Ankyrin-G in skeletal muscle: Tissue-specific alternative splicing contributes to the complexity of the sarcolemmal cytoskeleton.

Alexander Hopitzan, Anthony Baines, Marie-Aline Ludosky, Michel Recouvreur, Ekaterini Kordeli

Ankyrin-G in skeletal muscle: Tissue-specific alternative splicing contributes to the complexity of the sarcolemmal cytoskeleton.

Alexander A. Hopitzan[1], Anthony J. Baines[2], Marie-Aline Ludosky[1],

Michel Recouvreur[1] and Ekaterini Kordeli[1*]

[1]Institut Jacques Monod/CNRS/Universités Paris VI et VII; 2 Place Jussieu, F-75251 Paris cedex 05, France

[2]Department of Biosciences, University of Kent, Canterbury, Kent, CT2 7NJ, England

*Corresponding author: Ekaterini Kordeli

Email: kordeli@ijm.jussieu.fr, Telephone: +33 (1) 44 27 42 18, Fax: +33 (1) 44 27 59 94.

Footnote

The nucleotide sequences for Ank_{G109}, Ank_{G128}, Ank_{G130}, Ank_{G197}, Ank_{G217}, and partial *Ank3* genomic sequence have been deposited in the GenBank/EMBL database under GenBank accession nos. **AJ812019**, **AJ812020**, **AJ812021**, **AJ812022**, **AJ812023**, **AJ812024**, and **AJ812025**, respectively.

Abstract

Ankyrins are versatile adaptor proteins that join the spectrin-based cytoskeleton to transmembrane proteins, and have roles in organizing the microstructure of cell membranes. Molecular diversity of ankyrins in mammals arises from extensive alternative splicing of the products of three genes. There has been no systematic analysis of the diversity of expression of ankyrins-G, the widely expressed *Ank3* gene products, in a complex tissue. We previously described Ank_{G107}, the first muscle-specific ankyrin-G. Here we combined cDNA and database analyses to gain novel insight into the ankyrins-G of skeletal muscle. We find: (i) that *Ank3* is composed of at least 53 exons, many of which are subject to tissue-specific splicing; (ii) five novel full-length cDNAs encoding two canonical (Ank_{G197}, Ank_{G217}) and three small isoforms (Ank_{G109}, Ank_{G128}, Ank_{G130}) bring to six the number of ankyrins-G expressed in skeletal muscle; (iii) a 76-residue C-terminal domain is a 'signature' for muscle ankyrins; (iv) variably spliced sequences of 17/18 and 195 residues increase diversity in the C-terminal domains. Comparison of endogenous ankyrins-G with *in vitro* translated cDNAs revealed that small ankyrins account for the majority of the immunoreactivity for ankyrin-G in soleus muscle. The small ankyrins, when expressed *in vivo* in the rat muscle, are all targeted to sarcolemmal costameres. Our results demonstrate the tissue-dependent alternative splicing of *Ank3* in skeletal muscle and point to novel functions of small ankyrins-G in organizing microdomains of the plasma membrane.

Keywords:

Ankyrin-G; *Ank3*; Alternative splicing; Cytoskeleton; Skeletal muscle; Costamere; Sarcolemma; Bioinformatics; Gene structure; Direct gene transfer.

Introduction

Ankyrins, the versatile molecular adaptors of the spectrin-based skeleton, are responsible for localizing integral proteins to the appropriate membrane domains. In vertebrates, ankyrins are encoded by three genes, designated *Ank1, Ank2* and *Ank3* in rodents (*ANK1, ANK2* and *ANK3* in humans). Tissue-specific alternative splicing of transcripts results in expression of multiple isoforms (ankyrins-R, ankyrins-B and ankyrins-G respectively) in most tissues displaying different localizations and related but distinct functions. Gene knock-out experiments in mice [1-3], siRNA depletion [4] and natural mutations in humans [5-7] demonstrate a role of ankyrins in organizing membrane domains and in delivering ion channels and cell adhesion proteins to the functionally relevant membrane sites.

Specialized functions of ankyrins in the various tissues rely on important structural diversity. Ankyrins are modular polypeptides composed of highly conserved membrane-binding (N-terminal), spectrin-binding and death domains, and a C-terminal domain that is variable among the three gene products. The membrane-binding domain is composed of 24 tandem copies of 33-residue Ank-repeats which provide sites of protein-protein interaction in numerous proteins [8] and bind to the cytoplasmic domains of most ankyrin-associated integral proteins [9]. The central, spectrin-binding domain associates with β spectrins. Recently, the death domain of ankyrins-G has been shown to interact with the proapoptotic molecule Fas and to promote fas-mediated cell death in renal epithelia [10]. The divergent C-terminal domains are subject to extensive alternative splicing, and appear to dictate functional specialization of the various isoforms, including binding and targeting specificities. An alternatively spliced 17.3 kDa insert in the C-terminal domain of erythrocyte ankyrin-R modulates binding affinities for both spectrin and the anion exchanger [11, 12]. The ankyrin-B C-terminal domain is required for rescuing abnormal

inositol 1,4,5 triphosphate receptor localization in ankyrin-B(-/-) neonatal cardiomyocytes [13]; in these experiments, ankyrin-B and ankyrin-G C-terminal domains do not compensate for each other. Despite accumulating evidence pointing to the functional potential of C-terminal domains, presently only one interaction is known between the ankyrin-B C-terminal domain and a molecular chaperone, Hdj1/Hsp40 [14]. Different combinations of domains and the use of alternative transcription start sites generate three classes of ankyrin polypeptides: the canonical, Ank-repeat containing; the small, lacking all or part of Ank-repeats and spectrin-binding domains; the large, displaying huge insertions in the C-terminal domains, found so far only in the nervous system [9].

Alternative splicing has emerged as a major source of proteome diversity in mammals; according to large-scale genomics studies it may occur in 40 to 60% of human genes [15]. Pre-mRNA splicing combined with alternative promoter usage is a mechanism commonly used by components of the spectrin-based skeleton to increase functional diversity and to regulate cell type-specific expression [9, 16]. Elucidation of the structure and organization of the human *ANK1* gene allowed the understanding of the complex alternative splicing pattern of ankyrins-R [17]. However, little is known about the other two genes, and in particular the broadly expressed *Ank3*. This gene codes for a variety of alternatively spliced isoforms including large [18], canonical [19, 20] and small variants [21-23] with molecular masses ranging from ~100 kDa to 480 kDa.

We recently identified Ank_{G107}, a novel small ankyrin-G from rat skeletal muscle, which lacks the entire Ank-repeat domain, contains a unique muscle-specific 76-residue insertion in the C-terminal domain and is targeted to the sarcolemma [24]. In skeletal muscle fibers, assembly of specialized membrane domains is a functional requirement, both at the cell surface and in the cytoplasm where Ca^{2+}-regulated excitation-contraction coupling occurs. Ankyrins-B have been

lately implicated in human heart disease due to abnormal coordination of multiple functionally related ion channels and transporters [5, 7]. Multiple ankyrins are expressed in skeletal muscle fibers by all three genes [2, 21, 25-28] and localize to the postsynaptic membrane [29, 30], the costameres [3, 31, 32], the triads and the non junctional sarcoplasmic reticulum [27, 28, 33]. A muscle-specific promoter gives rise to sAnk1, a truncated ankyrin-R which is thought to link the sarcoplasmic reticulum to myofibrils [27, 34, 35]. To understand the complex functions of ankyrins in muscle cells, we sought to establish the tissue-specific expression of muscle ankyrins-G. Here we report five novel ankyrins-G including canonical and small variants. Annotation of their full length cDNAs together with all so far described ankyrin-G sequences to the rat *Ank3* gene led to the demonstration of a muscle-specific alternative splicing pattern. Moreover, the small isoforms appear to represent the majority of ankyrins-G in skeletal muscle and they are targeted to sarcolemmal costameres despite the lack of Ank-repeats.

Materials and methods

Isolation of cDNA clones

A rat skeletal muscle λgt10 cDNA library (Clontech) was screened with two *Ank3* cDNA fragments derived by PCR (polymerase chain reaction) amplification from 270-kDa ankyrin-G (GenBank Accession no. AF102552) membrane-binding (bp 1481-2381) and C-terminal (bp 6999-7867) domains. The probes were labelled with [α-^{32}P]dCTP using a random primer labelling kit (Stratagene). Plaque-lifted filters (Nytran Plus; Schleicher & Schuell) were hybridized overnight in hybridization buffer containing 6X SSC (1X SSC: 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 5X Denhardt's, 0.2% SDS, and 50 μg/ml sonicated salmon sperm at 65°C. The filters were washed at a maximum stringency 0.2X SSC, 0.1% SDS at 65°C. The positive phage clones were plaque purified via three rounds of screening. Phage DNA from selected clonal isolates was either extracted from plate lysates by standard methods (Qiagen Lambda Kit; Qiagen) or amplified by PCR using primers flanking the insert sequences (gt1, gt2; see Table1) and a high-fidelity polymerase mix (Advantage; Clontech). Restricted cDNA inserts were gel purified (Qiaquick Gel Extraction Kit; Qiagen) and subcloned into pUC18 vector (Qbiogene). DNA sequence determination was performed by MWG, Germany. The full length cDNAs not included in these phages were obtained by nested PCR amplifications using 0.5 ng of rat skeletal muscle cDNA (Marathon-Ready; Clontech) as a template, two pairs of untranslated region (UTR)-based primers (Table 1), and a high-fidelity polymerase mix (Advantage 2; Clontech). The primer combinations for the long isoforms were 5'UTRmb-ne1 and 13/2-1R for the first round, and 5'UTRmb-ne2 and 3'UTRn2 for the second round of PCR; the primer combinations for the short isoforms were 5'Malpa and 13/2-1R for the first round, and Malpan2 and 3'UTRn2 for the second round of PCR. PCR products were subcloned into the TOPO-pCR2.1 vector (Invitrogen). Clones were analyzed by standard colony PCR methods to

distinguish between cDNA inserts of different size and composition. DNA from selected clones was extracted by standard methods (Qiagen Plasmid Midi Kit; Qiagen) and sequenced.

Rapid amplification of cDNA ends (RACE) analysis

The 5' and 3' ends of the *Ank3* transcripts were determined using Marathon-Ready rat skeletal muscle cDNA (RACE system; Clontech). 5' and 3' RACE was performed following the recommended protocol except that the second PCR reactions were carried out with 1/100 of the material from the first reaction as a template. The gene-specific primers were as follows (Table 1). 5' RACE: 5'RACEn3 and 3'Malpa (initial PCR reactions), and 5'RACEn4 and 3'Malpan2 (nested PCR reactions) for the long and short isoforms respectively; 3' RACE: 9/3-1 (initial PCR reactions), and either 17Aup or 9/3-2 (nested PCR reactions). The amplified fragments were subcloned into the TOPO-pCR2.1 vector (Invitrogen) and sequenced.

Bioinformatics

Routine sequence analysis of the nucleotide and protein sequences was performed using the LASERGENE software package (DNASTAR). Cloned muscle ankyrin-G cDNAs were used for querying the NCBI BLAST server. Sequences from
various rat *Ank3* transcripts were extracted from the non-redundant (nr) DNA database. To determine the chromosomal location of the *Ank3* gene a BLAT query (genome.ucsc.edu) of the rat genome was performed using cloned (this paper) and database-derived ankyrin-G cDNA sequences. The genomic sequence corresponding to the *Ank3* gene plus ~180 kb extensions at both ends was extracted from the Ensembl database (Ensembl Rat release 14.2.1; www.ensembl.org/Rattus_norvegicus). Transcribed cDNA sequence data were aligned to the genomic *Ank3* sequence using the EMBOSS program est2genome (www.hgmp.mrc.ac.uk/Software/EMBOSS/). The textual outputs of est2genome were

graphically displayed and processed with the genome annotation tool Artemis release 5.0 (www.sanger.ac.uk/Software/Artemis/v5/). In all cases predicted matches and the presence of canonical splicing sites (GT/AG) was confirmed by eye. Poly-A sites in the genomic sequence were predicted using the program polyadq (argon.cshl.org/tabaska/polyadq_form.html) analyzing both canonical poly-A patterns AATAAA and ATTAAA. Rat expressed sequence tags (EST) sequences were identified by querying the Gene2EST BLAST Server (woody.embl-heidelberg.de/gene2est/) with the *Ank3* genomic sequence. Database-derived protein sequences of human and murine ankyrins-G were annotated to the rat *Ank3* gene using the program Wise2 (www.ebi.ac.uk/Wise2/).

Isolation of genomic clones

The genomic region containing the novel muscle *Ank3* exon 25 not present in the Ensembl database was amplified by PCR from rat genomic DNA (Clontech) using exon-based primers e24u and e26r1 (Table1) and a high-fidelity polymerase mix (Elongase Amplification System; Invitrogen) according to the manufacturer's recommendations. Cycling conditions were as following: 94 °C for 1 min; 8 cycles of 15 s at 94 °C and 10 min at 72 °C decreasing 0.5 °C per cycle; 27 cycles of 15 s at 94°C and 10 min at 68 °C; final extension at 68 °C for 15 min. The amplified fragment was subcloned into the TOPO-pCR2.1 vector (Invitrogen) and sequenced.

In vitro expression and immunoblot analyses

Coupled in vitro transcription and translation analyses were carried out employing a rabbit reticulocyte based T7 expression kit (Promega) according to the manufacturer's protocol. PCR-amplified cDNAs encoding all small muscle ankyrin-G isoforms previously subcloned via an EcoRI site into pcDNA3 vector (Invitrogen) were used. Products were resolved by SDS PAGE

electrophoresis and revealed by western blot analysis using polyclonal antibodies against the 76aa insert [24]. In control experiments for both the reaction and the specificity of the antibody, in vitro translation was performed with cDNAs encoding anti-sense Ank_{G107} and Ank_{G107} lacking the 76aa insert ($Ank_{G107}\Delta 76aa$). Both controls were negative.

Northern Blot Analysis

Rat poly(A)+ RNA multiple tissue Northern blots (Clontech, Origene) were consecutively hybridized with cDNA probes corresponding to epithelial insert C (Ank_{G128}, bp 2212-2703) and specrtrin-binding domain (Ank_{G128}, bp 68-988) sequences of rat muscle ankyrins-G. The probes were labeled with [α-^{32}P]dCTP using a random primer labeling kit (Stratagene) and purified by standard chromatography methods. Hybridizations were carried out according to the manufacturer's recommendations using ExpressHyb solution (Clontech). The membranes were washed at a maximum stringency of 0.1XSSC, 0.1% SDS, 50°C prior to autoradiography.

In vivo DNA injection and immunofluorescence confocal microscopy

Direct gene transfer experiments into rat muscles were performed as described [36, 37]. Briefly, pEGFP-N1 plasmids containing full length cDNAs for the small muscle ankyrins-G were diluted to a final concentration of 1-2 mg/ml in a 20% sucrose / 0.5x PBS (10 mM phosphate buffer, 150 mM NaCl, pH 7.4) solution. For each isoform, about 40 µl of DNA solution were injected into 12-day old rat anterior tibialis muscles *in vivo*. As a control, the vector without ankyrin-coding sequences was used. Two independent injection experiments per isoform were performed using rats from different liters. Ten days after injection, muscles were removed, fixed with 3% paraformaldehyde / 0.1 M phosphate buffer pH 7.4 for 1h at 4°C, cryoprotected with 0.7 M sucrose and frozen in liquid nitrogen-cooled isopentane. Cryostat 10 µm-thick sections of muscle

fibers expressing the fusion proteins were selected by direct GFP fluorescence and double-labeled for indirect immunofluorescence with primary antibodies diluted at 2-5 μg/ml as previously described [28]. Monoclonal antibodies to the green fluorescent protein (GFP) were from Sigma. Polyclonal antibodies to αII-spectrin were previously described [38] and were a generous gift from Dr. M. C. Lecomte (INTS, Paris). FITC- and Cy3-conjugated secondary antibodies were from Jackson Immunoresearch Laboratories (West Grove, PA). Confocal laser scanning microscopy was performed using a TCS-4D confocal imaging system (Leica instrument), equipped with a 63x objective (plan apo; numerical aperture = 1.32). For FITC and CY3 excitation, an Argon-Krypton ion laser adjusted at 488 and 568 nm, respectively, was used. The confocal system was adjusted to allow field depths of ~0.5 μm. Digital images were captured using MetaView Imaging System (Universal Imaging Corp.) and arranged using Adobe PhotoShop. Animal experiments were carried out according to relevant ethical protocols.

Results

Isolation and characterization of rat skeletal muscle ankyrin-G cDNAs

We previously reported the expression of multiple ankyrin-G isoforms in skeletal muscle, based on the detection of two clusters of 8 kb and 5.6 kb *Ank3* transcripts [24]. Northern blot analysis using domain-specific probes strongly suggested the presence of two types of ankyrins, the canonical Ank-repeat-containing, and the small variants. Only one small isoforms, Ank_{G107}, had been cloned so far. To isolate a broad range of ankyrins-G including the hypothetical canonical isoforms, a rat skeletal muscle cDNA library (Clontech) was screened with probes corresponding to the generic ankyrin-G C-terminal domain and to Ank-repeats (Figure 1A).

Two out of twelve independent clones contained Ank-repeats and covered sequences ranging from 5'UTR to the beginning of the spectrin-binding domain. Coding sequence begun with exon 1e previously described for epithelial and brain canonical ankyrins-G [2]. Three clones, including one full-length clone, covered a different 5'UTR and 5' coding sequence starting with amino acids MALPHS previously found in Ank_{G107} [24] and in the murine epithelial Ank3-5kb [19], and followed by the spectrin-binding domain. These results confirmed the presence of canonical and small ankyrins-G in skeletal muscle. Clones covering 3'coding and 3' UTR sequences contained two additional inserts as compared to Ank_{G107}: a 195-residue sequence originally described in epithelial ankyrins-G from kidney [19], and a novel 17/18-residue insert (17/18 aa) preceding the muscle-specific 76 aa stretch (Figure 1B and Supplementary Figure 1).

In order to determine 5' and 3' ends of muscle *Ank3* transcripts and to investigate further sequence variations we performed RACEs employing different pairs of nested primers (Table 1) and rat muscle Marathon-Ready cDNA as a template. 5'RACE for both canonical and small ankyrins-G resulted in cDNAs differing only in the length of their 5' UTRs (Supplementary Figure 2). No alternative 5´ coding sequence was identified. 3' RACE was performed with

primers corresponding either to sequences shared by all muscle transcripts, or to the 17/18 aa insert. Four types of cDNAs were isolated showing identical coding and UTR sequences differing only in their length beyond the stop codon.

To define the exact combinations of 5' and 3' sequences of each isoform, full length cDNAs were obtained by PCR using rat skeletal muscle Marathon-Ready cDNA (Clontech) as a template. Six muscle ankyrin-G isoforms were identified (Figure 1A) including Ank_{G107}. By analogy with Ank_{G107}, we termed the five novel isoforms according to their calculated molecular weight Ank_{G109} (ORF of 2571 bp), Ank_{G128} (ORF of 3471 bp), Ank_{G130} (ORF of 3522 bp), Ank_{G197} (ORF of 5421 bp), and Ank_{G217} (ORF of 5955 bp).

Previous studies [24] reported that the primary sequence of the first muscle isoform Ank_{G107} differed from homologous *Ank3* gene products isolated from other tissues at three major points: 1) residues PKI replace residues LRSF within the spectrin-binding domain (aa 48-50); 2) a deletion of 12 residues occurs in the end of the spectrin-binding domain at position aa 590; 3) a 76-residue insert (aa 864-939) is present near the COOH terminus (Figure 1A, Ank_{G107}: a, b and c, respectively). Sequence analysis of the five novel muscle isoforms showed that these features are common to all six muscle ankyrins-G (Figure 1A), strongly suggesting tissue-specific alternative splicing events. We refer to these features as muscle-specific signatures for ankyrins-G.

Sequence differences between the six isoforms occur in the N-terminal region (presence or absence of the 24 membrane-binding Ank-repeats), and in the C-terminal domain due to the 195 aa and 17/18 aa inserts. Sequence analysis of the muscle 195 aa revealed 88% identity on the protein level when compared to its murine homolog insert C in 7kb-Ank3 and 5kb-Ank3 kidney ankyrins-G (GenBank Accession no. AAB01607 and AAB01603, respectively). The difference between 17 aa and 18 aa is due to the additional N-terminal residue Gln (Figure 1B). Database

searches against the 17/18 aa yielded only one relevant hit, an EST clone from pooled human uterus adenocarcinoma (GenBank Accession no. AW439255). Canonical ankyrins contain either the 195 aa or the 18 aa inserts, whereas the small isoforms showed all four possible combinations with 17 aa and 195 aa either alternatively present, both missing (Ank_{G107}), or, in a single case, both present in the C-terminal domain. Another isolated case showed the presence of the 18 aa in a small isoform.

Annotation of cDNA sequences to the Ank3 locus

To determine exon usage and alternative splicing events leading to the muscle-specific expression of ankyrins-G, we examined the organization of the rat *Ank3* gene. A database search of rat genomic sequences employing all available sequence data of rat *Ank3* transcripts revealed that the *Ank3* gene spans a region of over 600 kb on rat chromosome 20p11. The genomic sequence assembly covering the *Ank3* locus was downloaded from the Ensembl database. Whenever necessary, we refined this genomic contig by additional cloning and sequence analysis to exclude any uncertainties caused by sequencing gaps.

The exon-intron structure was established by annotating our full length and RACE clones as well as previously cloned rat epithelial Ank_{G190} [20] and rat neuronal 270 kDa ankyrin-G [18] to the rat genomic sequence (Figure 2). Considering only rat ankyrin-G sequences our analysis identified 53 exons. Exons found in muscle transcripts were numbered 1 to 47 while non-muscle exons were indicated by roman numbers. All splice junctions displayed the canonical GT/AG dinucleotides. The complete list of exon/intron boundaries is available as supplementary data. The overall organization of the exons shed light on a complex pattern of alternative splicing. In addition to the two distinct first exons of muscle isoforms (exon 1 and 23), we annotated the two first alternative exons used by Ank_{G190} (exon I followed by exon II) and 270 kDa ankyrin-G (exon III). Exons 2 to 22 encod the 24 Ank-repeats; a 24 bp exon IV was only found in 270 kDa

ankyrin-G. Interestingly, muscle isoforms contain neither insert A nor insert B at the end of the membrane-binding domain previously identified in the mouse epithelial Ank3-7kb [19]; insert A (Figure 2, exon V) is present in both non-muscle Ank_{G190} and 270 kDa ankyrin-G.

The spectrin-binding domain is encoded by exons 24 to 36. The muscle-specific sequence PKI is encoded by a single exon of 12 bp (exon 25). Identification of this exon required cloning of the corresponding genomic sequence (Figure 3A) not yet present in the latest Ensembl Rat assembly 3 (v28.3e.1; February 3rd, 2005). Exon 26 contains an internal splice acceptor site used by all muscle transcripts. Part "a" of exon 26 (Figure 3B) encodes amino acids LRSF spliced out in muscle transcripts, and does not appear as an independent exon because of the lack of a splice donor site. In the end of the spectrin-binding domain, an internal splice donor in exon 36 causes the deletion of 12 residues observed in muscle isoforms. Part "a" of exon 36 lacks the acceptor splice site and consequently adjacent part "36b" does not appear as a distinct exon (Figure 2).

Between exons 36 and 37, a single exon VI, present in rat 270 kDa ankyrin-G, encodes the Ser/Thr rich sequence previously identified in neuronal ankyrins-G. These isoforms specifically localize at the nodes of Ranvier and initial axonal segments [18]. All so far muscle transcripts identified have this exon spliced out.

The C-terminal domain is encoded by exons 37 to 46. The conserved death domain sequences as defined by databases Pfam [39] (amino acids DIRMA…LEGPI) and SMART [40] (PQSPC… LEGPI) are present in all muscle isoforms. Exon 40 shows an internal splice acceptor site giving rise to transcripts containing part "b" either alone or in combination with part "a" (Figure 2). Part "a" encodes the entire 195 aa insert. Both variants have been found in muscle (this paper) and kidney [19]. The muscle-specific 76 aa insert is encoded by exactly three contiguous exons 43 to 45. Preceding exon 42 encodes the novel 17/18 aa segment. An internal splice donor site is responsible for the absence of the N-terminal Gln that distinguishes insert 18

aa from 17 aa. Interestingly, genomic sequence comparison revealed that rat exon 42 is highly conserved when compared to corresponding mouse, chimpanzee, and human sequence (96%, 81%, and 80% identity); however, in humans and chimpanzee only the 17 aa version is expressed due to a base exchange within the first acceptor splice site (data not shown). Exon 46 harbours the translation stop codon and 20 bp of 3'UTR sequence. Interestingly a second stop codon is found contiguous to the first one.

Further 3'UTR sequence is encoded by exon 47. This exon contains two "AATAAA" and one "ATTAAA" polyA signals 532 bp, 596 bp and 943 bp downstream of the first stop codon (Figure 2). 3'RACE clones extended beyond all three polyA signals. The longest clone ended 1070 bp downstream of the stop codon. Interestingly, all cDNAs terminated in A-rich sequences which were confirmed by comparison with the genomic sequence and could be the reason for annealing of oligo-d(T) primers during the first-strand cDNA synthesis. Several additional putative polyA signals were observed in the genomic region proximal to aligning 3' RACE sequences; the most likely to be a real poly-A site (70% probability) was found 1558 bp downstream of the stop codon. Two rat EST clones (GenBank Accession no. AA892174, kidney, and BE116901, whole embryo) extended our 3'RACE results and confirmed the usage of the predicted poly-A site (data not shown).

To extend insights into exon usage and splice patterns of muscle ankyrins-G, we performed a Gene2EST BLAST search of the rodent EST database employing the genomic rat *Ank3* sequence. EST sequences revealed no additional splice variants of muscle ankyrins-G. Further information on the gene structure was obtained by annotating all isoforms so far isolated from other species to the rat *Ank3* gene using the Wise2 program. Analysis of homologous sequences of human brain 480-kDa ankyrin-G (GenBank Accession no. AAA64834), human Golgi Ank_{G119} (AAB08437), mouse macrophage ankyrin-3 (AAB58381), mouse skin ank3-93

(AF212924), mouse epithelial Ank3-5kb (AAB01603) and Ank3-7kb (AAB01607) showed that the overall exon structure and splice sites are highly conserved.

Additional residues ETESDQDDE present only in the spectrin-binding domain of human 480-kDa ankyrin-G [18] corresponded to an individual exon between exons 35 and 36. Inserts A and B of epithelial ankyrins-G [19] are encoded by two individual exons. The huge 2604 aa tail domain of 480-kDa ankyrin-G is encoded by a single exon. Alternative splicing of this exon gives rise to the 850 aa tail domain of rat 270 kDa ankyrin-G (Figure 2, exon VI). The smaller known ankyrin-G is the skin-specific isoform ank-3_{93} [23] which lacks Ank-repeats and the first 75 residues of the spectrin-binding domain. Our analysis revealed that this variant results from alternative splicing joining rat exon 1 to exon 27.

Tissue expression of muscle ankyrin-G isoforms

Northern blot analysis was carried out to assess the tissue expression of ankyrins-G containing the novel inserts. High stringency hybridization of poly(A+)RNAs from several rat tissues was carried out with cDNAs corresponding to the195 aa insert (Figure 4A) and to the spectrin-binding domain (Figure 4B). These experiments showed that: 1) the tissue expression pattern of 195 aa-containing isoforms was comparable to the general pattern of ankyrins-G, with 195 aa-positive transcripts more abundant in heart, skin and testis, and 2) the relative hybridization signals of the two probes were different, suggesting that the expression of 195 aa-containing ankyrins-G varies among tissues with maximum relative levels in skin and reduced levels in kidney, brain, and testis. Skeletal muscle transcripts containing 195 aa corresponded to both 8 kb and 5.6 kb clusters of messages [24] (Figure 4C). Our attempts to get information on the tissue expression pattern of the 17/18 aa-containing isoforms using different oligonucleotide probes were not conclusive.

We further compared the electrophoretic mobility of *in vitro* translated small isoforms

Ank_{G107}, Ank_{G109}, Ank_{G128} and Ank_{G130}, with endogenous ankyrins-G in 18-day-old rat skeletal muscle homogenates (Figure 5). Western blot analysis using anti-ank_G76aa antibodies showed that the two endogenous polypeptides of 190 kDa and 130 kDa [24] co-migrated with isoforms either containing or not the 195 aa insert, respectively. The presence of the 17 aa insert did not significantly alter migration in this gel system. These results confirmed that the cloned isoforms carrying the novel inserts are indeed expressed in skeletal muscle fibers.

Intracellular targeting of muscle ankyrin-G isoforms

To gain insight into possible functional consequences of the different exon combinations of the C-terminal domains, we followed the subcellular targeting of the four small muscle isoforms when expressed as GFP fusion proteins in the rat skeletal muscle. cDNAs encoding Ank_{G107}, Ank_{G109}, Ank_{G128}, and Ank_{G130} were injected *in vivo* into the tibialis anterior muscle of 12 day-old rats and the distribution of expressed proteins was assessed ten days later on cryosections of injected muscles by indirect immunofluorescence and confocal microscopy (Figure 6). No evident differences in targeting between isoforms were observed at this level of analysis. All isoforms displayed a predominantly sarcolemmal distribution, as previously shown for Ank_{G107}. Low levels of intracellular staining did not show distinct patterns among isoforms. The only difference in localization occurred upon deletion of the 76 aa insert which resulted in increased cytoplasmic distribution in agreement with previous observations [24]. Colocalization of GFP fusion proteins (Figure 6A) with αII spectrin and superficial confocal optical sections (Figure 6B) strongly suggested a costameric pattern of distribution. In control experiments, GFP alone remained in the cytoplasm.

Discussion

This study provides the first insights into the organization of the *Ank3* gene, and a systematic analysis of the diversity of ankyrin-G expression in the rat skeletal muscle. In the case of large genes like *Ank3* undergoing extensive alternative splicing, the application of either gene finding programs or EST annotation alone does not yield satisfactory results [41]. On the other hand, cloning of partial cDNA fragments usually derived from either library screening or RT-PCR do not provide ultimate information on the exon composition of transcripts as expressed in tissues. We combined cloning of full length cDNAs and computing sequence analysis to show that i) the rat *Ank3* gene, located on chromosome 20p11, is composed of at least 53 exons spanning a region of over 600 kb known to be syntenic to human chromosome 10q21[42], ii) these exons exhibit a complex, tissue-dependent pattern of alternative splicing, and iii) in skeletal muscle two alternative first exons 1 and 23 give rise to transcripts encoding canonical and small ankyrins-G respectively. Exon 1 is expressed in several tissues including striated muscle, brain, kidney and skin [2, 23]. No other first exon was found in skeletal muscle transcripts by 5'RACE using primers within Ank-repeats. Exon 23 has been so far reported only in kidney [19] and skeletal muscle [24]. Ank-3_{93}, the smaller known ankyrin-G isoform [23] or even further truncated variants similar to the *Ank1* gene product sAnk1 [34, 35] were not detected in skeletal muscle by 5'RACE.

In this study, two canonical and four small muscle ankyrins-G have been isolated and characterized, including the previously reported Ank_{G107} [24]. We termed the novel isoforms Ank_{G109}, Ank_{G128}, Ank_{G130}, $Ank_{G197,}$ and Ank_{G217} based on their calculated molecular weight. The expression of these isoforms in muscle tissue was assessed by comparing the pattern of *in vitro* translated cDNAs on western blots with endogenous ankyrin-G polypeptides, as well as by Northern hybridization. The four small muscle isoforms comigrated with 130 kDa and 190 kDa

muscle polypeptides. Higher apparent molecular weight polypeptides previously reported in muscle lysates [24] should account for the canonical isoforms. Finally, expression of GFP fusion proteins in rat skeletal muscle fibers *in vivo* injected with cDNAs encoding the four small ankyrins-G showed preferentially targeting to the sarcolemma where they co-localize with spectrin at costameres.

One major conclusion of this study is that common sequence features distinguish muscle isoforms from ankyrins-G expressed in other tissues. Ank_{G107}, the first muscle ankyrin-G identified, previously revealed three unique sequence features as compared to all so far cloned ankyrins-G, namely the presence of residues PKI and deletion of the last 12 residues in the spectrin-binding domain, and the presence of the 76aa insert near the C-terminus [24]. These features were found in all five novel isoforms, strongly suggesting that they constitute muscle-specific "signatures" of ankyrins-G. The establishment of intron/exon boundaries demonstrated that all three features derive from alternative splicing (Figure 2).

Diversity within muscle isoforms occurs in the C-terminal domains. We identified two additional inserts issued from alternative splicing, 17/18 aa and 195 aa, which do not appear to be muscle-specific. The 195 aa sequence is homologous to the previously described epithelial insert C, a highly acidic alternative sequence of the mouse ankyrin-G C-terminal domain found in kidney transcripts [19]. The usage of the alternative splice site giving rise to 195 aa is not limited to certain tissues. Isoforms with and without insert have been previously observed in kidney [19], skin [23] and macrophages [22]. Accordingly, our Northern blot data show that 195 aa-containing transcripts are present in all ankyrin-G-expressing rat tissues here analyzed, including skeletal muscle and heart. However, their relative expression levels as compared with total *Ank3* transcripts varied among tissues, with an important enrichment in skin and reduced levels in kidney, brain, and testis. The novel 17/18 aa sequence immediately preceding the 76 aa insert is

encoded by a single 54 bp exon. Rat multi-tissue Northern blot analysis employing two different oligonucleotide probes was not conclusive. Comparison between vertebrate genomic sequences revealed that the 17/18 aa is highly conserved (data not shown). Nevertheless, extensive database searches identified only one EST clone (human uterus, pooled ardenocarcinoma) and showed no significant homology with known proteins including products of other ankyrin genes. Taken together these observations suggest that the 17/18 aa sequence is specific for ankyrins-G and it is weakly expressed in a limited number of tissues.

The high degree of conservation of vertebrate ankyrin-G protein sequences allowed us to precisely annotate all so far known isoforms to the rat genome, with three exceptions: i) sequence preceding the putative start codon of mouse macrophage ankyrin-3 [22] corresponds to three contiguous upstream exons in canonical ankyrins and therefore it is unlikely to be 5'UTR: ii) 5'UTR and the alternative first exon of human Ank_{G119} [21] could not be assigned to human chromosome 20; iii) the unique short C-terminal and 3'UTR sequences of Ank_{G119} could not be assigned to the human genome. This approach showed that the overall exon structure and splice sites are highly conserved among species. Examples of tissue-dependent preferential exon usage include the lack of both epithelial inserts A and B from muscle isoforms and Ank_{G119}, the only non-muscle isoform also lacking amino acids LRSF.

The here identified alternative splicing events address the functional consequences of structural diversity for muscle ankyrins-G. Various combinations of the two alternatively spliced 195 aa and 17/18 aa inserts in the C-terminal domain have been found in both types, canonical and small isoforms. The two inserts may provide novel ligand-binding capacities by either creating new binding sites or influencing interactions of neighbouring domains of the protein. Moreover, the 76 aa insert appears as an entity encoded by three complete exons that are spliced out together in non muscle transcripts, and may act as a separate module or protein-protein

interaction domain. Interestingly, the binding sites of the muscle-specific sarcoplasmic reticulum-associated sAnk1 isoform [27] for muscle proteins titin and obscurin [43-45] reside within the region of high homology with the 76 aa of muscle ankyrins-G. By analogy, the 76 aa may be involved in interactions with other sarcolemmic and/or cytoplasmic muscle proteins, as previously proposed for Ank_{G107} [24]. Such interactions as well as their potential modulation by the presence of the preceding 17/18 aa sequence remain to be established. Another interesting question to address will be the significance of the muscle-specific alternative sequence PKI in the spectrin-binding domain. Even very short alternative sequences may have major functional consequences, like for instance in the case of agrin, the heparan sulfate proteoglycan that organizes the postsynaptic differentiation of the neuromuscular junction [46].

One consequence of sequence variability could be a differential subcellular targeting. We previously showed that Ank_{G107} is targeted to the sarcolemma when transfected in primary cultures of skeletal muscle cells [24]. Here we applied direct gene transfer in skeletal muscle, an established approach for expression studies of muscle proteins [36, 37]. We followed the localization of the four GFP-tagged small muscle isoforms in living cells during final maturation of muscle fibers. Interestingly, all isoforms were preferentially addressed to the sarcolemma where they displayed a costameric pattern of distribution (Figure 6). Our results directly demonstrate that Ank-repeats are not required for membrane localization. The spectrin-binding domain alone is apparently capable of addressing ankyrins to the plasma membrane, as previously suggested in neurons and muscle cells [24, 47].

Costameres, the sarcolemmic structures that physically couple the force-generating sarcomeres to the sarcolemma and beyond, to neighbouring muscle cells via the extracellular matrix [48], are made of several multi-protein assemblies capable of binding actin: Na/K ATPase and the spectrin/ankyrin complex; integrins and the vinculin/talin/α-actinin complex; the

dystrophin complex. One obvious partner for ankyrins at costameres is the Na/K ATPase [20, 49-52]. However, in skeletal muscle fibers, products of all three ankyrin genes have been localized at the costameres [3, 24, 28, 32]. Moreover, our targeting studies suggest that several alternatively-spliced ankyrins-G are present at these membrane sites. Williams and colleagues [32] showed that an ankyrin-G isoform of apparent molecular mass 190 kDa co-localized at costameres and co-immunoprecipitated with Na,K-ATPase and alpha-II /beta-I spectrins of adult sarcolemma [53]. Data in this paper demonstrate that the small costameric isoforms lacking Ank-repeats are the major ankyrin-G polypeptides in skeletal muscle, they co-localize with sarcolemmal alpha-II spectrin at costameres, and two of them, Ank_{G128} and Ank_{G130}, have apparent molecular masses on western blots close to 190 kDa. Responsible for their association with the sarcolemma could be the direct interaction of the spectrin-binding domain not only with spectrin but also with other sarcolemmal proteins: Interestingly, two previous studies report that, in addition to and independently from Ank-repeats, the spectrin-binding domain also binds to the Na,K-ATPase [20, 50]. Taken together these results may suggest a novel role for small ankyrins, so far predominantly found in intracellular sites [19, 21, 22], in organizing the costameric microdomains by linking the Na,K-ATPase to muscle spectrin; alternatively, they may be important for localizing relevant cytoplasmic proteins to these membrane sites.

In conclusion, this study revealed important structural features of the *Ank3* gene and elucidated the molecular basis for the complexity of ankyrin-G transcripts in skeletal muscle. Furthermore, the isolation of five novel ankyrin-G isoforms provides the basis for identification of interacting proteins and determination of the functional role(s) of ankyrins-G in the skeletal muscle under normal and pathogenic conditions.

Acknowledgements

We are grateful to S. Paturance for carrying out the *in vivo* injection experiments, and to C. Chamot and A. Jobart for assistance with confocal microscopy. We also thank Drs G. Nicolas and MC Lecomte for providing antibodies against spectrin. Drs S. Lefèvre and M. Castañon are gratefully acknowledged for critically reading the manuscript.

This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Universities Paris 6 and 7, and by grants from EC (HPRN-CT-2000-00096) and from the Association Française contre les Myopathies (AFM).

Supplementary Data

Supplementary Table. Exon-Intron boundaries of the rat *Ank3* gene.

Supplementary Figure 1. Sequence of the rat 195aa insert of the C-terminal domain.

Supplementary Figure 2. 5'UTRs of muscle ankyrins-G.

Partial sequences of the canonical (A) and small (B) rat ankyrin-G isoforms obtained by 5'RACE and cDNA library screen. The translation start codon is underlined. Double underlined nucleotides indicate the beginning of cDNA clones; numbers show the nucleotide position relative to the ATG codon. PCR primers are designated by arrows.

Abbreviations

PCR (polymerase chain reaction)

RT (reverse transcriptase)

UTR (untranslated region)

RACE (rapid amplification of cDNA ends)

nr (non-redundant)

EST (expressed sequence tag)

ORF (open reading frame)

GFP (green fluorescent protein)

References

[1] P. Scotland, D. Zhou, H. Benveniste, V. Bennett, Nervous System Defects of AnkyrinB (-/-) Mice Suggest Functional Overlap between the Cell Adhesion Molecule L1 and 440-kD AnkyrinB in Premyelinated Axons, J. Cell Biol. 143 (1998) 1305-1315.

[2] D. Zhou, S. Lambert, P.L. Malen, S. Carpenter, L.M. Boland, V. Bennett, AnkyrinG Is Required for Clustering of Voltage-gated Na Channels at Axon Initial Segments and for Normal Action Potential Firing, J. Cell Biol. 143 (1998) 1295-1304.

[3] S. Tuvia, M. Buhusi, L. Davis, M. Reedy, V. Bennett, Ankyrin-B Is Required for Intracellular Sorting of Structurally Diverse Ca2+ Homeostasis Proteins, J. Cell Biol. 147 (1999) 995-1008.

[4] K. Kizhatil, V. Bennett, Lateral Membrane Biogenesis in Human Bronchial Epithelial Cells Requires 190-kDa Ankyrin-G, J. Biol. Chem. 279 (2004) 16706-16714.

[5] P.J. Mohler, I. Splawski, C. Napolitano, G. Bottelli, L. Sharpe, K. Timothy, S.G. Priori, M.T. Keating, V. Bennett, A cardiac arrhythmia syndrome caused by loss of ankyrin-B function, PNAS 101 (2004) 9137-9142.

[6] W.T. Tse, S.E. Lux, Red blood cell membrane disorders., Br J Haematol 104 (1999) 2-13.

[7] P.J. Mohler, J.-J. Schott, A.O. Gramolini, K.W. Dilly, S. Guatimosim, W.H. duBell, L.-S. Song, K. Haurogne, F. Kyndt, M.E. Ali, T.B. Rogers, W.J. Lederer, D. Escande, H.L. Marec, V. Bennett, Ankyrin-B mutation causes type 4 long-QT cardiac arrhythmia and sudden cardiac death, Nature 421 (2003) 634-639.

[8] E.R. Main, S.E. Jackson, L. Regan, The folding and design of repeat proteins: reaching a consensus, Curr Opin Struct Biol. 13 (2003) 482-489.

[9] V. Bennett, A.J. Baines, Spectrin and Ankyrin-Based Pathways: Metazoan Inventions for Integrating Cells Into Tissues, Physiol. Rev. 81 (2001) 1353-1392.

[10] M. Del Rio, A. Imam, M. DeLeon, G. Gomez, J. Mishra, Q. Ma, S. Parikh, P. Devarajan,

The Death Domain of Kidney Ankyrin Interacts with Fas and Promotes Fas-Mediated Cell Death in Renal Epithelia, J Am Soc Nephrol 15 (2004) 41-51.

[11] T.G. Hall, V. Bennett, Regulatory domains of erythrocyte ankyrin, J. Biol. Chem. 262 (1987) 10537-10545.

[12] L.H. Davis, J.Q. Davis, V. Bennett, Ankyrin regulation: an alternatively spliced segment of the regulatory domain functions as an intramolecular modulator, J. Biol. Chem. 267 (1992) 18966-18972.

[13] P.J. Mohler, A.O. Gramolini, V. Bennett, The Ankyrin-B C-terminal Domain Determines Activity of Ankyrin-B/G Chimeras in Rescue of Abnormal Inositol 1,4,5-Trisphosphate and Ryanodine Receptor Distribution in Ankyrin-B (-/-) Neonatal Cardiomyocytes, J. Biol. Chem. 277 (2002) 10599-10607.

[14] P.J. Mohler, J.A. Hoffman, J.Q. Davis, K.M. Abdi, C.-R. Kim, S.K. Jones, L.H. Davis, K.F. Roberts, V. Bennett, Isoform Specificity among Ankyrins: An amphipathic alpha-helix in the divergent regulatory domain of ankyrin-b interacts with the molecular co-chaperone Hdj1/Hsp40, J. Biol. Chem. 279 (2004) 25798-25804.

[15] B. Modrek, C. Lee, A genomic view of alternative splicing, Nature 30 (2002) 13-19.

[16] V.C. Hou, J.G. Conboy, Regulation of alternative pre-mRNA splicing during erythroid differentiation, Curr Opin Hematol. 8 (2001) 74-79.

[17] P.G. Gallagher, W.T. Tse, A.L. Scarpa, S.E. Lux, B.G. Forget, Structure and Organization of the Human Ankyrin-1 Gene. Basis for complexity of pre-mRNA processing, J. Biol. Chem. 272 (1997) 19220-19228.

[18] E. Kordeli, S. Lambert, V. Bennett, Ankyrin(G), J. Biol. Chem. 270 (1995) 2352-2359.

[19] L.L. Peters, K.M. John, F.M. Lu, E.M. Eicher, A. Higgins, M. Yialamas, L.C. Turtzo, A.J. Otsuka, S.E. Lux, Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin

gene family and the major ankyrin in kidney, is expressed in alternatively spliced forms, including forms that lack the repeat domain, J. Cell Biol. 130 (1995) 313-330.

[20] S. Thevananther, A.H. Kolli, P. Devarajan, Identification of a Novel Ankyrin Isoform (AnkG190) in Kidney and Lung That Associates with the Plasma Membrane and Binds alpha - Na,K-ATPase, J. Biol. Chem. 273 (1998) 23952-23958.

[21] P. Devarajan, P.R. Stabach, A.S. Mann, T. Ardito, M. Kashgarian, J.S. Morrow, Identification of a small cytoplasmic ankyrin (AnkG119) in the kidney and muscle that binds beta I sigma spectrin and associates with the Golgi apparatus, J. Cell Biol. 133 (1996) 819-830.

[22] T.C. Hoock, L.L. Peters, S.E. Lux, Isoforms of Ankyrin-3 That Lack the NH2-terminal Repeats Associate with Mouse Macrophage Lysosomes, J. Cell Biol. 136 (1997) 1059-1070.

[23] B. Peters, H.W. Kaiser, T.M. Magin, Skin-specific expression of ank-3(93), a novel ankyrin-3 splice variant, J Invest Dermatol 116 (2001) 216-223.

[24] C. Gagelin, B. Constantin, C. Deprette, M.-A. Ludosky, M. Recouvreur, J. Cartaud, C. Cognard, G. Raymond, E. Kordeli, Identification of AnkG107, a Muscle-specific Ankyrin-G Isoform, J. Biol. Chem. 277 (2002) 12978-12987.

[25] R.T. Moon, J. Ngai, B.J. Wold, E. Lazarides, Tissue-specific expression of distinct spectrin and ankyrin transcripts in erythroid and nonerythroid cells, J. Cell Biol. 100 (1985) 152-160.

[26] C.S. Birkenmeier, R.A. White, L.L. Peters, E.J. Hall, S.E. Lux, J.E. Barker, Complex patterns of sequence variation and multiple 5' and 3' ends are found among transcripts of the erythroid ankyrin gene, J. Biol. Chem. 268 (1993) 9533-9540.

[27] D. Zhou, C.S. Birkenmeier, M.W. Williams, J.J. Sharp, J.E. Barker, R.J. Bloch, Small, Membrane-bound, Alternatively Spliced Forms of Ankyrin 1 Associated with the Sarcoplasmic Reticulum of Mammalian Skeletal Muscle, J. Cell Biol. 136 (1997) 621-631.

[28] E. Kordeli, M. Ludosky, C. Deprette, T. Frappier, J. Cartaud, AnkyrinG is associated with

the postsynaptic membrane and the sarcoplasmic reticulum in the skeletal muscle fiber, J Cell Sci 111 (1998) 2197-2207.

[29] E. Kordeli, Introduction to organization of the neuromuscular junction: From structure to function, Microsc. Res. Tech. 49 (2000) 1-107.

[30] S.J. Bailey, M.A. Stocksley, A. Buckel, C. Young, C.R. Slater, Voltage-Gated Sodium Channels and AnkyrinG Occupy a Different Postsynaptic Domain from Acetylcholine Receptors from an Early Stage of Neuromuscular Junction Maturation in Rats, J. Neurosci. 23 (2003) 2102-2111.

[31] W.J. Nelson, E. Lazarides, Goblin (ankyrin) in striated muscle: identification of the potential membrane receptor for erythroid spectrin in muscle cells., Proc Natl Acad Sci U S A 81 (1984) 3292-3296.

[32] M.W. Williams, W.G. Resneck, T. Kaysser, J.A. Ursitti, C.S. Birkenmeier, J.E. Barker, R.J. Bloch, Na,K-ATPase in skeletal muscle: two populations of beta-spectrin control localization in the sarcolemma but not partitioning between the sarcolemma and the transverse tubules, J Cell Sci 114 (2001) 751-762.

[33] M.M. Flucher BE, Froehner SC, Daniels MP, Localization of the alpha 1 and alpha 2 subunits of the dihydropyridine receptor and ankyrin in skeletal muscle triads, Neuron 5 (1990) 339-51.

[34] P.G. Gallagher, B.G. Forget, An Alternate Promoter Directs Expression of a Truncated, Muscle-specific Isoform of the Human Ankyrin 1 Gene, J. Biol. Chem. 273 (1998) 1339-1348.

[35] C.S. Birkenmeier, J.J. Sharp, E.J. Gifford, S.A. Deveau, J.E. Barker, An Alternative First Exon in the Distal End of the Erythroid Ankyrin Gene Leads to Production of a Small Isoform Containing an NH2-Terminal Membrane Anchor, Genomics 50 (1998) 79-88.

[36] J.A. Wolff, R.W. Malone, P. Williams, W. Chong, G. Acsadi, A. Jani, P.L. Felgner, Direct

gene transfer into mouse muscle in vivo, Science 247 (1990) 1465-8.

[37] A. Duclert, N. Savathier, J. Changeux, An 83-Nucleotide Promoter of the Acetylcholine Receptor {varepsilon}- Subunit Gene Confers Preferential Synaptic Expression in Mouse Muscle, PNAS 90 (1993) 3043-3047.

[38] G. Nicolas, C.M. Fournier, C. Galand, L. Malbert-Colas, O. Bournier, Y. Kroviarski, M. Bourgeois, J.H. Camonis, D. Dhermy, B. Grandchamp, M.-C. Lecomte, Tyrosine Phosphorylation Regulates Alpha II Spectrin Cleavage by Calpain, Mol. Cell. Biol. 22 (2002) 3527-3536.

[39] A. Bateman, L. Coin, R. Durbin, R.D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E.L.L. Sonnhammer, D.J. Studholme, C. Yeats, S.R. Eddy, The Pfam protein families database, Nucl. Acids. Res. 32 (2004) 138-141.

[40] I. Letunic, R.R. Copley, S. Schmidt, F.D. Ciccarelli, T. Doerks, J. Schultz, C.P. Ponting, P. Bork, SMART 4.0: towards genomic data integration, Nucl. Acids. Res. 32 (2004) 142-144.

[41] L.I. Boue S, Bork P, Alternative splicing and evolution, BioEssays 25 (2003) 1031-1034.

[42] D. Kapfhamer, D.E. Miller, S. Lambert, V. Bennett, T.W. Glover, M. Burmeister, Chromosomal Localization of the AnkyrinG Gene (ANK3/Ank3) to Human 10q21 and Mouse 10, Genomics 27 (1995) 189-191.

[43] P. Bagnato, V. Barone, E. Giacomello, D. Rossi, V. Sorrentino, Binding of an ankyrin-1 isoform to obscurin suggests a molecular link between the sarcoplasmic reticulum and myofibrils in striated muscles, J. Cell Biol. 160 (2003) 245-253.

[44] A. Kontrogianni-Konstantopoulos, R.J. Bloch, The Hydrophilic Domain of Small Ankyrin-1 Interacts with the Two N-terminal Immunoglobulin Domains of Titin, J. Biol. Chem. 278 (2003) 3985-3991.

[45] A. Kontrogianni-Konstantopoulos, E.M. Jones, D.B. van Rossum, R.J. Bloch, Obscurin Is a

Ligand for Small Ankyrin 1 in Skeletal Muscle, Mol. Biol. Cell 14 (2003) 1138-1148.

[46] G. Bezakova, M.A. Ruegg, New insights into the roles of agrin, Nat Rev Mol Cell Biol 4 (2003) 295-309.

[47] X. Zhang, V. Bennett, Restriction of 480/270-kD Ankyrin G to Axon Proximal Segments Requires Multiple Ankyrin G-specific Domains, J. Cell Biol. 142 (1998) 1571-1581.

[48] J.M. Ervasti, Costameres: the Achilles' Heel of Herculean Muscle, J. Biol. Chem. 278 (2003) 13591-13594.

[49] W.J. Nelson, P.J. Veshnock, Ankyrin binding to (Na+ + K+)ATPase and implications for the organization of membrane domains in polarized cells, Nature 328 (1987) 533 - 536.

[50] J.Q. Davis, V. Bennett, The anion exchanger and Na+K(+)-ATPase interact with distinct sites on ankyrin in in vitro assays, J. Biol. Chem. 265 (1990) 17252-17256.

[51] C. Jordan, B. Püschel, R. Koob, D. Drenckhahn, Identification of a binding motif for ankyrin on the alpha-subunit of Na+,K(+)-ATPase, J. Biol. Chem. 270 (1995) 29971-29975.

[52] Z. Zhang, P. Devarajan, A.L. Dorfman, J.S. Morrow, Structure of the Ankyrin-binding Domain of alpha -Na,K-ATPase, J. Biol. Chem. 273 (1998) 18681-18684.

[53] D. Zhou, J.A. Ursitti, R.J. Bloch, Developmental Expression of Spectrins in Rat Skeletal Muscle, Mol. Biol. Cell 9 (1998) 47-61.

Figure Legends

Figure 1. (A) Schematic representation of rat skeletal muscle ankyrin-G isoforms.

Domain organization of the five novel ankyrin-G isoforms and their comparison with brain 270 kDa ankyrin-G and muscle Ank_{G107}. Dotted boxes on top of 270 kDa ankyrin-G indicate the probes used for screening the cDNA library. The three specific sequence features distinguishing all muscle isoforms from non-muscle ankyrins-G are indicated as follows: presence of residues PKI (a), lack of a 12-residue stretch at the end of the spectrin-binding domain (b), and presence of the 76 residue insert in the C-terminal domain (c). Ankyrins are named according to their predicted molecular weight.

(B) Sequence of the novel 17/18aa insert of the C-terminal domain.

Figure 2. Genomic organization of rat *Ank3* sequences.

The exon/intron structures were determined by comparing sequences of rat cDNAs with the genome. Exons are represented by closed boxes and numbered consecutively below (arabic numbers: exons found in muscle transcripts; roman numbers: non-muscle exons). Untranslated sequences are shown in white and coding sequences in black (muscle) or grey (non-muscle). Alternative splicing within exons is indicated by letters (a,b). Asterisks indicate polyadenylation signals. Introns are not drawn to scale.

Figure 3. Sequence (A) and in scale schematic representation (B) of the genomic region surrounding muscle specific exon 25.

Exon sequences are shown in capital or in lowercase (not present in muscle transcripts) letters, intron sequences are in italic, lowercase letters. Written in boldface is the sequence of exon 25. Dinucleotides of splice acceptor/donor sites are underlined. PCR primers are underscored with

arrows. Closed boxes represent muscle (in black) and non-muscle (in grey) coding sequences. Splicing patterns of muscle (solid line) and non-muscle (dotted line) transcripts are shown.

Figure 4. Tissue expression of *Ank3* transcripts.

A rat multiple tissue Northern blot (Origene) was consecutively hybridzed with cDNA probes corresponding to the 195 aa insert (A) and derived from the spectrin-binding domain shared by all ankyrins-G (B). The expression pattern of 195 aa-containing transcripts in kidney, skeletal muscle, and heart is shown in (C), (Clontech, rat multiple tissue Northern blot). Lower panel shows hybridization with control beta actin probe.

Figure 5. In vitro transcription/translation of full-length cDNAs encoding the small muscle ankyrins-G.

cDNAs were cloned in pcDNA3. Products of in vitro transcription/translation were resolved by SDS PAGE, revealed by western blot analysis using the anti-ank$_G$76aa antibody, and compared to endogenous immunoreactive polypeptides in 18 days-old total rat soleus muscle lysates. As a control, Ank$_{G107}$ lacking the 76aa insert (Ank$_{G107}\Delta$76aa), was not revealed by this antibody.

Figure 6. Subcellular targeting of the small muscle ankyrins-G in muscles fibers.

(A) The tibialis anterior of 12 day-old rats was *in vivo* injected with cDNA constructs encoding the GFP-tagged small muscle isoforms of ankyrin-G. Ten days later the recombinant polypeptides (Ank$_{G107}$, Ank$_{G109}$, Ank$_{G128}$, and Ank$_{G130}$) were immunolocalized in muscle cryosections by confocal microscopy using anti-GFP antibodies (green). Sections were double labeled with anti-αII spectrin antibodies (red). All muscle isoforms colocalized with spectrin at the sarcolemma. In control experiments, the lack of the 76aa (Ank$_{G107}\Delta$76aa) resulted in greatly

enhanced cytoplasmic distribution. The green fluorescent protein alone (GFP) remained cytoplasmic.

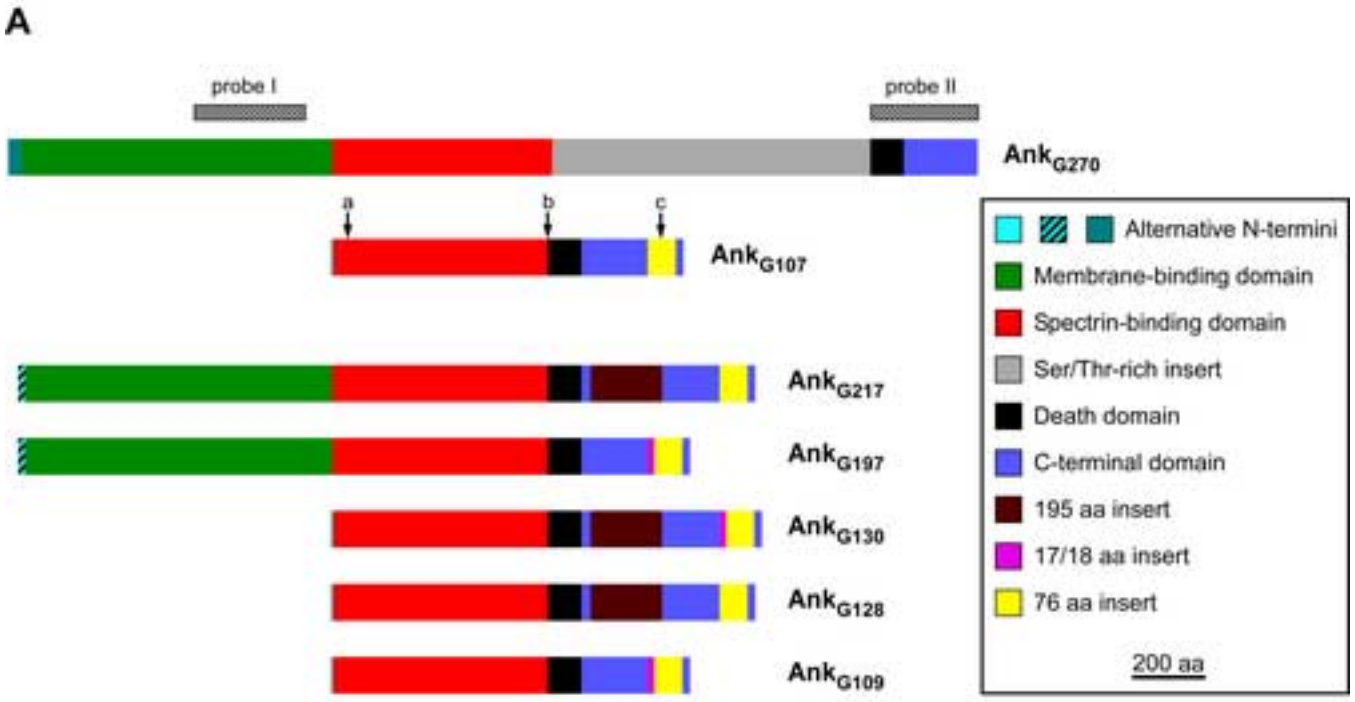
(B) Costameric distribution of muscle ankyrins-G. Confocal superficial optical sections of muscle fibers expressing GFP-tagged Ank_{G107} show a rectilinear costameric pattern of labeling (green, anti-GFP) and partial colocalization of the muscle ankyrin-G with spectrin (red, anti-αII spectrin) at the sarcolemma. Bars: 10 μm.

Table 1
Hopitzan et al.

Table 1
Nucleotide sequences of primers

Primer	Sequence (5'→3')	Orientation
gt1	AGCAAGTCCAGCCTGGTTAAGT	forward
gt2	CTTATGAGTATTTCTTCCAGGGTA	reverse
5′UTRmb-ne1	TATCTTGGGGTAATGACTTGCTGACA	forward
5′UTRmb-ne2	TCTTTGTCGGGGGTTTGGATGTG	forward
13/2-1R	CCGGTGGCACAGGTGTCGTTGTT	reverse
3′UTRn2	AAAAACCCAAAACCCCAAACATG	reverse
5′Malpa	AAAGCATTCCTCAACCCAGCGTG	forward
Malpan2	CAGCGTGGTGGAAATGGCTTTG	forward
5′RACEn3	GGCACCATTGTCCAGAAGAAACCT	reverse
5′RACEn4	AGACCTTGACCACTTCCGCTTGT	reverse
3′Malpa	GCATCTTCACTGTGTGGCAAAGC	reverse
3′Malpan2	CACGCTGGGTTGAGGAATGCTTT	reverse
9/3-1	CGGGAAGAATGCCACAACTGATG	forward
17Aup	GGATGACCGCCTGCTGTTACAAG	forward
9/3-2	ACGAGACGCCAAGTGGAAGCCTA	forward
e24u	GCCATCACAGGGGACACTGACAA	forward
e26r1	GCTCCTCGATCATCATGCTGTCC	reverse
e8up	ATCCCGATTCTGTCAGCCATTTCT	reverse
e8lo	ACCATATTTGCGTTTCCCCGTTTC	forward

Figure 1



B

17/18 aa insert

1	CAG	GTC	AAA	AGC	CCG	GGC	GAA	GCG	TTC	ACT	CGG	ATG	ACC	GCC	TGC	TGT	TAC	AAG	54
1	Q	V	K	S	P	G	E	A	F	T	R	M	T	A	C	C	Y	K	18

Figure 1
Hopitzan et al.

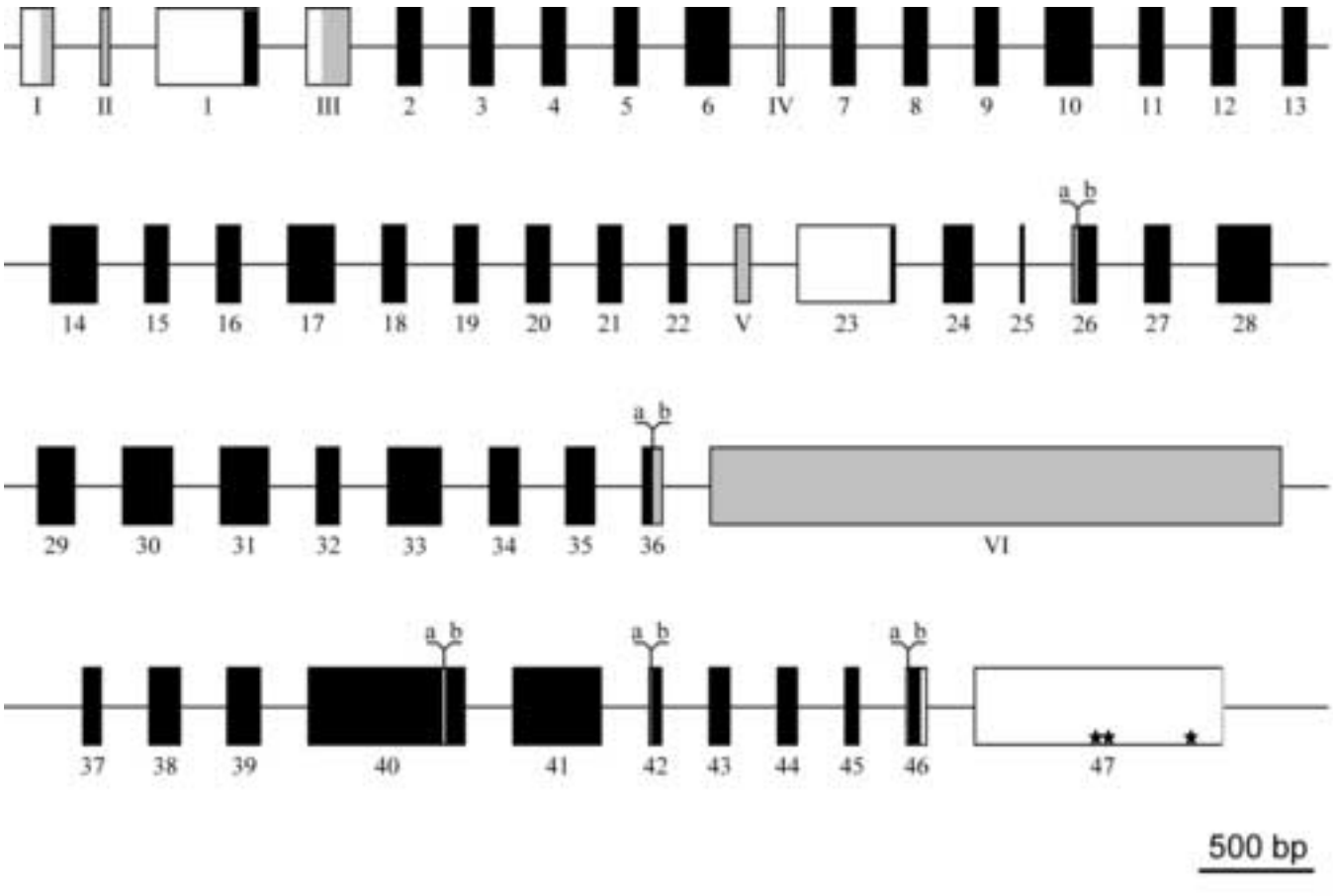


Figure 2
Hopitzan et al.

Figure 3

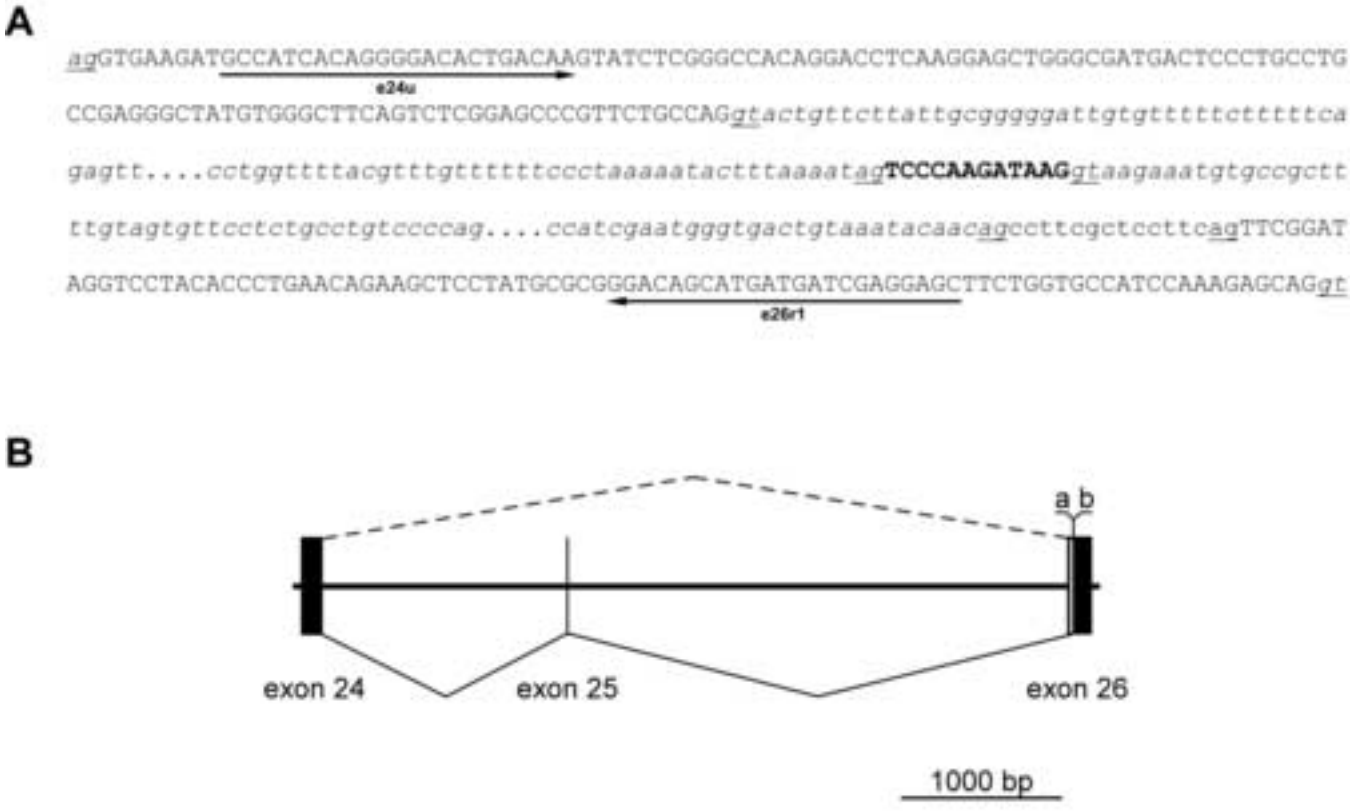


Figure 3
Hopitzan et al.

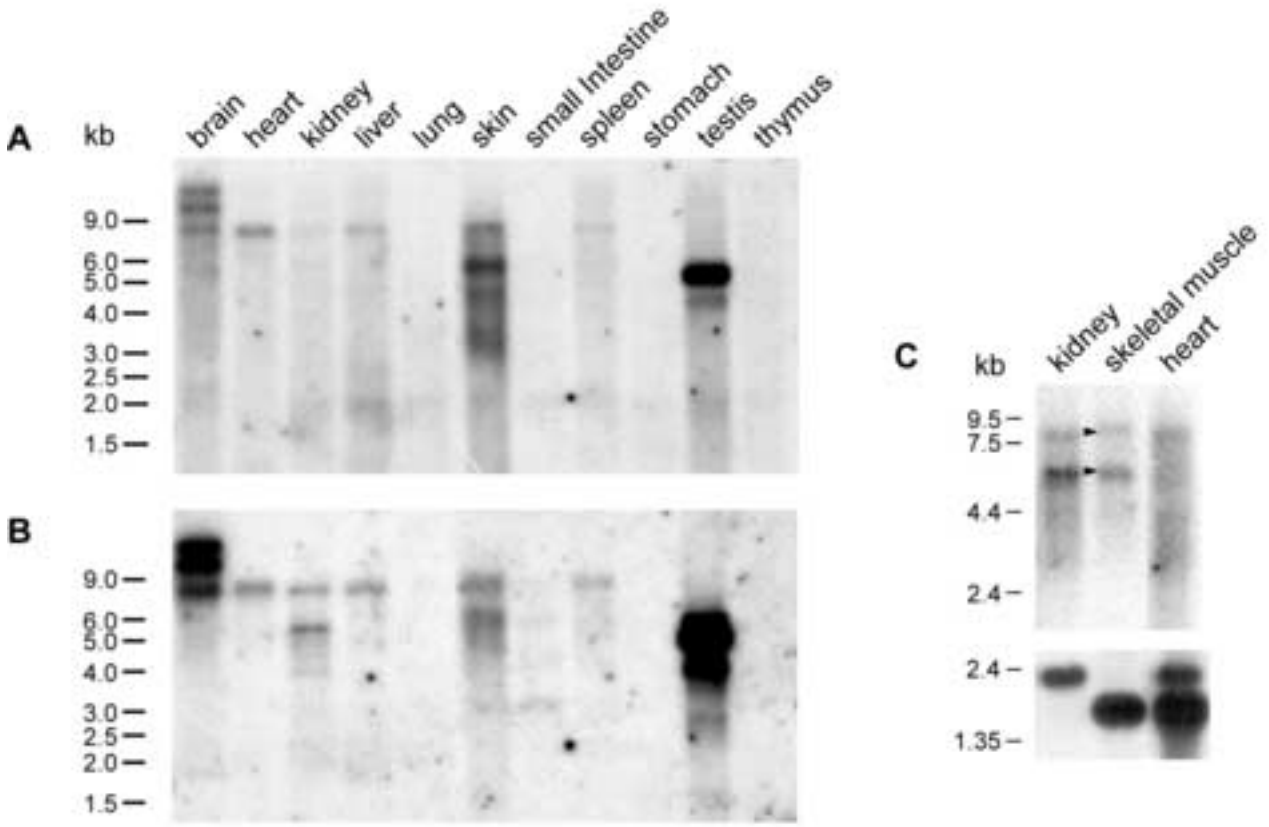


Figure 4
Hopitzan et al.

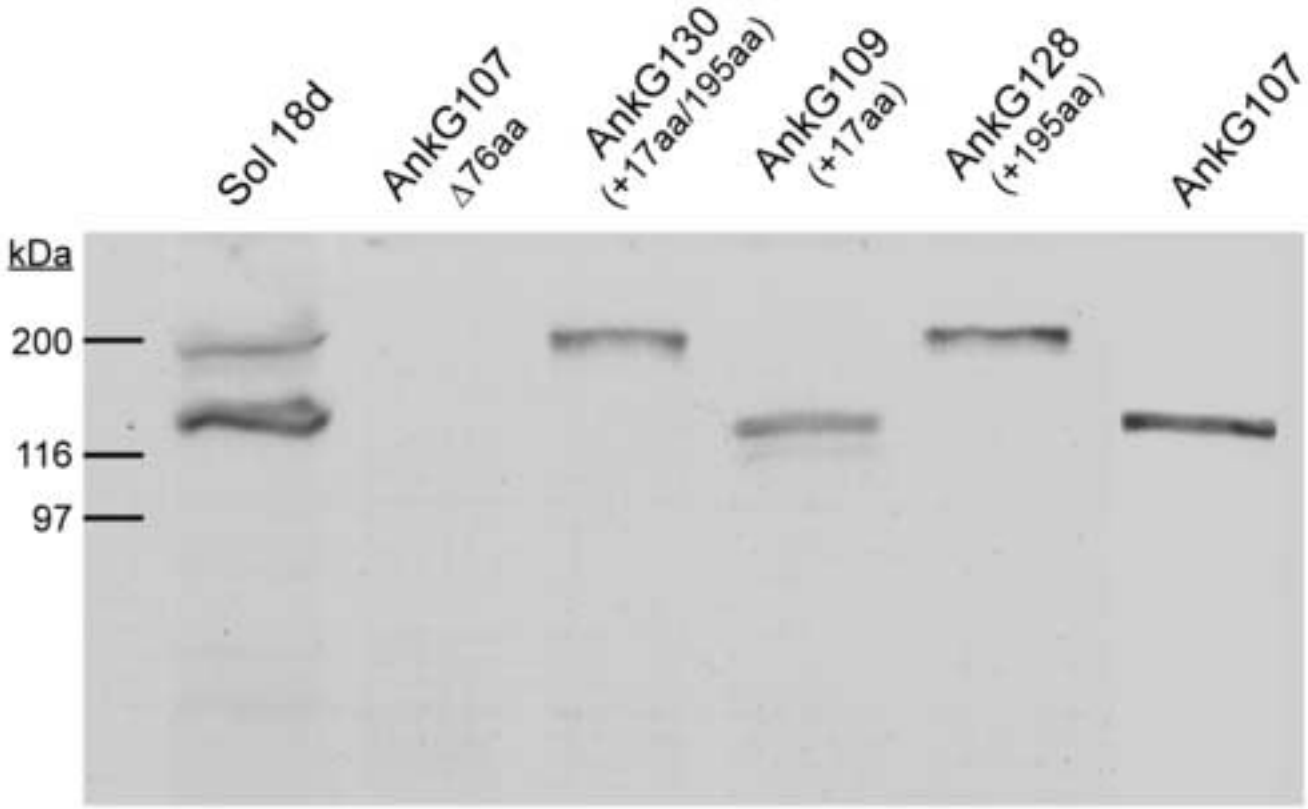


Figure 5
Hopitzan et al.

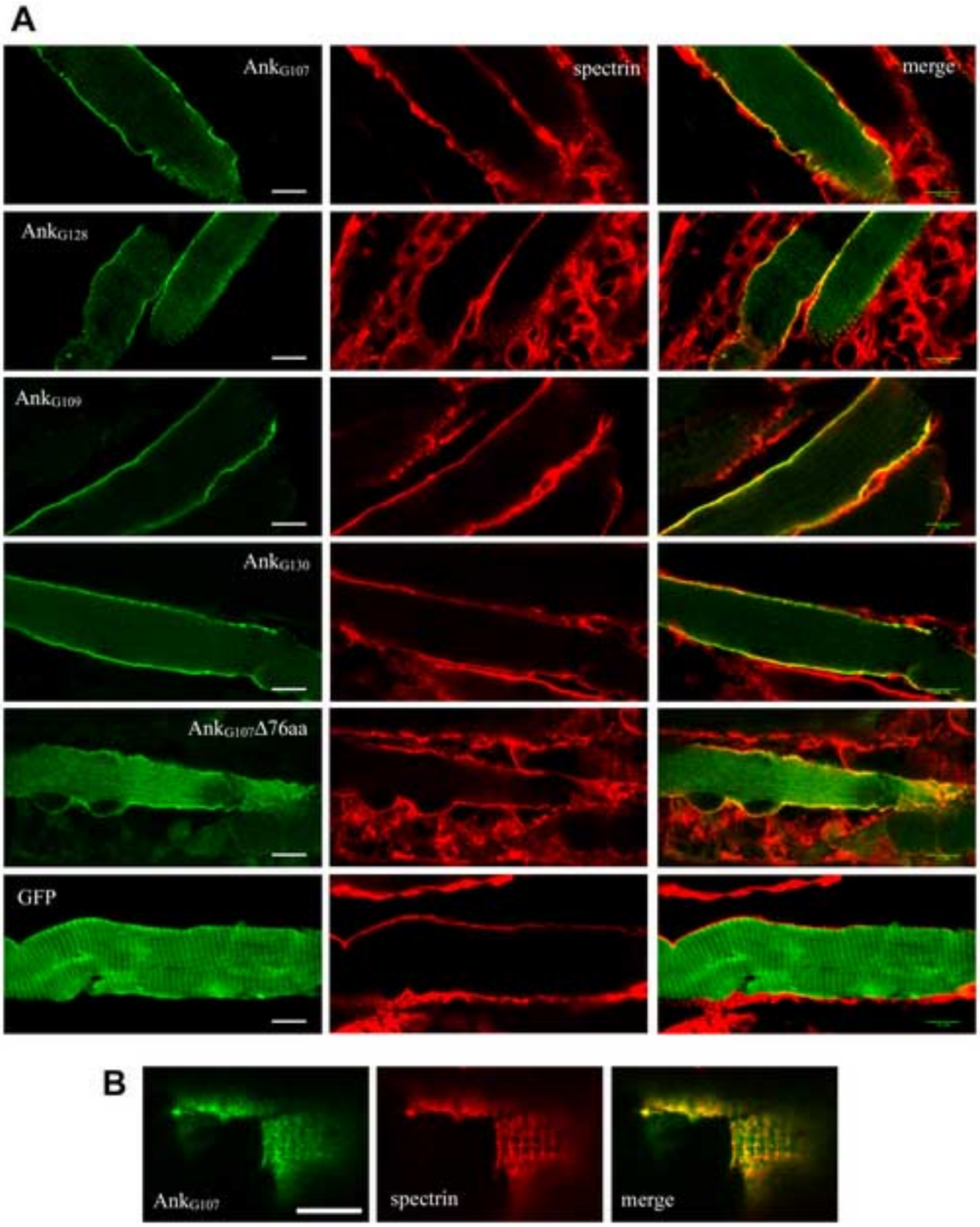


Figure 6
Hopitzan et al.