC components (for review, see Refs. 6 and 7). Although the DGC is known to be essential for normal muscle function, the precise role of this multi-protein complex remains to be determined.

Purification of the DGC has led to the identification of many of its constituent polypeptides, which range in size from 25 kDa to over 400 kDa and include glycosylated as well as non-glycosylated proteins (1–4). Characterization of these proteins has increased our understanding of how the DGC is oriented in the sarcolemma and has provided clues to its function. In addition to dystrophin, the DGC consists of α/β-dystroglycan, the sarco- glycans (α, β, γ, and δ subunits), and the syntrophins. While most of the “dystrophin-associated proteins” (DAPs) have been identified, one of them, a 25-kDa protein (also called A5) (1–3, 5), has remained an enigma.

In the present work, we have determined the amino acid sequence of two peptides derived from 25DAP and have isolated the corresponding human cDNA. Previous analysis of the 25DAP indicated that it is an integral membrane protein, since it reacts strongly with a probe for protein hydrophobicity (4). We now report that 25DAP is a novel component of the DGC and is predicted to span the sarcolemma four times. This is unusual for integral membrane proteins of the DGC, which possess only a single transmembrane span. We have renamed 25DAP “sarcospan” in reference to its multiple sarcolemma-spanning domains. The predicted topology of sarcospan is similar to that of the tetraspan superfamily of proteins (8, 9), in which all members have four transmembrane domains and a large extracellular loop.

EXPERIMENTAL PROCEDURES

DGC Purification and Protein Sequencing—The DGC was extracted from rabbit skeletal muscle membranes (10) and then purified by SWGA affinity chromatography followed by ion exchange on a DEAE-column, as described previously (1, 2, 5). Purified DGC was electrophoresed on 3–15% gradient gels (0.7 mm thick) and stained with Coomassie Brilliant Blue as described by Ervasti et al. (5). The 25-kDa band was cut from the gel and sent to the Biopolymers Laboratory at the Massachusetts Institute of Technology for in-gel Lys-C digestion (11) and peptide microsequence analysis (12).

cDNA Isolation and Sequencing—Sarcospan cDNA clones were isolated by hybridization screening of a CLONTECH human adult skeletal muscle cDNA library (HL5002a) with a PCR-derived sarcospan cDNA probe encoding exons 1 to 3. The exon structure of the cDNA clones was confirmed by PCR, and the 5’ sequence was determined by direct analysis of biotinylated PCR products, as described previously (13). Sequencing analysis of the clones was performed using dye terminator cycling and analyzed on a 373A Stretch Fluorescent Automated sequencer (Applied Biosystems). The 5’ end of rabbit sarcospan was sequenced from a CLONTECH rabbit skeletal muscle cDNA library, as described for the human clone.

Computational Analysis—Multiple sequence alignment of sarcospan
Sarcospan, a Unique Component of the DGC

RESULTS AND DISCUSSION

To determine the identity of the 25DAP (1–3) (also called A5 spn1), proteins of the DGC were purified from rabbit skeletal muscle and separated using preparative SDS-PAGE. Peptide fragments of the 25DAP were obtained by Lys-C digestion. Microsequencing of the resultant peptides gave the following sequences: peptide I, KHRYQVFYVGV; peptide II, KDRQPRGQQRQGDAAGPDGPG. The amino acid sequence deduced from the human cDNA predicts a protein identical to the human sarcospan. Predictions of the topology and transmembrane helix orientation of this region demonstrates an exact match between the human and mouse sarcospan. As verification of this, partial sequencing of the 5’-terminal peptide coupled to bovine serum albumin was performed as described previously.

Immunofluorescence—Immunofluorescence on muscle biopsies from control and several DMD patients was accomplished as described except that muscle sections were incubated with affinity-purified rabbit anti-sarcospan antibody (1:10) at room temperature for 12 h. The sections were observed under a Bio-Rad MRC-600 laser scanning confocal microscope, and the digitized images were captured under identical conditions.

Sucrose Gradient Centrifugation—DGC fractions from the DEAE-column were pooled and concentrated to 800 μl. The samples were separated by centrifugation through a 5–20% sucrose gradient (2, 5). The gradients were fractionated into 1-ml fractions, and 50 μl of each fraction was resolved by 3–15% SDS-PAGE and either stained with silver or transferred to nitrocellulose. Immunoblot staining with antibodies to DGC (goat 20, 1:500) and sarcospan (rabbit 216 and 217; Knapp Creek Farms) at intramuscular and subcutaneous sites with 500 μg of an N-terminal sarcospan-2 synthetic peptide, CAADRQPRGQQRQGDAAGPD (Research Genetics), coupled to keyhole limpet hemocyanin in an emulsion of Freund’s complete adjuvant (Sigma). The amino acids Cys-Ala-Ala were added to the peptide for coupling purposes. Affinity purification of sarcospan antibodies was accomplished using Immobilon-P (Millipore) strips containing the N-terminal peptide coupled to bovine serum albumin. This antibody did not react with mouse tissue.

Immunofluorescence—Immunofluorescence on muscle biopsies from control and several DMD patients was accomplished as described except that muscle sections were incubated with affinity-purified rabbit anti-sarcospan antibody (1:10) at room temperature for 12 h. The sections were observed under a Bio-Rad MRC-600 laser scanning confocal microscope, and the digitized images were captured under identical conditions.

Sucrose Gradient Centrifugation—DGC fractions from the DEAE-column were pooled and concentrated to 800 μl. The samples were separated by centrifugation through a 5–20% sucrose gradient (2, 5). The gradients were fractionated into 1-ml fractions, and 50 μl of each fraction was resolved by 3–15% SDS-PAGE and either stained with silver or transferred to nitrocellulose. Immunoblot staining with antibodies to DGC (goat 20, 1:500) and sarcospan (rabbit 216 and 217, 1:20) was performed as described previously.

DAAGPDDPG and the amino acid sequence of peptide II (shown above). Furthermore, we confirmed sarcospan expression in rabbit skeletal and cardiac muscle by Northern analysis with a full-length cDNA probe (data not shown).

The primary amino acid sequence deduced from the human cDNA predicts a protein with four transmembrane spanning domains and a topology similar to that of the tetranspan proteins (Fig. 1B). Based on the predicted membrane topology of the human protein and its residence in the sarcolemma, we have named this protein sarcospan. The tetranspans, also referred to as the transmembrane 4 superfamily (TM4SF) and the tetratraspans, have four transmembrane spanning helices that are conserved among family members (8, 9). Although there is no clear function for the TM4SF members, many of them are associated with integrins and have been implicated in the regulation of cell proliferation (9). Using phylogenetic analysis, we find that sarcospan is more closely related to the divergent family members Rom-1, peripherin, and uroplakin (data not shown). We refer to the two cloned cDNAs as sarcospan-1 (spn1; GenBankTM accession number X89105) and sarcospan-2 (spn2; GenBankTM accession number AF016028), where Spn2 includes 26 amino acids at its N terminus. In light of the identification of sarcospan, we have redesignated KRAG (13) as the spn1 gene.

To examine the tissue distribution of sarcospan, we performed RNA hybridization analysis. We probed human multiple tissue Northern blots with a full-length spn1 cDNA probe. As shown in Fig. 2, a 6.5-kb transcript is present exclusively in skeletal and cardiac muscles. However, a 4.5-kb transcript is also found in skeletal and cardiac muscle, as well as in the thymus, prostate, testis, ovary, small intestine, colon, and spleen. Our data suggest the presence of tissue-specific sarcospan transcripts. The expression of sarcospan in a broad array of tissues is most similar to that of dystroglycan (18, 19), suggesting that these proteins play important roles in muscle as well as non-muscle tissues.
Molecular size markers are indicated on the left. We confirmed that the 25-kDa Coomassie-stained band in the DGC was sarcospan by staining immunoblots containing purified DGC with sarcospan antibodies. Although a 25-kDa band is present in purified DGC from rabbit skeletal muscle, it does not elicit an antibody response when DGC preparations were used to immunize either sheep or goat (Fig. 3A). Thus, we generated a polyclonal antibody to sarcospan by immunizing rabbits with peptide II. This affinity-purified antibody specifically stains a 25-kDa band on immunoblots of purified DGC from rabbit skeletal muscle membranes (Fig. 3A).

As a first demonstration that sarcospan is an integral component of the DGC, we show that sarcospan enriches in purification of the complex. Immunoblots of purified DGC stained with DGC antibodies show a clear enrichment of these purified proteins relative to the levels in rabbit skeletal muscle membranes (Fig. 3A). Identical immunoblots illustrate that sarcospan also enriches in DGC prepared from rabbit membranes (Fig. 3A).

As another test for association between sarcospan and the DGC, we examined sarcospan expression in muscle from DMD patients. We examined several DMD patients with primary mutations in dystrophin, resulting in loss of the entire dystrophin protein. Indirect immunofluorescence assays with antibodies against dystrophin's N and C termini and the central rod domain verify that dystrophin is completely absent from the DMD sarcolemma (data not shown). Without dystrophin, the other members of the complex are reduced, perhaps the result of premature protein degradation, improper assembly of the complex, or aberrant transportation to the sarcolemma. We show that sarcospan is present at the sarcolemma in normal muscle and is dramatically reduced in DMD muscle (Fig. 3B). The reduced staining of sarcospan in DMD muscle provides further evidence that this protein is complexed with the DGC.

The tight association of sarcospan with the DGC is illustrated by centrifugation of the DGC through sucrose gradients. Sucrose gradient sedimentation separates the DGC from any proteins which might bind with low affinity to the complex. Proteins from the sucrose gradient fractions were separated by SDS-PAGE. The resultant polyacrylamide gels were stained with silver to better visualize sarcospan (Fig. 4), which stains weakly with Coomassie Brilliant Blue (Fig. 3A). Immunoblotting with sarcospan antibodies confirms that the intensely silver-stained band at 25 kDa is sarcospan. The peak of DGC proteins migrates in fractions 9 and 10 as seen by the silver-stained gels and Western blotting of these same fractions with anti-DGC antibodies (Fig. 4). Sarcospan migrates in the same fractions as the DGC during sedimentation through sucrose gradients, confirming that sarcospan is an integral member of this complex (Fig. 4).

The structural connection between the extracellular matrix and the intracellular actin network is dependent on the integrity of the DGC. Disruption of this linkage eventually progresses to muscle cell necrosis, as evidenced by the dystrophic phenotype that results from defects in several of the DGC components, including the dystrophin and the sarcoglycans (for review, see Refs. 6 and 7). Mutations in sarcospan would feasibly also give rise to a dystrophic phenotype. The gene encoding sarcospan maps to human chromosome 12p11.2 (13). Congenital fibrosis of the extraocular muscle (CFEOM) is an autosomal dominant disorder which primarily affects the ocular muscles, rendering patients with little or no eye movement. Linkage analysis of two large unrelated families with CFEOM indicate that this disease locus lies within 12p11.2-q12 (20, 21).
We propose that sarcospan is a prime candidate gene for this disease. Sarcospan is the first identified member of the DGC that spans the membrane more than once. In fact, over 60% of sarcospan’s amino acids are predicted to be within the membrane and this unique characteristic may be providing clues to sarcospan’s function. For instance, sarcospan’s transmembrane domains are expected to hold this protein firmly within the lipid bilayer, in which case it could provide a solid anchorage for the rest of the DGC. Additionally, the multiple transmembrane regions of sarcospan might form a pore in the sarcolemma and thereby serve as a membrane channel. The latter scenario is particularly attractive in light of the emerging concept that the DGC performs functional as well as structural roles in muscle. The charged residues in sarcospan’s transmembrane domains may be important for protein-protein interactions, perhaps with integrins or other DGC components. Finally, the observation that sarcospan expression is amplified in some human tumors, in combination with the proposal that tetraspan proteins play a role in cell growth, supports the idea that sarcospan itself has essential functions that extend beyond its role in muscle.

Acknowledgments—We are indebted to Richard Cook at the Biopolymers Facility at MIT. We thank the University of Iowa Diabetes and Endocrinology Research Center, funded through National Institutes of Health Grant DK25285, and the University of Iowa DNA Sequencing Core Facility.

REFERENCES

1. Campbell, K. P., and Kahl, S. D. (1989) Nature 338, 259–262
2. Ervasti, J. M., Ohlendieck, K., Kahl, S. D., Gaver, M. G., and Campbell, K. P. (1990) Nature 345, 315–319
3. Yoshida, M., and Ozawa, E. (1990) J. Biochem. (Tokyo) 108, 748–752
4. Ervasti, J. M., and Campbell, K. P. (1991) Cell 66, 1121–131
5. Ervasti, J. M., Kahl, S. D., and Campbell, K. P. (1991) J. Biol. Chem. 266, 9161–9165
6. Campbell, K. P. (1995) Cell 80, 675–679
7. Straub, V., and Campbell, K. P. (1997) Curr. Opin. Neurol. 10, 168–175
8. Wright, M. D., and Tomlinson, M. G. (1994) Immunol. Today 15, 588–594
9. Määcker, H. T., Todd, S. C., and Levy, S. (1997) FASEB J. 11, 428–442
10. Ohlendieck, K., Ervasti, J. M., Snook, J. B., and Campbell, K. P. (1991) J. Cell Biol. 112, 135–148
11. Williams, K. L., and Stone, K. L. (1997) Techniques in Protein Chemistry VI (Crabb, J., ed) Academic Press, New York, in press
12. Stone, K. L., and Williams, K. R. (1993) in A Practical Guide to Protein and Peptide Purification for Microsequencing (Matsudaira, P. T., ed) 2nd Ed., pp. 43–69, Academic Press, New York
13. Heighway, J., Betticher, D. C., Hoban, P. R., Altermatt, H. J., and Cowen, R. (1996) Genomics 35, 207–214
14. Hofmann, K., and Stoffel, W. (1993) Biol. Chem. 374, 166
15. Jung, D., Dukowski, F., Apostol, B., Straub, V., Lee, J. C., Allaamand, V., Venzke, D. P., Sunada, Y., Moosmayer, C. R., Leveille, C. J., Slaughter, C. A., Crawford, T. O., McPherson, J. D., and Campbell, K. P. (1996) J. Biol. Chem. 271, 32321–32329
16. Altshul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
17. Scott, A. F., Elizaga, A., Morrell, J., Bergen, A., and Penno, M. B. (1994) Genomics 20, 227–230
18. Ibraghimov-Beskrovnaya, O., Ervasti, J. M., Leveille, C. J., Slaughter, C. A., Sernett, S. W., and Campbell, K. P. (1992) Nature 355, 696–702
19. Ibraghimov-Beskrovnaya, O., Milatovich, A., Ozcelik, T., Yang, B., Koepnick, K., Francke, U., and Campbell, K. P. (1993) Hum. Mol. Genet. 2, 1651–1657
20. Engle, E. C., Kunkel, L. M., Specht, L. A., and Beggs, A. H. (1994) Nat. Genet. 7, 69–73
21. Engle, E. C., Marondel, I., Houthman, W. A., de Vries, B., Loewenstein, A., Lazar, M., Ward, D. C., Kucherlapati, R., and Beggs, A. H. (1995) Am. J. Hum. Genet. 57, 1086–1094