Human Molecular Genetics, 1993, Vol. 2, No. 10 1651–1657

Human dystroglycan: skeletal muscle cDNA, genomic structure, origin of tissue specific isoforms and chromosomal localization

Oxana Ibraghimov-Beskrovnaya, Athena Milatovich[1], Tayfun Ozcelik[2], Bin Yang, Kevin Koepnick, Uta Francke[1,2] and Kevin P.Campbell*

Howard Hughes Medical Institute and Department of Physiology and Biophysics, University of Iowa College of Medicine, Iowa City, IA 52242, [1]Department of Genetics and Pediatrics and [2]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA

Received June 1, 1993; Revised and Accepted July 19, 1993

GenBank accession no. L19711

Dystroglycan is a novel laminin binding component of the dystrophin-glycoprotein complex which provides a linkage between the subsarcolemmal cytoskeleton and the extracellular matrix. Here we report the cDNA and genomic structure of human dystroglycan. The human dystroglycan is encoded by a single gene (DAG1) mapped to chromosome 3 band p21. The coding sequence is organized into two exons, separated by a large intron. The predicted amino acid sequence of human and rabbit dystroglycan are 93% identical with predicted glycosylation sites being conserved. Human dystroglycan is expressed in a variety of fetal and adult tissues. Our data suggest that muscle and non-muscle isoforms of dystroglycan differ by carbohydrate moieties but not protein sequence. Therefore, we hypothesize that variable glycosylation of the conserved protein core might modulate laminin binding. The relationship of dystroglycan to human diseases is discussed.

INTRODUCTION

Duchenne muscular dystrophy (DMD) is a progressive muscle disorder which results from rearrangements of the DMD gene, encoding the large cytoskeletal protein dystrophin (1,2). Dystrophin is localized to the sarcolemma (3–5) where it is tightly associated with several membrane glycoproteins (6). Biochemical characterization of the dystrophin-glycoprotein complex has demonstrated that four transmembrane proteins provide a linkage between dystrophin and a highly glycosylated 156 kDa glycoprotein which interacts with laminin, a major component of the extracellular matrix (7).

We have previously shown that the 43 kDa dystrophin-associated transmembrane glycoprotein and the extracellular 156 kDa component of rabbit skeletal muscle are posttranslational products of a single 97 kDa precursor protein, named dystroglycan (7). Dystroglycan is a novel laminin receptor, suggesting that one function of dystrophin-glycoprotein complex is to provide a linkage between the subsarcolemmal cytoskeleton and the extracellular matrix in skeletal muscle. We have demonstrated that dystroglycan is expressed not only in muscle tissues, but in non-muscle tissues as well, suggesting it may play an important role in maintaining membrane integrity in different cell types. Recently, the expression and subcellular distribution of the dystrophin-glycoprotein complex and laminin were examined in cardiac muscle (8). It was found that cardiac muscle contains proteins antigenically similar to all of the skeletal muscle dystrophin-associated proteins and that these proteins codistribute with laminin in papillary myofibers and Purkinje fibers. However, the 156 kDa component of dystroglycan in cardiac muscle was lower in molecular weight. Since the extracellular 156 kDa dystroglycan component differs in molecular weight among various tissues (7), we have named the extracellular component 'α-dystroglycan' and the transmembrane component 'β-dystroglycan' to avoid confusion.

The loss of the linkage between the subsarcolemmal cytoskeleton and the extracellular matrix due to the absence of dystrophin in DMD patients and mdx mice was demonstrated by the great reduction of all dystrophin-associated proteins (9,10). The importance of the integrity of all components of the complex in skeletal muscle was further demonstrated by analysis of biopsies of patients with a form of autosomal recessive muscular dystrophy (SCARMD), in which the absence of the 50 kDa dystrophin-associated glycoprotein apparently causes the dysfunction of the complex, despite the presence of dystrophin (11). Since alterations in dystroglycan could also lead to an abnormal expression of the complex which, in turn, could result in muscle cell necrosis, it was important to investigate the structural and functional properties of human dystroglycan.

To better define the domain structure of dystroglycan and to understand the mechanism underlying the synthesis of human dystroglycan isoforms in different tissues, we isolated and characterized the human dystroglycan gene and cDNA. The dystroglycan gene (DAG1) was localized to human chromosome 3p21. The chromosomal location of dystroglycan together with the availability of the cDNA and gene sequence will allow the evaluation of this locus as a candidate gene in various neuromuscular diseases.

* To whom correspondence should be addressed at: Howard Hughes Medical Institute, University of Iowa College of Medicine, 400 EMRB, Iowa City, IA 52242, USA

RESULTS

Cloning and analysis of human dystroglycan cDNA

In order to isolate a human dystroglycan skeletal muscle cDNA, a human skeletal muscle cDNA λgt10 library was screened with ^{32}P-labeled rabbit β-dystroglycan cDNA. Approximately 10^6 recombinant phage were analyzed and two independent clones were obtained. One clone, HD-1, is represented in Fig. 1a. By several rounds of screening the cDNA library with human-specific probes corresponding to the 5′ or 3′ end of HD-1, several additional clones were isolated. Four overlapping clones covering the entire mRNA were completely sequenced (Fig. 1a). The full-length human cDNA consists of 5510 nucleotides, of which 2685 nucleotides represent an open reading frame. A polyadenylation sequence and poly(A) tail were also identified. The deduced amino acid sequence predicts a polypeptide of a calculated Mr of 97,552 with a signal sequence of 27 amino acids, a single transmembrane domain close to the C-terminal region, four potential N-glycosylation sites and many potential sites for O-glycosylation.

Conservation of dystroglycan between species

Alignment of amino acid sequences for human and rabbit dystroglycan are shown in Fig. 1b. Both proteins contain 895 amino acids with overall sequence identity of 93%. Ninety percent of the amino acid substitutions are conservative. The transmembrane domain of human dystroglycan is identical to that of rabbit dystroglycan. The intracellular C-terminal region of human and rabbit dystroglycan is highly conserved and is enriched in proline (23%). Both proteins have identical consensus sites for N-glycosylation and have high content of threonine and serine as potential sites for O-glycosylation. It is interesting to note that the most probable candidate site for attachment of glycosaminoglycan chains, a serine-glycine followed by a negatively charged amino acid is conserved at position 90 (Fig. 1b). Another Ser-Gly dipeptide sequence at position 485 is also conserved. Ser-Gly at position 398 is present only in the rabbit sequence. High homology between rabbit and human dystroglycan suggests its functional importance, especially in terms of carbohydrate chain attachment sites, since carbohydrates may play an important role in laminin binding (Ervasti and Campbell, unpublished results).

To characterize further the conservation of dystroglycan among mammalian species, Southern hybridization was performed on DNA isolated from a variety of species using the human HD-2 cDNA and high stringency conditions (Fig. 1c). The same conditions were used for all Southern blots as described in the methods section. Strong hybridization signals of similar intensities were detected on all DNA preparations.

Dystroglycan is encoded by a single gene

The human gene for dystroglycan was characterized by genomic Southern analysis. DNA isolated from lymphocytes of peripheral

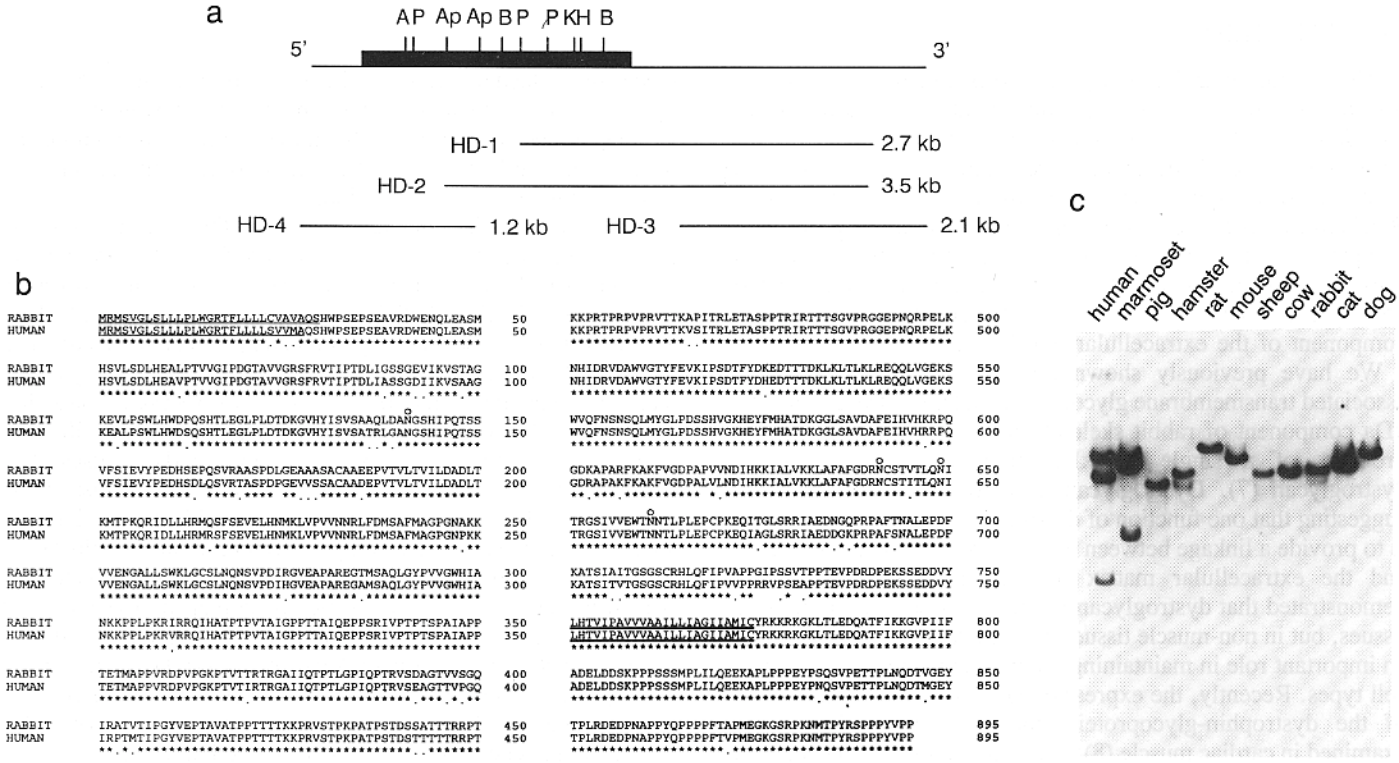

Figure 1. Cloning and primary structure analysis of human dystroglycan. **a.** Restriction map and overlapping cDNA clones to human dystroglycan. The protein coding region is indicated by a solid box. Restriction sites are indicated: A-AccI, Ap-ApaI, B-BamHI, H-HindIII, K-KpnI, P-PstI. **b.** Alignment of human and rabbit dystroglycan. Both human and rabbit polypeptides consist of 895 amino acids and share 93% identity. Identical amino acids are indicated by asterisks and conservative amino acids by dots. The putative signal peptide is underlined and the single transmembrane domain is boldly underlined. Potential N-linked glycosylation sites are indicated by an open circle. The nucleotide sequence of human dystroglycan will appear in GENBANK accession number L19711. **c.** Conservation of dystroglycan in mammalian group. Each lane contains 8 μg of EcoRI-digested DNA from the various mammals (BIOS Corp.) indicated above. The blot was probed with ^{32}P-labeled human cDNA HD-2.

blood from two individuals (DNA1 and DNA2) was digested with EcoRI, HindIII, KpnI and PstI and transferred on to nitrocellulose. The blot was probed with ^{32}P-labeled HD-2 cDNA (Fig. 2). The patterns of hybridization with DNA1 and DNA2 digested with HindIII, KpnI, PstI were identical. Hybridization with EcoRI digested samples showed one 12 kb EcoRI band in DNA1, but two EcoRI fragments (9 kb and 3 kb) in DNA2, suggesting EcoRI polymorphism. The simple predictable hybridization pattern suggests the existence of a single gene for human dystroglycan.

Genomic organization of human dystroglycan

Two independent libraries were constructed from partially EcoRI-digested human DNA1 and DNA2 using the EMBL3 vector. Approximately 2×10^6 independent clones were obtained. Libraries were screened with ^{32}P-labeled cDNA HD-1. Several hybridizing clones were isolated and characterized. Two clones, G1 and G2 originating from DNA1 and DNA2 respectively, were chosen for further analysis. Restriction analysis demonstrated several EcoRI fragments within each recombinant phage, including 12 kb in G1 and 9 kb in G2 which were detected previously by Southern blot hybridization. Southern blot analysis using HD-2 cDNA as a probe identified these fragments specifically. The 12 kb and 9 kb EcoRI inserts were subcloned into pUC19 for further analysis. A physical map of these fragments was generated as shown in Fig. 3. Based on this map, the 12 kb EcoRI fragment was digested into small subclones and partially sequenced. The combination of sequencing, hybridization and PCR analysis of cloned genomic fragments and comparison with cDNA sequence allowed us to propose the genomic structure of the 12 kb EcoRI-fragment (Fig. 3). Interestingly, the majority of coding sequence of human dystroglycan (2.4 kb out of 2.7 kb) together with part of 3′-untranslated region is organized in one large exon, which starts at nucleotide 286 of the cDNA sequence. The rest of the 3′-untranslated region is organized in other exons and needs further characterization. Upstream of this exon a large intron was identified which spans more than 6 kb of genomic sequence. The polymorphic EcoRI site was mapped downstream of the coding region (Fig. 3). To isolate a genomic DNA fragment containing exon sequence encoding the initiator methionine, the genomic library constructed from DNA1 was screened with ^{32}P-labeled cDNA representing the first 280 nucleotides of coding sequence as well as the 5′ untranslated region. Southern blot analysis of EcoRI-digested DNA1 and DNA2 with the above probe showed specific hybridization to a 6.5 kb EcoRI-fragment. This 6.5 kb genomic EcoRI-fragment (clone G3) was isolated, mapped and partially sequenced. The exon, encoding part of the 5′-untranslated region and 285 nucleotides of coding sequence, including the initiating methionine codon was identified (Fig. 3). The exon(s) containing the rest of the 5′-untranslated region, transcription initiation site and upstream region were not analyzed. Clones G1 and G3 do not overlap in their intronic sequence, suggesting that the intron is more than 10 kb in size. Thus, the human dystroglycan coding sequence is organized into two exons, a small first exon separated from a large second exon by a large intron (Fig. 3).

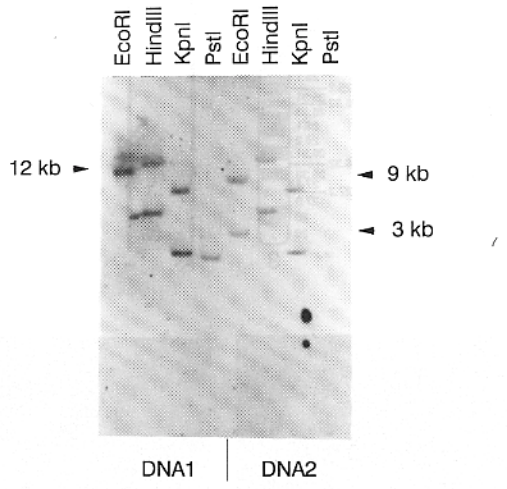

Figure 2. Human genomic Southern hybridization. 8 μg of DNA was digested to completion with indicated restriction endonucleases. The first four lanes correspond to DNA1 and the last four lanes correspond to DNA2 as indicated. The filter was probed with ^{32}P-labeled HD-2 cDNA. Indicated by an arrows are the 12 kb EcoRI fragment of DNA1 and two fragments of 9 kb and 3 kb of DNA2 showing polymorphism characterized by additional EcoRI site.

Tissue-specific expression of human dystroglycan

The tissue-specific expression of dystroglycan was examined by northern blot analysis. Radioactively labeled cDNA HD-2, which is specific to exon 2, was hybridized to mRNA from human adult

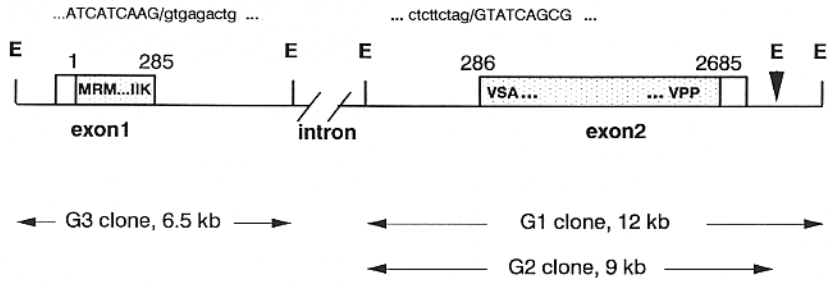

Figure 3. Diagram of the genomic structure of human dystroglycan. The schematic structure of the dystroglycan gene is not drawn to scale. The gene is represented by the line, interrupted at the location of the intron. Coding sequences within exons are shown by the shaded boxes, corresponding to the nucleotides of the cDNA sequence shown above. The amino acid sequences within each exon are indicated. The nucleotide sequences at exon/intron junctions are presented. The nucleotides in uppercase are exon sequence while the nucleotides in lowercase are intron sequence. A polymorphic EcoRI site, which is present within clone G2 but not G1 is indicated by the arrowhead.

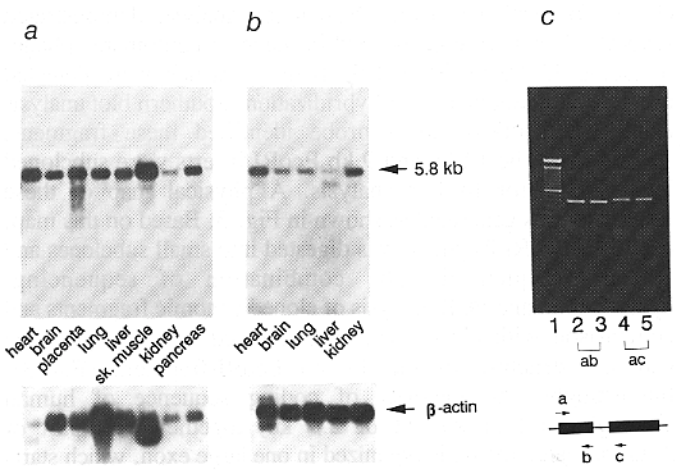

Figure 4. Tissue specific expression of human dystroglycan. **a**. Northern blot analysis of human adult tissues. Each lane contains 2 μg of human poly(A)$^+$ RNA from tissues indicated. RNA was probed with ^{32}P-labeled cDNA HD-2 (exon2-specific). **b**. Northern blot analysis of human fetal tissues. Each lane contains 2 μg of human fetal poly(A)$^+$RNA from tissues indicated. Exon-1 specific probe from HD-4 cDNA was used for hybridization. 5.8 kb mRNA is indicated by an arrow. The RNA in each lane was shown to be intact by probing the blots with a ^{32}P-labeled cDNA probe for human β-actin as indicated. **c**. RT-PCR analysis of tissue specific dystroglycan isoforms. lane 1—100 b.p. ladder; lane 2 and 3, RT-PCR amplification products from brain and heart RNA respectively using exon1 specific primers *a* and *b* (primer *a*: 5′- GGAGCAGG-TGTGCAGAGGG -3′, primer *b*: 5′-CTTGATGATATCTCCACTGGAGGC-3′); lane 4 and 5, amplification products from brain and heart RNA respectively using exon 1 and exon 2 specific primers *a* and *c* (primer *c*: 5′-AACATGGCAAAG-CCTCCTTCCCTG-3′). A faint larger band in lanes 2 and 4 can be attributed to the nonspecific amplification from brain RNA since the same band appears in control brain RT-PCR amplification using primer *a* alone (not shown). Schematic representation of primers is shown.

muscle and non-muscle tissues as indicated in Fig. 4a. A band of 5.8 kb was detected in all tissues examined. The mRNA is most abundant in skeletal muscle and heart and less abundant in non-muscle tissues. mRNA from several fetal tissues was probed with exon 1 sequence (Fig. 4b). Interestingly, the same size dystroglycan specific transcript was detected in all tissues examined, which demonstrates that dystroglycan is expressed in fetal and adult muscle and non-muscle tissues. Since the 5.8 kb band is detected in all tissues examined by using probes specific for exon 1 or exon 2, we conclude that dystroglycan transcripts are identical in these tissues. In addition, RT-PCR was used to amplify skeletal-muscle specific exon 1 from adult brain and cardiac RNA using two sets of primers (Fig. 4c). The first set contained a sense primer upstream of ATG codon and an antisense primer corresponding to the end of exon 1 (Fig. 4c, lanes 2 and 3). The second set contained the same sense primer and antisense primer, corresponding to the beginning of the exon 2 (Fig. 4c lanes 4 and 5). The identical amplification products of predicted size based on the skeletal muscle cDNA were detected in brain and heart for each set of primers (Fig. 4c). This further demonstrates that tissue specific isoforms do not differ by the primary structure.

Chromosomal localization of human dystroglycan

The chromosomal localization of the dystroglycan gene was first determined by Southern blot analysis of a panel of human/hamster cell hybrids using a radiolabeled dystroglycan cDNA probe. A

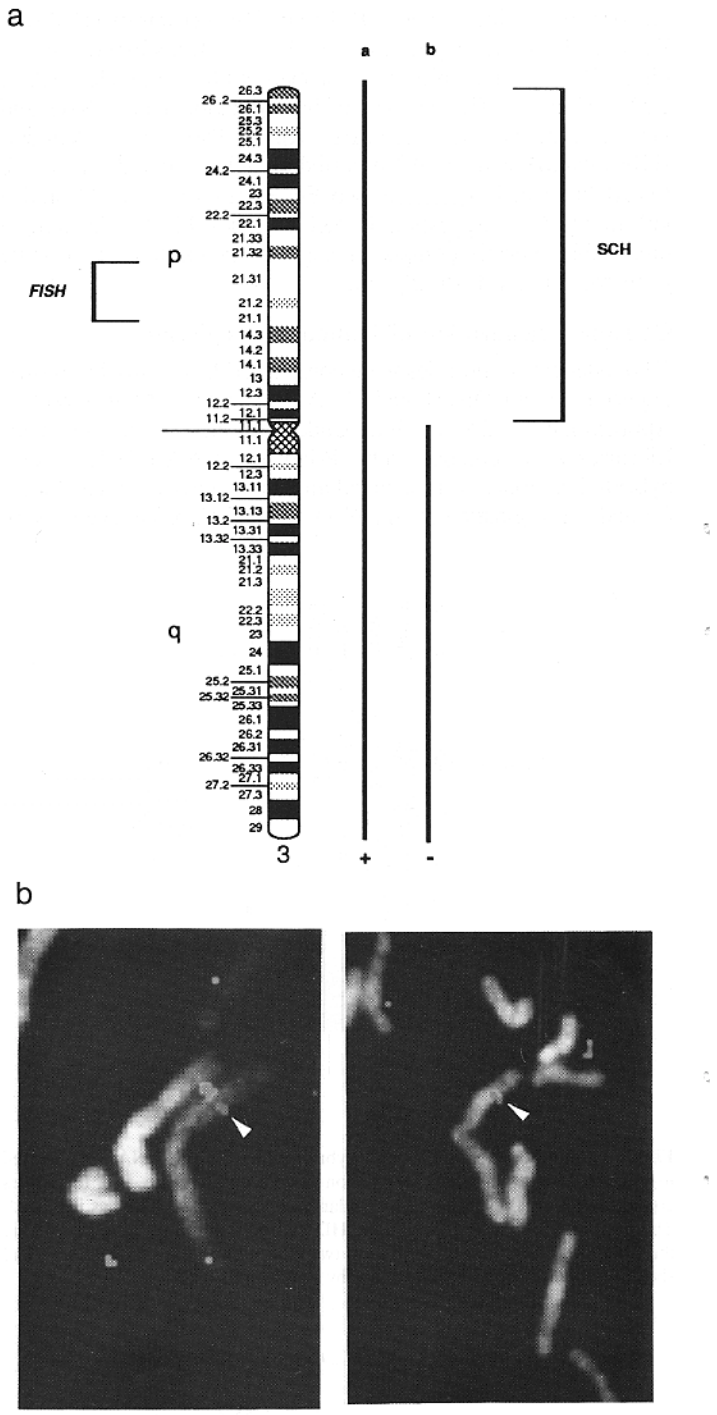

Figure 5. Fluorescence in situ hybridization. **a.** An 850 band level ideogram of a G-banded human chromosome 3. Vertical lines labeled 'a' and 'b' represent somatic cell hybrids. Line 'a' represents the human×Chinese hamster somatic cell hybrid line which retained an intact copy of human chromosome 3. The human-specific fragment was detected in this hybrid denoted by a '+'. Line 'b' represents the human×Chinese hamster somatic cell hybrid, which contained an isochromosome for the long arm of human chromosome 3 (i(3q)). The human-specific fragment was not detected in this hybrid, denoted by '−'. The right-hand bracket labeled 'SCH' represents the localization of the human DAG1 gene to the short arm of human chromosome 3 (3pter−cen) using somatic cell hybrid analysis. The left-hand bracket labeled 'FISH' represents the localization of DAG1 gene to human chromosome 3p21.1−p21.31 using fluorescence chromosomal in situ hybridization. **b.** Partial representative metaphase spreads with specific hybridization signals on human chromosome 3 at approximately 3p21.1−p21.31. Arrowheads indicate the hybridization signals.

Table 1. Comparison of dystroglycan sequences and human chromosomes in human×chinese hamster somatic cell hybrids

Dystroglycan sequence/ Chromosome	Human chromosome																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
Concordant hybrids																							
+/+	1	3	7	2	1	3	1	2	1	1	5	4	3	5	4	3	0	3	3	4	5	3	2
−/−	2	3	3	2	2	2	2	1	3	4	1	3	2	1	1	3	1	2	0	2	1	1	1
Discordant hybrids																							
+/−	5	4	0	4	6	3	6	5	5	6	2	3	4	1	3	3	7	4	4	3	2	3	1
−/+	1	1	0	2	2	1	1	3	1	0	2	1	1	3	3	1	3	2	4	2	2	3	3
Total discordant hybrids	6	5	0	6	8	4	7	8	6	6	4	4	5	4	6	4	10	6	8	5	4	6	4
Total informative hybrids*	9	11	10	10	11	9	10	11	10	11	10	11	10	10	11	10	11	11	11	11	10	10	7
Percent discordant hybrids	67	45	0	60	73	44	70	73	60	55	40	36	50	40	55	40	91	55	73	45	40	60	57

*Chromosomes with rearrangements or present at a frequency of 0.1 or less were excluded.

10 kb Bgl II human specific fragment was detected. This fragment was seen in the normal human control and in all hybrids which retained an intact copy of human chromosome 3. All other chromosomes were excluded by at least four discordant hybrids (Table 1). One hybrid in the mapping panel contained an isochromosome of the long arm of human chromosome 3 [i(3q)] which spontaneously arose in the subcloning of this hybrid. The 10 kb human-specific Bgl II fragment was not detected in this hybrid. Therefore, the dystroglycan gene locus was localized to the short arm of chromosome 3 (Fig. 5a).

This regional assignment was confirmed and further defined by fluorescence in situ hybridization (FISH) (Fig. 5b). A plasmid containing a genomic 9 kb EcoRI-fragment (G2 clone) was biotin-11-dUTP labeled and hybridized to human metaphase chromosomes. In 11 of 17 metaphase spreads examined, specific signals were observed on at least one chromosome 3 homolog at band p21. Seven of these 11 metaphases with specific signals had both chromosome 3 homologs labeled, and no other chromosomes had specific signals. The somatic cell hybrid analysis and in situ hybridization discovered only one site of specific hybridization, which further supports the conclusion that DAG1 is a single copy gene.

DISCUSSION

We present here the cloning and sequence analysis of the cDNA for human dystroglycan, a component of the dystrophin-glycoprotein complex. The cDNA encodes a precursor protein of 97 kDa which is posttranslationally processed into α-dystroglycan (156 kDa dystrophin-associated glycoprotein) and β-dystroglycan (43 kDa dystrophin-associated glycoprotein). Dystroglycan is encoded by a single gene (DAG1). The gene has been isolated and its structure determined. A polymorphic EcoRI-site was identified within the dystroglycan gene and mapped downstream of the coding region. The coding region is organized into two exons, the first consists of 285 b.p. of coding sequence and the second incorporates the rest of coding sequence of 2400 b.p. These exons are separated by a single large intron. We could not estimate the exact length of the intron, since neither a continuous fragment of DNA with both exons nor overlapping clones were available. Attempts to PCR amplify the intron were not successful most likely due to its large size. Similar genomic structure was described for the human α_{1B}-adrenergic receptor (12). It consists of two exons and a single large intron, which is different from the genes for other G-protein-coupled receptors, which are intronless (13,14).

Comparison of the amino acid sequence for human and rabbit dystroglycan showed 93% identity and with the majority of amino acid substitutions being conservative. Interestingly, all four N-glycosylation sites are identical and a high content of threonine and serine is found in both proteins, which represent potential O-glycosylation sites. We had proposed that rabbit dystroglycan may be a proteoglycan based on the fact that sugar moieties constitute 2/3 of its molecular weight (15). The consensus site for glycosaminoglycan chain attachment consists of the dipeptide serine-glycine, most often followed by a negatively charged amino acid (16,17). Two out of three such sites in rabbit dystroglycan sequence are found in the human sequence as well, namely at position 90 (ser-gly-negatively charged amino acid) and at position 485. Unfortunately, our attempts to demonstrate the presence of glycosaminoglycan chains biochemically by using GAG lyases were unsuccessful, most probably because the sugar composition is very complex.

Similarly to the previously characterized rabbit dystroglycan, the mRNA for the human isoform is 5.8 kb in size and is expressed in many different muscle and non-muscle tissues. As we demonstrated previously, the α-dystroglycan differs in molecular weight between muscle and non-muscle rabbit tissues. Based on alteration of antibody reactivity we suggested that tissue specific isoforms of rabbit dystroglycan differ by sugar composition, but not by core protein structure (7). The data presented in this report also favor this idea. Southern blot hybridization and chromosomal localization rule out the possibility of multiple genes encoding human dystroglycan. Additionally, the proposed relatively simple two-exon genomic structure of the human dystroglycan gene does not leave much opportunity for alternative splicing. It is possible that additional exons exist in the 5′-untranslated region or even within the large intron we have identified, since the two exons were identified based only on the skeletal muscle specific cDNA sequence. However, we have found that the same 5.8 kb mRNA could be detected using exon 1 and exon 2 specific probes in muscle and non-muscle tissues. These results were confirmed by RT-PCR as described in the result section. Taken together, these data suggest carbohydrate-mediated isoform differences of dystroglycan in muscle and non-muscle tissues.

We have mapped the dystroglycan locus (DAG1) to human chromosome 3 by analysis of somatic cell hybrids. In situ

hybridization of the biotin-labeled genomic dystroglycan probe sublocalized the gene to the 3p21.1−p21.31 region. Other gene loci mapped to human chromosome band 3p21 include: aminoacylase 1 (ACY1), aminolevulinate, delta, synthases 1 and 3 (ALAS1 and ALAS3), N-acylaminoacyl-peptide hydrolase (APEH), aplysia ras-related homolog 12 (ARH12), collagen type VII, alpha 1 (COL7A1), guanine nucleotade binding protein, alpha (GNAI2 and GNAT1), inter-alpha (globulin) inhibitor, H3 polypeptide (ITIH3), macrophage stimulating 1 or hepatocyte growth factor-like (MST1), myosin, light polypeptide 3 (MYL3), pyruvate dehydrogenase, E1 beta polypeptide (PDHB), and zinc finger proteins 35, 52 and 64 (ZNF35, ZNF52 and ZNF64) (19). Of particular interest here is COL7A1 which is mutated in some patients with epidermolysis bullosa (both autosomal dominant and recessive forms). In addition, there are reported cases of siblings and other individuals with epidermolysis bullosa and an associated muscular dystrophy (20,21). Because mutations in dystroglycan could be responsible for some types of muscular dystrophy, the dystroglycan locus could be closely linked to the COL7A1 locus and could possibly be involved in deletions or rearrangements in patients who have both epidermolysis bullosa and muscular dystrophy.

A conserved region of synteny exists on human chromosome 3p21−q23 and distal mouse chromosome 9 that includes the homologous loci MYL3/Mylc, ACY1/Acy-1, GNAI2/Gnai-2, GNAT1/Gnat-1, already mentioned above at human chromosome 3p21, in addition to galactosidase beta 1 (GLB1/Bgl) at 3p23−p22, cholecystokinin (CCK/Cck) at 3pter−p21, retinol binding proteins 1 and 2 (RBP1/Crbp and RBP2/Crbp-2) at 3q21−q22 and 3p11−qter, respectively, transferrin (TF/trf) at 3q21, and lactotransferrin (LTF/Ltf) at 3q21−q23.

Interestingly, two neurological mouse mutations, ducky (du) at 60 cM and tippy (tip) at 61 cM are located on distal mouse chromosome 9 close to GNAI2/Gnai-2 and GNAT1/Gnat-1 which lie at 58 cM on the linkage map and whose homologs map to human chromosome 3p21 (22). Both du and tip are autosomal recessive mutations and could be models for human inherited diseases. Homozygous ducky mice (du/du) show a waddling or reeling gait and have a tendency to fall to one side. Histologically they have a severe dysgenesis of the hindbrain and spinal cord, myelin deficiency, demyelination and axonal dystrophy in selective fiber systems, such as the spinocerebellar and vestibulospinal tracts. Homozygous tippy mice (tip/tip) mice are small, hyperactive and cannot stand or walk without falling over (23). Given dystroglycan expression in brain, it is not unreasonable to consider these mutations for dystroglycan involvement.

Chromosome band 3p21 is also involved in a number of malignancies by either deletion and consequently, loss of heterozygosity (LOH), or as a breakpoint in chromosomal rearrangements that are associated with particular malignancies (24−27).

Finally, the dystroglycan gene might be considered as possible candidate gene for a Fukuyama-type congenital muscular dystrophy (FCMD), an autosomal recessive muscular dystrophy which is not yet linked to a particular chromosome. The phenotype of the disease consists of a combination of severe muscular dystrophy and anomaly of the brain (28). In addition, recent observations suggest the possible interaction between dystrophin and the FCMD gene product (29). Interestingly, dystroglycan is expressed in both muscle and brain, which are affected in FCMD patients. Moreover, recent findings demonstrate that all dystrophin-associated proteins are reduced in skeletal muscle, but the reduction of dystroglycan is most prominent despite near-normal expression of dystrophin (30). Knowledge of the genomic structure will greatly facilitate the scanning of exons of dystroglycan for sequence alterations in different forms of neuromuscular diseases with possible involvement of dystroglycan as a primary target of genetic abnormality.

MATERIALS AND METHODS

Cloning and sequence analysis of human dystroglycan cDNA

A human fetal skeletal muscle λgt10 cDNA library (kindly provided by Dr. Michel Koenig) was initially screened with ^{32}P-labeled R43-B cDNA (random primed labelling kit, Boehringer Mannheim) encoding rabbit dystroglycan (7). The HD-1, HD-2, HD-3 clones were isolated from the same library, and the HD-4 clone was isolated from a human skeletal muscle cDNA lgt11 library (Clontech). All cDNA inserts were sequenced on both strands either on an Applied Biosystems Automatic Sequencer or manually by the dideoxy chain termination method (31). Sequences were analysed with the Genetics Computer Group (Wisconsin package) and PCgene (Intelligenetics) software.

Northern blot analysis

A multiple human tissue mRNA blot (Clontech) was prehybridized at 42°C in 5×SSC, 5×Denhardt's solution, 50% formamide, 10% dextran sulphate and 100 mg/ml of salmon sperm DNA. The membrane was hybridized with ^{32}P-labeled HD-2 cDNA overnight at 42°C at a 1×10^6 c.p.m./ml. Membranes were washed at 62°C in 2×SSC, 0.1% SDS and were exposed to film (X-OMAT AR, Kodak) at −80°C.

Southern blot analysis

High molecular weight DNA was isolated from peripheral blood lymphocytes, digested with several restriction enzymes, electrophoresed on a 0.8% agarose gel and transferred onto nylon membrane. Prehybridization was done in 6×SSC, 5×Denhardt's solution, 10% dextran sulfate, 1%SDS and 100 μg/ml salmon sperm DNA at 65°C. For overnight hybridization at 65°C, 2×10^6 cpm/ml of ^{32}P-labelled HD-2 cDNA was used. Washes were performed as follows: twice in 2×SSC, 0.5%SDS for 10 minutes at room temperature, once in 1×SSC, 1%SDS for 15 minutes at 65°C, twice in 0.1×SSC, 1%SDS for 15 minutes at 65°C. The blot was exposed to X-ray film at −80°C.

RT-PCR analysis

1st-strand cDNA was synthesized from 0.5 μg of either human brain or cardiac mRNA. RNA was preheated at 70°C for 2 minutes and incubated in 50 mM Tris−HCl, 75 mM KCl, 3 mM $MgCl_2$, 0.5 mM each dNTP, 20 pmole random hexamer primer, 1 unit/μl RNase inhibitor and 200 units/μg RNA M-MLV Reverse transcriptase at 42°C for 1 hour in total volume of 100μl. For each amplification 5 μl of 1st-strand cDNA was used. PCR was carried out by 25 cycles of denaturation at 94°C (1 min), annealing at 65°C (1 min), extension at 72°C (1 min) under conditions recommended by the manufacturer (Perkin Elmer Cetus).

Somatic cell hybrid analysis

Radioactively labeled dystroglycan cDNA was hybridized to a Southern blot of Bgl II-digested genomic DNA from normal human and Chinese hamster controls and 11 human×Chinese hamster somatic cell hybrid lines derived from six independent fusion experiments (32). Prehybridization and hybridization solutions consisted of 4×SSC/5×Denhardt's solution/100 mM Na_2HPO4, 1% SDS and 0.1 mg/ml salmon sperm DNA. Hybridization, post-hybridization washes and film exposure were as previously described (32).

Fluorescence *in situ* hybridization (FISH)

A plasmid containing a genomic 9 kb insert (clone G2) specific for dystroglycan was biotin-11-dUTP labeled by nick-translation (Boehringer Mannheim) and hybridized to human metaphase chromosomes prepared by a 17 hr synchronization with methotrexate and a 5 hr release with BrdU and standard cytogenetic methods. This procedure generates an R-banding pattern when chromosomes are counterstained with propidium iodide. Hybridization solution consisted of 6.5 ng/μl labeled probe, 200 ng/μl each of human placental and salmon sperm DNA as competitors, 50% formamide, 10% dextran sulfate and 2×SSC. FISH was carried out essentially as previously described using a biotin/avidin/FITC detection system (33). The location of fluorescent signals was recorded with reference to the R-banding pattern on chromosomes. Metaphase spreads were examined using

a Zeiss Axiophot microscope equipped with epifluorescence. Images were captured using a cooled charge coupled device (CCD) camera (PM512, Photometrics) and a Macintosh computer using the GeneJoin program (developed by Tim Rand, Yale University). These PICT images were made into color slides and converted into photographs.

ACKNOWLEDGEMENTS

We gratefully acknowledge Richard Anderson for an excellent technical assistance and Dr. Michel Koenig for the kind gift of the human skeletal muscle cDNA library. We would also like to thank Dr. Val Sheffield for many helpful discussions and Drs. Michael Solursh, Gregory Kitten, Kiichiro Matsumura and Steven Roberds for their comments on the manuscript. K.P.C. and U.F. are Investigators of the Howard Hughes Medical Institute.

ABBREVIATIONS

DMD—Duchenne muscular dystrophy
SCARMD—Severe childhood autosomal recessive muscular dystrophy
FCMD—Fukuyama type congenital muscular dystrophy
DAG1—gene symbol for dystroglycan (dystrophin-associated glycoprotein 1)
RT-PCR—reverse transcription of RNA followed by the Polymerase Chain Reaction

REFERENCES

1. Koenig, M., Monako, A.P., Kunkel, L.M. (1988) Cell 53, 219–228.
2. Anderson, M.S., Kunkel, L.M. (1992) Trends Biochem. Sci. 17, 289–292.
3. Watkins, S.C., Hoffman, E.P., Slayter, H.S., Kunkel, L.M. (1988) Nature 333, 861–866.
4. Arahata, K., Ishiura, S., Ishiguro, T., Tsukahara, T., Suhara, Y., Eguchi, C., Ishihara, T., Nonaka, I., Ozawa, E., Sugita, H. (1988) Nature 333, 861–866.
5. Byers, T.J., Kunkel, L.M., Watkins, S.C. (1991) J. Cell Biol. 115, 411–421.
6. Ervasti, J.M., Campbell, K.P. (1991) Cell 66, 1121–1131.
7. Ibraghimov-Beskrovnaya, O., Ervasti, J.M., Leveille, C.J., Slaughter, C.A., Sernett, S.W., Campbell, K.P. (1992) Nature 355, 696–702.
8. Klietsch, R., Ervasti, J.M., Arnold, W., Campbell, K.P., Jorgensen, A.O. (1993) Circ. Research 72, 349–360.
9. Ervasti, J.M., Ohlendieck, K., Kahl, S.D., Gaver, M.G., Campbell, K.P. (1990) Nature 345, 315–319.
10. Ohlendieck, K., Campbell, K.P. (1991) J. Cell. Biol. 115, 1685–1694.
11. Matsumura, K., Tome, F.M.S., Collin, H., Azibi, K., Chaouch, M., Kaplan, J.C., Fardeau, M., Campbell, K.P. (1992) Nature 359, 320–322.
12. Ramarao, C.S., Kincade, Denker, J.M., Perez, D.M., Gaivin, R.J., Riek, R.P., Graham, R.M. (1992) J. Biol. Chem. 267, 21936–21945.
13. Kobilka, B.K., Frielle T., Dohlman, H.G., Balanowski, M.A., Dixon, R.A., Keller, P., Caron, M.G., Lefkowitz, R.J. (1987) J. Biol. Chem. 262, 7321–7327.
14. Emorine, L.J., Marullo, S., Briend-Sutren, B., Patey, G., Tate, K., Delavier-Klutchko, C., Strosberg, A.D. (1989) Science 245, 1118–1121.
15. Hook, M., Kjellen, L., Johansson, S., Robinson, J. (1984) Ann. Rev. Biochem. 53, 847–869.
16. Hassel, J.R., Kimura, J.H., Hascall, V.C. (1986) Ann. Rev. Biochem. 55, 539–567.
17. Bourdon, M.A., Shiga, M., Ruoslahti, E. (1986) J. Biol. Chem. 261, 12534–12537.
18. Nishiyama, A.N., Dahlin, K.J., Prince,J.T., Johnstone, S.R., Stallcup, W.B. (1991) J. Cell. Biol. 114, 359–371.
19. Naylor, S., and Carritt, B. (1991) Cytogenet. Cell Genet. 58, 170–230.
20. Niemi, K.M., Sommer, H., Kero, M., Kanerva, L., Haltia, M. (1988) Arch. Dermatol. 124, 551–554.
21. Fine, J.D., Stenn, J., Johnson, L., Wright, T., Bock, H.G.O., Horiguchi, Y. (1989) Arch. Dermatol. 125, 931–938.
22. Kingsley, D.M. (1991) Mamm. Genome 1, 127–145.
23. Lyon, M.F. and Searle, A.G. Eds. (1989) Genetic Variants and Strains of the Laboratory Mouse. 2nd edition. Oxford Univ. Press pp64 and 367.
24. Mitelman, F., Kaneko, Y., Trent, J. (1991) Report of the committee on chromosome changes in neoplasia. Cytogenet. Cell Genet 58, 1053–1079.
25. Mori, N., Yokota, J., Oshimura, M., Cavenee, W.K., Mizoguchi, H., Noguchi, M., Shimosato, Y., Sugimura, T., Terada, M. (1989) Cancer Res. 49, 5130–5135.
26. Yamakawa, K., Morita, R., Takahashi, E., Hori, T., Ishikawa, J., Nakamura, Y. (1991) Cancer Res. 51, 4707–4711.
27. van der Hout, A.H., van der Vlies, P., Wijmenga, C.L.F.P., Oosterhuis, J.W., Buys, C.H. (1991) Genomics 11, 537–542.
28. Nonaka, I., Chou, S.M. (1979) Congenital muscular dystrophy. In: Vinken PJ, Bruyn GW (Eds) Handbook of Clinical Neurology, North-Holland Publ Co, Amsterdam, New York, Oxford 41, 27–50.
29. Beggs, A.H., Neumann, P.E., Arahata, K., Arikawa, E., Nonaka, I., Anderson, M.S. and Kunkel, L.M. (1992) Proc.Nat. Acad.Sci. U.S.A. 89, 623–627.
30. Matsumura, K., Nonaka, I., Campbell, K.P. (1993) The Lancet 34, 521–522 (1993).
31. Sanger, F., Nickler, S., Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A 74, 5463–5467.
32. Francke, U., Yang-Feng, T.L., Brissenden, J.E., Ullrich, A. (1986) Cold Spring Harbor Symp. Quant. Biol. 51, 855–866.
33. Milatovich, A., Travis, A., Grosschedl, R., Francke, U. (1991) Genomics 11, 1040–1048.