Molecular cloning and functional expression of a novel human gene encoding two 41–43 kDa skeletal muscle internal membrane proteins

Sophie BOUJU*, Marie-Françoise LIGNON*, Geneviève PIÉTU†, Martine LE CUNFF*, Jean-J. LÉGER*, Charles AUFFRAY† and Claude A. DECHESNE*[1]

*Laboratoire de Physiopathologie Cellulaire et Moléculaire, INSERM Unité 300, Faculté de Pharmacie, 34060 Montpellier Cedex, France, and †Genexpress, Génétique Moléculaire et Biologie du Développement, CNRS UPR 420, 19 rue Guy Mocquet, 94801 Villejuif, France

Systematic analysis of gene transcript repertoires prepared from libraries made with various specific human tissues permitted isolation of many partially sequenced cDNA clones. A few of these represented novel genes with limited or no similarity to known genes from humans or other species. The present study set out to isolate and sequence the full-length cDNA corresponding to one of these novel human transcripts, and identify the corresponding protein product at the subcellular level. Current sequence analyses have revealed that the protein contains a hydrophobic N-terminal segment and an internal leucine-zipper motif. Numerous sites of putative post-translational modifications, such as N-linked glycosylation, myristoylation and phosphorylation sites, were also identified. Using one monoclonal antibody raised against a recombinant fragment, two different 41–43 kDa proteins were detected in human skeletal muscle, heart and placenta homogenates at various ratios. Both immunodetected protein products of the novel human gene were distributed in the transverse tubules and/or near the junctional sarcoplasmic reticulum within skeletal muscle cells. Both proteins had physical properties believed to be attributable to integral membrane components. Finally, the GENX-3414 gene was chromosomally localized at position 4q24–q25.

INTRODUCTION

Over the last two decades, there have been successive efforts to obtain information on the sequences, chromosomal positions and expression patterns of about 80000 genes that are expressed in humans. Following initial physical and genetic mapping studies on the human genome, gene transcript repertoires have been systematically analysed from partially sequenced cDNA clones isolated from libraries derived from a variety of specific human tissues [1]. Many of these transcripts are novel genes with limited or no similarity to known genes from humans or other species. More recently, human gene transcripts were classified on the basis of their tissue-specific expression and other selective criteria, such as cellular differentiation or disease features. Hybridization signatures from many gene transcripts were thus collected by analysing arrayed cDNA clones, dotted on to a membrane at high density using a robotic device, with complex cDNA probes obtained by reverse transcription of mRNA from different tissues [2,3].

The present study represents the step following identification of tissue-specific gene transcripts, within the general context of the IMAGE (Integrated Molecular Analysis of the human Genome and its Expression) consortium [4,5]. It focuses on a novel gene with a high expression level in human striated muscle, as revealed by quantitative and differential hybridization of high-density cDNA with probes derived from nine different human tissues. We first set out to isolate and sequence the full-length cDNA. Current sequence analyses have revealed that the protein contains a hydrophobic N-terminal segment, an internal leucine-zipper motif and numerous sites for putative post-translational modification. One monoclonal antibody (mAb) raised against a recombinant protein fragment highlighted the fact that the human gene encodes two immunorelated proteins with apparent molecular masses of 41 and 43 kDa. The subcellular membrane localization of both proteins was assessed through confocal microcopy analysis of cryostat sections of human skeletal muscle and by their immunodetection in muscle-membrane-enriched preparations. According to currently available information, both protein products of the novel human gene are distributed in the transverse tubules and/or near the junctional sarcoplasmic reticulum within skeletal muscle cells. The observed membrane subcellular localization is consistent with sequence analysis data. Chromosome mapping indicated that the gene is mapped within the 4q24–q25 chromosomal band.

MATERIALS AND METHODS

Northern blots

Human multiple tissue Northern blots, containing 2 μg of pure polyadenylated [poly(A)$^+$] (RNA) per lane, were hybridized with [α-^{33}P]dATP-labelled bu-07b12 (nt 683–2334) cDNA, prepared by random priming, and purified on Sephadex G-50 columns, according to the manufacturer's instructions (Clontech, Palo Alto, CA, U.S.A.). The transcript sizes were estimated by comparison with the size of myosin heavy chain, glyceraldehyde-3-phosphate dehydrogenase, 28 S and 18 S ribosomal transcripts. The hybridization signals were detected on phosphorus screens after 5 h exposure with the PhosphoImager imaging plate system (Molecular Dynamics, Sunnyvale, CA, U.S.A.).

Abbreviations used: cM, centimorgan; DHPR, dihydropyridine receptor; EST, expressed sequence tag; mAb, monoclonal antibody; ORF, open reading frame; poly(A)$^+$, polyadenylated; RH, radiation hybrid; RT-PCR, reverse transcriptase-PCR; RyR, ryanodine receptor; STS, sequence-tagged site; T-tubule, transverse tubule.

[1] To whom correspondence should be addressed (e-mail cdechesn@balard.pharma.univ-montp1.fr).

The nucleotide sequence results reported here will appear in the EMBL and GenBank nucleotide sequence databases under the accession number AF062534.

Cloning of the full-length human GENX-3414 cDNA

To clone the full-length GENX-3414 cDNA, nested PCR was used to obtain a sequence extension based on a non-oriented λgt11 phage cDNA library made from normal whole skeletal thigh muscle from a 45-year-old Caucasian male (Clontech #HL5002b). The template used was bacteriophage cDNA (50 ng) prepared from 200000 plaques of the library [6]. For the nested PCR, two different forward primers were used, chosen in the λgt11 left arm (5′-TTCAACATCAGCCGCTACAG-3′ and 5′-ACTGATGGAAACCAGCCATC-3′) and the two respective reverse primers (R1 and R2) in clone bu-01d05 (R1, 5′-GGCAGGAAAATGGAATGAGA-3′; R2, 5′-CCGGTCCTGACTGTTCAAT-3′). After a 2 min denaturation at 95 °C, both amplifications were performed with Ampli *Taq* polymerase (Perkin–Elmer, Courtaboeuf, France) using a 30 cycle reaction on a Biometra DNA thermal cycler (95 °C for 1 min, 55 °C for 1 min, and 72 °C for 1 min). The second amplification was performed using 1% of the volume of the first PCR product. Several reverse transcriptase (RT)-PCRs, using the Titan one tube RT-PCR kit (Boehringer, Mannheim, Germany), were then performed on total RNAs prepared from various human deltoid muscles from different subjects, using as forward primer (F1) 5′-GCAGAAACCGTAAAAGGCTG-3′, and the R1 reverse primer. A full-length cDNA sequence was finally obtained by a unique one-step PCR amplification using the 5′-GGTGGCGACGACTCCTGGAGCCCGTCAGTATCG-3′ forward sequence in the λgt11 left arm and F1 as the reverse primer. The resulting PCR product was cloned in pBluescript vector. All PCR products were purified with the QIAquick Spin kit (Qiagen, Courtaboeuf, France), and sequenced.

Sequencing and sequence analysis

Fluorescence-based DNA sequencing was performed by Genome Express S.A. (Grenoble, France). cDNA, translated sequences, expressed sequence tags (ESTs) and sequence-tagged sites (STSs) were compared with sequences in public nucleotide and protein databases using BLAST and FASTA programmes via the National Center for Biotechnology Information WWW server [7]. Sequences were aligned with ClustalW using the Baylor College of Medicine server. Protein motifs were explored with the ProfileScan programme in the Prosite library through the Swiss Institute for Experimental Cancer Research server [8].

Production of mAbs against the GENX-3414 protein products

A chimaeric protein was produced by cloning the bu-07b12 cDNA into the pMal-c2 plasmid (New England Biolabs, Beverly, MA, U.S.A.), consisting of an N-terminal maltose-binding protein fused to 173 amino acids of the sequence of interest, and expressed in the TB1 *Escherichia coli* strain. The recombinant component was purified by affinity chromatography using an amylose resin column (New England Biolabs) and used as an immunogen to obtain mAbs [9]. The procedures for mouse immunization, fusion, screening, production and purification of antibodies are described previously [10]. The mAb against the GENX-3414 protein products (hereafter called mAb 4D11) was selected for immunoblots and immunofluorescence experiments.

Other mAbs

The following antibodies were diluted as follows: anti-dystrophin mAb 5G5 [11] was diluted 1:1, anti-α-actinin (Boehringer) 1:800, anti-α1-dihydropyridine receptor (DHPR) IIF7 [12] 1:100 and anti-triadin IIG12 A8 [13] 1:50. Both of the latter antibodies were generously given by Dr. K. P. Campbell (University of Iowa, Howard Hughes Medical Institute, Iowa City, IA, U.S.A.). The antibodies were diluted in 5% (v/v) non-fat milk and reacted for 30 min at 37 °C with the Western-blotting membranes. Secondary antibodies, conjugated with alkaline phosphatase at 1:1000 dilution, were purchased from Euromedex (Souffelweyersheim, France).

Immunoblots and immunofluorescence

Human (deltoid) skeletal muscle from a 27-year-old male was obtained by orthopaedic surgery for non-muscle-related reasons. A myocardial left-ventricle biopsy sample was obtained from a transplanted heart of a 53-year-old patient with ischaemic cardiomyopathy. Placenta was obtained from a 24-year-old woman after delivery. For immunoblot analysis of total tissue extracts, 20 serial 10-μm-thick cryostat sections, adjacent to the sections sampled for the immunofluorescence studies, were homogenized in 40 μl of gel-loading SDS-containing buffer, according to the Western-blotting microprocedure described by Dechesne et al. [10]. Human skeletal muscle (7.5 μl), heart (5 μl) and placenta (5 μl) were analysed with SDS/PAGE (4–15% gradient gels). Transfer and immunostaining procedures, using the anti-mouse mAb 4D11 raised against the GENX-3414 protein products, were performed by a standard procedure [10]. Alkaline-phosphatase-labelled goat anti-mouse antibodies were used for immunolabelling. Subcellular localization of the studied protein was carried out on 10–12 μm cryostat sections of unfixed human skeletal muscle. Single labelling of muscle tissue sections involved a first incubation with pure supernatant specific mAb, raised against either the GENX-3414 protein products (mAb 4D11) or anti-(fast human myosin heavy-chain) mAb [10], followed by incubation with a Cy3-labelled anti-mouse serum (1/1000 dilution). Confocal laser microscopy analyses were performed with a Leica TCS 4D laser confocal microscope.

Isolation of human skeletal muscle membranes and biochemical analysis

KCl-washed light microsomes were prepared from approx. 5 g of human (deltoid) skeletal muscle from a 27-year-old male, as described by Ohlendieck et al. [14]. Occasionally, the microsomal preparation was further fractionated in crude surface-membrane vesicles obtained at the 0.303 M/0.878 M sucrose interface following ultracentrifugation in an SW41 swinging-bucket rotor (Beckman Instruments) at 160000 ***g*** for 17 h. The light microsome and crude surface-membrane preparations were collected, diluted with buffer A [0.3 M sucrose/0.6 M KCl/20 mM Tris/maleate (pH 7.0)] and centrifuged in a 50Ti rotor for 30 min at 145000 ***g***. All preparations were made in the presence of antiproteases (Sigma Cocktail, St Quentin Fallavier, France). All membrane preparations were resuspended in buffer B [0.303 M sucrose/20 mM Tris/maleate (pH 7.0)] and stored frozen at −20 °C. For KCl washing, alkaline treatment and Triton X-100 extraction, human KCl-washed light microsomes in 50 mM Tris/HCl, pH 7.5, and 5% (w/v) sucrose were further treated, as described previously [15,16]. After 1 h incubation at room temperature, samples were centrifuged for 1 h at 100000 ***g*** using a Beckman Airfuge. The protein contents of the supernatants and pellets were analysed by SDS/PAGE (4–15% gradient gels); their GENX-3414 protein-product content was immunoanalysed with the specific mAb 4D11.

Refined mapping

Mapping was performed according to three frameworks: the Whitehead radiation hybrid (RH) Map (July 97), the RH

Consortium Gene Map 96 (to be updated with the Gene Map 98), both of which were deduced from mapping experiments performed on the Genebridge-4 RH panel [17–19], and the G3 RH map, built at the Stanford Human Genome Center using the G3 RH panel [20]. Integration within the cytogenetic map was established via the genetic markers identified in the most immediate vicinity to the gene of interest. The genetic markers were cytogenetically localized through the genome database.

RESULTS

Identification of a novel human gene transcript highly expressed in striated muscle

A cDNA clone corresponding to a novel gene, registered in the Genexpress index as GENX-3414 [3], was derived from a 19-year-old female human skeletal muscle cDNA library. The PCR-amplified inserts from a set of 1091 human skeletal muscle cDNA clones of this library, representing more than 800 human gene transcripts, were initially arrayed on high-density filters hybridized in parallel with total cDNA probes prepared from various human tissues: brain, adult and fetal heart, liver, placenta, skeletal muscle, testis, thymus and uterus [3]. On the basis of quantitative comparisons made among hybridization signals detected on phosphorus screens, cDNAs hybridizing preferentially with the muscle probes were selected and confirmed by Northern blot analysis. Of the original set, 48 cDNA clones were found with an expression pattern either restricted to or highly specific for striated muscle. After partial sequencing, 41 of the clones were identified as corresponding to novel gene transcripts, i.e. with no detectable database homology.

A 1651-bp cDNA clone, a partial copy of the full-length GENX-3414 transcript, was selected from the cDNA collection on the basis of two criteria: its novel nucleotide sequence and its high expression level in human skeletal muscle. Another 2117-bp cDNA clone, sharing the same 1651-bp 3′-end sequence as clone bu-07b12, was also obtained from the same cDNA collection. Human multiple-tissue Northern blots hybridized with cDNA clone bu-07b12 revealed a mRNA migrating at 2.4 kb in different human tissues. It was present at very high levels in human skeletal and cardiac muscles, and at moderate levels in liver and placenta. No hybridization with other human tissue mRNAs from pancreas, kidney, liver and lung was detected (Figure 1).

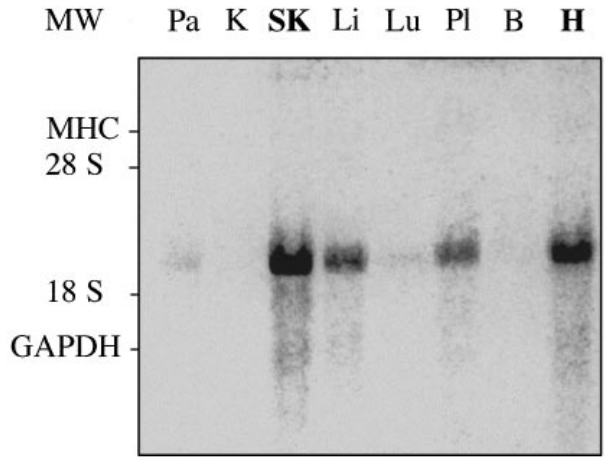

Figure 1 Tissue-specific expression of the novel human GENX-3414 gene

Each lane contained 2 μg of poly(A)$^+$ RNAs from pancreas (Pa), kidney (K), skeletal muscle (SK), liver (Li), lung (Lu), placenta (Pl), brain (B) and heart (H). They were probed with a ^{33}P-labelled human probe generated from the cDNA clone bu-07b12 (nt 683–2334). The positions of the RNA molecular-mass markers are indicated on the left: MHC, myosin heavy chain; GAPDH, glyceraldehyde-3-phosphate dehydrogenase. A single 2.4 kb mRNA band was present in the different human tissues analysed, at the highest levels in skeletal and cardiac muscles.

Isolation and sequencing of the full-length cDNA for the GENX-3414 transcript

Sequencing of both cDNA clones available initially (bu-07b12 and bu-01d05) revealed that clone bu-01d05 extended 466 bp upstream from the 5′-end of the bu-07b12 clone. The two cDNAs terminated with an identical 3′-end containing a poly(A)$^+$ tail, 20 bp downstream of the consensus AATAAA polyadenylation signal. Clone bu-01d05 contained a single open reading frame (ORF) of 328 amino acids, starting at the second nucleotide and ending with a stop codon 984 bp downstream. Further upstream, a 5′-sequence was obtained with PCR extension of cDNA templates from a non-oriented λgt11-phage cDNA library made from normal whole skeletal thigh muscle of a 45-year-old Caucasian male. Nested PCR, using two λgt11-phage left-arm-specific forward primers with two different bu-01d05 clone-specific reverse primers (R1 and R2 successively), gave rise to one cDNA product of around 700 bp. Sequencing studies revealed that it extended 216 bp from the 5′-end of clone bu-01d05, and shared the identical 502-bp 3′-sequence. Then using a forward primer in the newly extended sequence (F1) and the reverse primer (R1) used previously, a few one-step RT-PCRs were performed using total RNAs derived from muscle of 12 different subjects. Their sequences were identical with that previously detected in the initial nested PCR product. Finally, using one primer in the λgt11-phage left arm and a primer in the 5′-end of the nested PCR product (primer F1), one PCR product of about 2.5 kb in length was obtained from the human skeletal muscle phage library, which was subsequently cloned and fully sequenced.

Nucleotide and/or primary protein structure analysis

Sequence assembly of the bu-01d05 clone with the different PCR products led to a unique 2334-bp-long cDNA sequence for the GENX-3414 gene (Figure 2). The ORF sequence observed in clone bu-01d05 was thus extended in the N-terminal direction up to the potential initiation codon (ATG) encoding the first methionine. The sequence around this ATG fitted poorly with the Kozak translation initiation consensus sequence, but contained the required purine in position -3 [21]. The sequence upstream of the ATG codon did not contain a potential 3′-splice site between the first in-frame stop codon and ATG. These different features suggest that this ATG is the translation initiation methionine codon. The sequenced transcript was close to the 2.4 kb size of the mRNA band (Figure 1), which was unique even after a lengthy Northern blot exposure. The 1074-bp ORF, extending from the initial codon at nt 128 to the stop codon at nt 1202, encoded a putative protein of 358 amino acid residues with a predicted unmodified M_r of 39007. The complete nucleotide sequence of the human GENX-3414 cDNA characterized in this study will be deposited in the GenBank/EMBL data bank (accession number AF062534).

The homology investigations confirmed that the GENX-3414 transcript, as well as the encoded protein, presented no significant sequence similarities with any gene or protein registered in the present GenBank and SwissProt databases. Alignment of the full-length GENX-3414 nucleotide sequence with different ESTs from the dbEST data bank [7] showed sequence homologies that

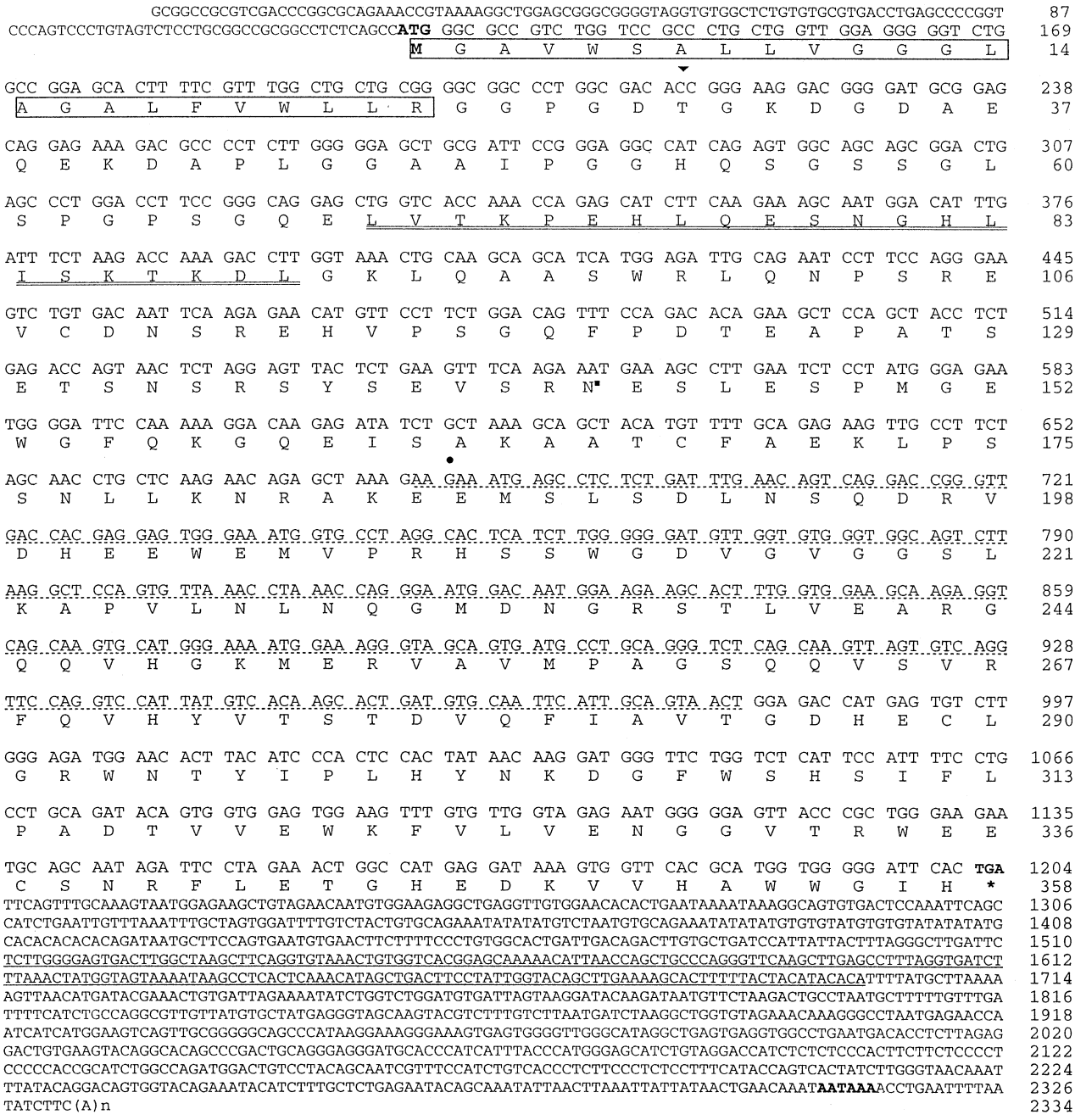

Figure 2 Sequence of the human full-length GENX-3414 gene transcript and the deduced sequence of the encoded protein

The initiation codon (ATG) at nt 128, and the polyadenylation signal (AATAAA) at the 3′-end are in bold typeface. The stop codon is represented by an asterisk; the 24 hydrophobic N-terminal amino acids are boxed, the leucine-zipper motif sequence is double-underlined, the STS sequence #STSG-63501 is dot-underlined and the STS sequence #WI-17734 is single-underlined. ▼, the first nucleotide of the clone bu-01d05 (nt 217); ●, the first nucleotide of the clone bu-07b12 (nt 683); ■, the predicted N-linked glycosylation site (amino acid 143). The sequence of human GENX-3414 cDNA has been deposited in GenBank under accession number #AF062534.

were close to human and mouse ESTs in the coding and long non-coding 3′-part of the sequence. One EST sequence derived from human placenta (referred to as #T47519) matched (98% homology) 408 nucleotide residues around the 5′-end of the longest GENX-3414 transcript, encoding a placental 117-amino-acid-long sequence identical with the human skeletal muscle N-terminal sequence. The amino acid content of the predicted protein encoded by the GENX-3414 transcript revealed its slightly acidic nature with a pI of 5.46. The primary protein structures were scanned for consensus motifs or sites using the Prosite command in the ProfileScan [8]. One leucine-zipper motif spanning from residues 69–90 was identified (Figure 2). One potential N-linked glycosylation site (residue 143) and five different potential myristoylation sites were identified in the predicted amino acid sequence (Figure 2). The putative protein contained multiple potential phosphorylation sites, including four protein kinase C sites, and 12 casein kinase II sites (not shown). Hydropathy analysis of the predicted amino acid se-

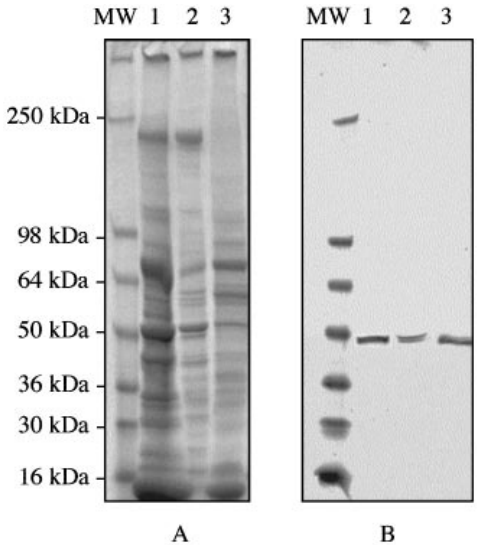

Figure 3 Immunodetection of native GENX-3414 protein products in different human tissues

Human extracts from skeletal muscle (lane 1), heart (lane 2) and placenta (lane 3) were analysed with SDS/PAGE (4–15% gradient gels). They were either stained with Coomassie Blue (**A**) or analysed by Western blotting (**B**) with the specific mAb 4D11 raised against a recombinant fragment encoded by a partial GENX-3414 cDNA. The apparent molecular masses of the double band, observed at differing ratios in the three human tissues analysed, are 41–43 kDa. Positions of the molecular-mass markers are indicated on the left.

quence was performed using an algorithm based on the properties of each individual amino acid [22]. Using a 17-residue window size, one hydrophobic stretch was detected at the N-terminal tip of the protein (results not shown); this comprised the first 24 residues, which could function as a potential domain for membrane fixation. The rest of the protein seemed to be very hydrophilic.

Immunoblot and immunofluorescence analysis of the novel human protein

A fusion protein in which a protein fragment corresponding to the 173-amino-acid-long C-terminal part of the novel protein encoded by GENX-3414 fused with the maltose-binding protein, was produced by genetic engineering. The corresponding recombinant protein fragment, which was recognized as a band with an apparent molecular mass of 62 kDa (results not shown), was used to produce one mouse mAb (called mAb 4D11) specific for the novel protein. Total homogenates (in SDS-containing buffer) of human (deltoid) skeletal muscle, myocardial ventricle and placenta were separated by SDS/PAGE (4–15% gels) and analysed by Western blotting using mAb 4D11 (Figure 3). One or two bands of about 41–43 kDa were immunodetected in the analysed human total extracts; they migrated slightly below the intense actin band. Both bands were present at about the same intensity in the skeletal and cardiac extracts, whereas the 41 kDa band was essentially present in placental extracts. Longitudinal and transversal cryostat sections of human skeletal deltoid muscle were processed with mAb 4D11. Examination of several transverse sections with optical microscopy at regular magnification showed discontinous immunostaining within the cytoplasm and no staining at the periphery (sarcolemmal membrane) of muscle cells. A close examination of the same immunolabelled transverse sections, on reconstituted confocal microscopy images, showed a

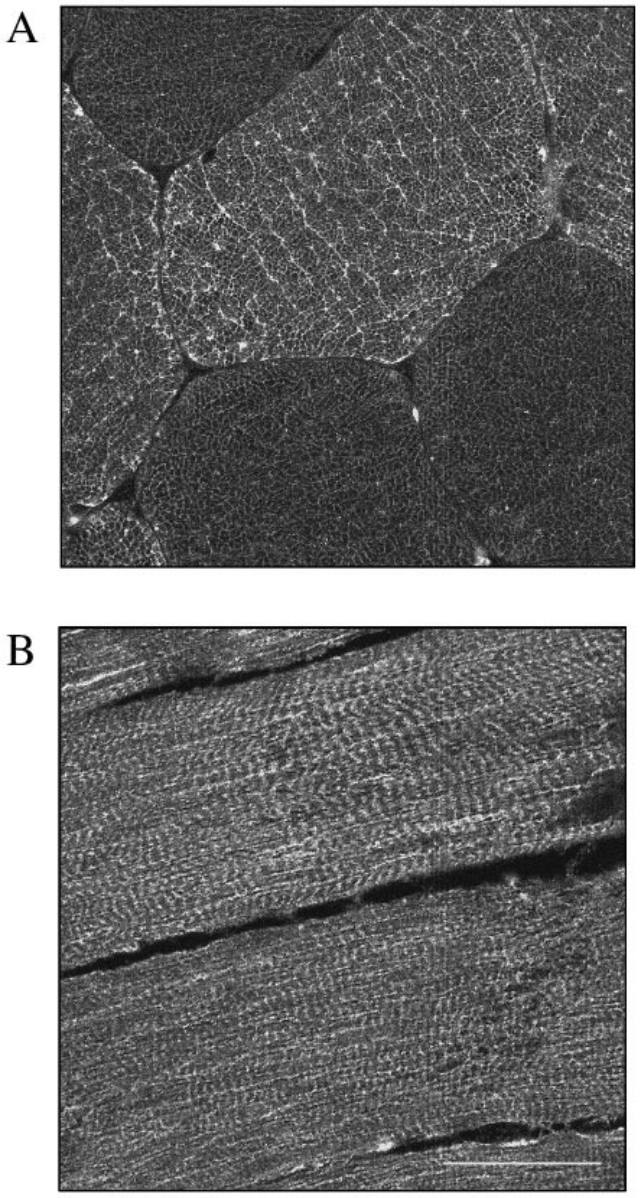

Figure 4 Subcellular localization of GENX-3414 protein products in internal muscle membranes

Transversal (**A**) and longitudinal (**B**) cryostat sections of (deltoid) skeletal muscle of a 27-year-old male, stained with the specific mAb 4D11 raised against a recombinant fragment encoded by a partial GENX-3414 cDNA, and analysed by confocal microscopy. The reacting mAb was identified with a Cy3 dye-labelled anti-mouse serum. Bar = 50 μm. A regular polygonal-staining pattern was observed throughout the myofibre cytoplasm. The internal staining pattern of mAb 4D11 was more intense in fast than in slow fibres, which agrees with the higher abundance of T-tubules in fast fibres compared with slow fibres.

regular polygonal staining pattern throughout the myofibre cytoplasm (Figure 4A). Although mAb 4D11 internally stained all muscle fibres, fast-twitch (type II) fibres were stained more intensely than slow-twitch (type I) fibres, as detected by specific anti-fast and anti-slow myosin heavy-chain antibodies (results not shown). The immunofluorescence-staining pattern observed in longitudinal cryosections appeared as transversely oriented rows of continuous staining, and occasionally as longitudinal

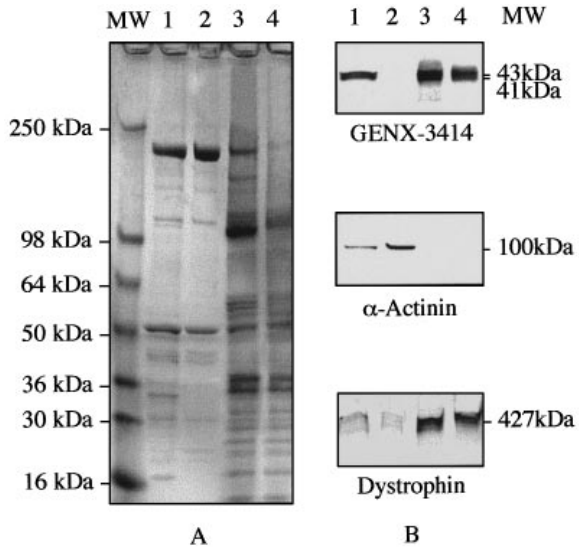

Figure 5 Comparative immunoblot analysis of GENX-3414 protein products in different human skeletal muscle preparations

Human muscle membranes were prepared as described by Ohlendieck et al. [14], and separated with SDS/PAGE (4–15% gradient gels), followed by immunoblotting with different mAbs. A Coomassie-Blue-stained gel (**A**) and three similar immunoblots (**B**) are shown, labelled with either the specific MAb 4D11 raised against a recombinant fragment encoded by a partial GENX-3414 cDNA or one mAb against α-actinin, or a mAb against human dystrophin respectively. Lanes 1–4 consisted of total muscle extracts, purified myofibrils, KCl-washed light microsomes and crude surface-membrane vesicles from human deltoid muscle respectively (about 35 μg of protein/lane). The molecular-mass standards are indicated on the left of the Coomassie-Blue-stained gels, and on the right of the Western blots. High concentrations of the GENX-3414 protein products were immunodetected in both muscle membrane-enriched preparations, but not in the purified myofibrillar preparation.

traces (Figure 4B). The row distributions corresponded to the interface between the A-band and the I-band (results not shown).

Membrane topology analysis

The membrane topologies of both novel proteins were further assessed by their immunodetection in different membrane preparations, compared with their absence in myofibrillar preparations of the same human skeletal muscle. As shown in Coomassie-Blue-stained gels in Figure 5(A), preparations of KCl-washed light microsomes (lane 3), as well as of crude surface-membrane vesicles (lane 4) from a few grams of human skeletal muscle, contained sarcoplasmic-reticulum markers, e.g. the 109 kDa calcium pump and transverse (T)-tubules, as shown by the double 36-kDa band [13,23–26]. By comparison, these membrane markers were not (or only slightly) present in myofibrils (lane 2) prepared from the same human muscle biopsy; the main myofibrillar markers were the 200 kDa myosin heavy chains and 43 kDa actin (Figure 5A, lane 2). Using mAb 4D11, both GENX-3414 protein products were clearly enriched in both membrane preparations when compared with the initial total extract (Figure 5B). They were not detectable in the myofibrillar preparations (Figure 5B, lane 2). Although a few Coomassie-Blue-stained bands with molecular masses close to those of the studied proteins were visible in the membrane-enriched preparations (Figure 5A), none of them could be conclusively identified as being either of the related proteins.

Other muscle markers were also used to further characterize the different muscle preparations. Anti-α-actinin antibody revealed that this sarcomeric marker was immunodetected in the muscle myofibrillar preparations only, and not in the enriched membrane preparations where the GENX-3414 protein products were more readily detected. In contrast, anti-dystrophin antibody [11] revealed that this sarcolemmal (external) membrane marker was highly present in both human muscle membrane-enriched samples, but not in the myofibrillar preparations (Figure 5B). Two other mAbs against markers of T-tubules, such as mAb IIF7, an antibody raised against the 170 kDa α1 subunit of skeletal DHPR [12], or of triads, such as mAb IIG12A8, an antibody against the 94 kDa junctional-specific glycoprotein (triadin) [13], were tested by both immunofluorescence analysis of human muscle cryostat sections and Western blotting experiments with the human muscle-membrane-enriched preparations. The anti-DHPR antibody reacted with human tissue in a similar manner to mAb 4D11, showing the same regular polygonal staining pattern throughout the myofibre cytoplasm (Figure 4A). Unfortunately, antibodies reacted either poorly or not at all with the human membrane extracts. Considering the relative specificity of the antibodies known previously [12,13], new membrane preparations similar to the human preparations (but with more concentrated protein contents) were made from more than 100 g of rabbit skeletal muscle. Under the same experimental conditions, mAb IIF7, mAb IIG12A8 and mAb 4D11 reacted well with their corresponding counterparts in the rabbit microsomal preparations (results not shown). These observations confirmed that the human and rabbit microsomal preparations used were composed of a mixture of internal (T-tubules and junctional sarcoplasmic reticulum) and external (sarcolemmal components of the dystrophin system) muscle membranes [14]. Overall, the results presented here indicated that the GENX-3414 protein products were preferentially found in the internal membrane system of the skeletal muscle.

To investigate whether the novel proteins were integral membrane proteins or peripheral proteins anchored to some component of the muscle membrane, KCl-washed light microsomes were subjected to washing with 0.8 M KCl or alkaline extraction [15]. The corresponding proteins were not detected in the soluble supernatant after the vesicles were washed with 0.8 M KCl (Figure 6A) or extracted in buffer at pH 11.5 (Figure 6B): they remained in the insoluble membrane fraction. The proteins were fully dissociated from the membranes by 0.5% (v/v) Triton X-100 treatment of the membrane extracts (Figure 6C), as were other integral membrane components of the muscle, such as the calcium pump [16] (results not shown).

Chromosomal location and genomic mapping

In initial sequence and clustering studies, the gene corresponding to GENX-3414 was assigned to chromosome 4 using monochromosomal somatic cell hybrids [1]. Three RH markers (RH9827, Bda66a06, Genbank F00445; RH20291, NIB1224, Genbank T16380; and RH22238, WI-17734, Genbank H26168) were located using the Genebridge 4 RH panel [19] and one on the G3 panel (STSG-63501, Genbank Z24879) [20]. Markers Bda66a06 and WI-17734 correspond to the same Unigene cluster Hs.109590. Both human-origin STSs (referred to as WI-17734 and STSG-63501 respectively, which presented partial nucleotide sequences identical with that of the GENX-3414 marker) are noted in Figure 2 (see single-underlined and dotted-underlined sequences respectively). The Bda66a06 marker was located between D4S392 and D4S395 in a 13 centimorgan (cM) interval, whereas the NIB1224 marker was located between D4S392 and D4S2958 in a 4 cM interval and the WI-17734 marker was located between D4S2958 and D4S400 in a 6 cM interval. Marker

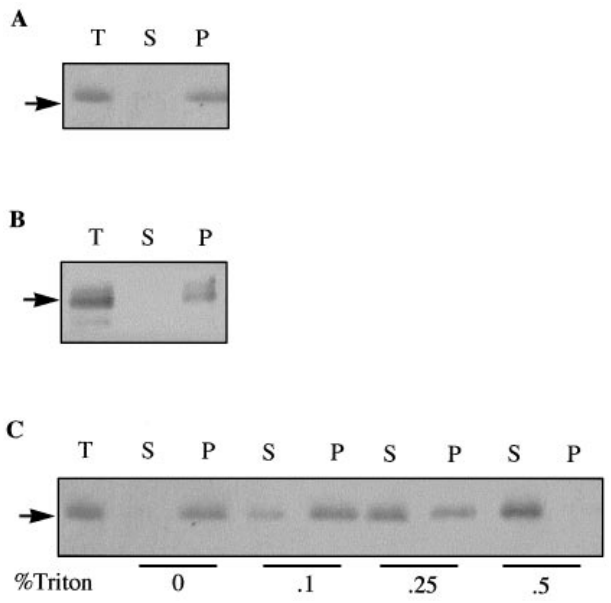

Figure 6 Membrane-binding studies on GENX-3414 protein products

KCl washing (**A**), alkaline treatment (**B**) and Triton X-100 extraction (**C**) were performed using human KCl-washed light microsomes [15,16]. The contents of the GENX-3414 protein products in the supernatants (S) and the pellets (P), produced from diverse treatments of the initial microsomal preparations (T), were immunodetected with the specific mAb 4D11 raised against a recombinant fragment encoded by a partial GENX-3414 cDNA. After KCl washing and alkaline treatment, the GENX-3414 protein remained bound to the membranous pellets. With increasing Triton X-100 concentrations (up to 0.5%), progressive solubilization of the protein (see supernatants) was observed. The arrows indicate the GENX-3414 protein products.

STSG-63501 was linked 0.1299 Gy 100 (the '100' refers to intensity of irradiation in Gy) from marker SHGS4-1008 with a lod score of 8.6. WI-17734 was mapped 4.96 Gy from the top of the chromosome linkage group. It corresponded to an extended 4q24–q25 chromosomal position, compatible with the previously determined chromosomal location of the GENX-3414 gene. Finally, analysis of the On-line Mendelian Inheritance in Man (OMIM) map, which presents the cytogenetic map locations of disease genes, revealed that a gene candidate for the long QT syndrome with ventrical tachyarrythmia type 4 presented significant positive linkages with respect to markers on 4q25–q27 [27].

DISCUSSION

This was a follow-up to a large-scale differential expression study that identified 41 novel gene transcripts with a specific and/or highly intense expression profile in human skeletal muscle [3]. Our aim was to characterize one member of this cDNA collection, i.e. the GENX-3414 gene transcript. Transcript analyses in different human tissues from different subjects indicated that the GENX-3414 gene gave rise to one specific 2.4-kb-long transcript highly expressed in skeletal and cardiac muscles with a unique 1074 bp ORF. Co-purification of the two 41–43 kDa immunodetected GENX-3414 protein products, together with skeletal muscle membrane fragments containing markers of the sarcoplasmic reticulum, T-tubules and sarcolemmal membrane, experimentally demonstrated their direct or indirect attachment to some muscle membranes. Immunostaining confined solely within the cytoplasm showed that GENX-3414 protein products were co-localized with both sarcoplasmic reticulum and T-tubules, i.e. the main constituents of the internal membrane system of skeletal muscle cells. Their precise local distribution as a subtle intracytoplasmic polygonal cluster was identical with that observed with different transverse tubule markers, such as the TS28 protein [28] or the DHPR [14,29], and for the 95 kDa triadin molecule, another major muscle membrane protein [29–32]. The latter protein is localized specifically in the junctional sarcoplasmic reticulum of skeletal muscle, where it binds both the ryanodine receptor (RyR) of the sarcoplasmic reticulum and the α1 subunit of the DHPR in the T-tubule membrane [33,34]. Preferential expression of the novel protein in muscle fast fibres was similar to that documented previously [13,28] for other components of the internal muscle membrane, such as DHPR and triadin molecules. The increased abundance of T-tubules and junctional sarcoplasmic reticulum was responsible for the observed overexpression. The results of the present analysis, using classical optical and confocal microscopy, clearly indicated that the GENX-3414 protein product is distributed throughout the tubular transverse membranes and at the junctional complexes between the sarcoplasmic reticulum and T-tubules.

The capacity of both related GENX-3414 protein products to bind to the internal membrane system of skeletal muscle cells could be due to the hydrophobic 24-amino-acid-long segment present at their N-terminal ends. Detailed examination of the membrane-binding domain, however, revealed that it was not a conventional transmembrane domain, because it was bordered by a long C-terminal hydrophilic cytoplasmic segment without any luminal N-terminal domain. The persistent fixation of GENX-3414 protein products to the membrane vesicles after high-salt-concentration washing or alkaline extraction [15], along with their extraction by 0.5% (v/v) Triton X-100 alone [16], provided conclusive evidence that this protein is an integral membrane protein.

Most muscle components of the sarcomeric apparatus, such as myosin subunits [10] or titin [35], or of the membrane system, such as the calcium pump or triadin [36], exist in different isoforms, and many muscle membrane proteins are also often either glycosylated or phosphorylated. Regarding the novel GENX-3414 protein products, it is currently not known whether the two quite similar proteins are two different isoforms, or whether they co-exist as a native form and a post-translationally modified form. Northern blot analysis is currently too insensitive and, from the number of sequencing studies performed, it seems likely that the differences between the two GENX-3414 protein products are due to post-translational modifications rather than an alternative splicing mechanism. No glycoconjugate in the GENX-3414 protein products was detected using a digoxigenin glycan detection kit (Boehringer). Further investigation will tell whether either of the potential post-translational sites, scanned in the GENX-3414 primary protein structure, could be involved in the development of the two protein products detected from one transcript.

Many different protein components are visible on polyacrylamide gels analysing the content of the many muscle microsomal preparations reported over the last three decades (reviewed in [26]). Until the present study, the GENX-3414 protein product was one of many other components of muscle microsomes that had neither been identified nor determined either at the genetic level or in terms of their amino acid sequences and functions. In fact, only a few of the membrane components, localized in the sarcolemmal membrane, T-tubules, sarcoplasmic reticulum or connecting zones such as triads, have been well characterized. Studies on their individual roles as functional receptors (calcium pump, DHPR, RyR) or constitutive structural proteins (triadin, TS28) have provided definitive evidence that this complex membrane system has an essential role in excita-

tion–contraction coupling and in calcium exchanges in skeletal muscle cells. The characteristics of the GENX-3414 gene revealed initially and the subcellular localization of its protein product(s) have not yet revealed its exact function in this complex environment of internal skeletal muscle membranes. The leucine-zipper motif identified in the GENX-3414 sequence could permit involvement of novel protein(s) in protein–protein interactions with either known or unknown partners, also localized along the same internal muscle membranes. Other identified phosphorylation and glycosylation sites could also contribute towards the functional activity of the novel protein.

Among the few disease genes mapped in the 4q24–q25 chromosomal region where the GENX-3414 gene was located, only the long QT syndrome candidate gene [27] could be connected with previously identified muscular diseases. The membrane localization and high cardiac expression of the GENX-3414 protein products are clearly compatible with its putative functional involvement in a specific cardiac channel. Its intricate structure with a unique N-terminal membrane-binding domain does not, however, correspond to that known for other classical potassium channels, which are usually composed of six transmembrane domains. The role of GENX-3414 protein products could be either structural or regulatory, interacting with, or in the vicinity of, an as yet unknown channel. Future genetic work using the now-isolated full-length GENX-3414 gene should determine whether GENX-3414 is associated with cardiac disease. More generally, future work on expression of the novel human protein in heart and other non-muscle tissues, e.g. placenta or liver, should help towards highlighting whether it is also bound to any membrane system analogous with the internal membrane system of skeletal muscle cells and, in particular, if it plays a role in voltage-gated potassium-channel-controlled cardiac repolarization.

We thank G. Hugon, G. Le Cam, N. Lautrédou, N. Fayein and C. Matingou for excellent technical assistance, C. Jacquet for attentive animal care, and F. Pons and A. Tkatchenko for helpful discussions. We are very grateful to the orthopaedics division of Clinique Clémentville for the skeletal muscle biopsies of patients and to H. Le Marec (Centre Hospitalier Universitaire, Nantes, France) for the heart biopsy. This work was supported by the Institut National de la Santé et de la Recherche Médicale, and the Centre National à la Recherche Scientifique and l'Association Française contre les Myopathies.

REFERENCES

1 Houlgatte, R., Mariage-Samson, R., Duprat, S., Tessier, A., Bentolila, S., Lamy, B. and Auffray, C. (1995) Genome Res. **5**, 272–304
2 Nguyen, C., Rocha, D., Granjean, S., Baldit, M., Bernard, K., Naquet, P. and Jordan, B. (1995) Genomics **29**, 207–216
3 Piétu, G., Alibert, V., Guichard, B., Lamy, F., Bois, E., Leroy, R., Mariage-Samson, R., Houlgatte, P., Soularue, P. and Auffray, C. (1996) Genome Res. **6**, 492–503
4 Auffray, C., Behar, G., Bois, F., Boucher, C., Da Silva, C. D., Devignes, M., Duprat, S., Houlgatte, R., Jumeau, M., Lamy, B. et al. (1995) C. R. Acad. Sci. Ser. III **318**, 263–272
5 Lennon, G. G., Auffray, C., Polymeropoulos, M. and Soares, B. (1996) Genomics **33**, 151–152
6 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
7 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Nucleic Acids Res. **25**, 3389–3402
8 Bairoch, A., Bucher, P. and Hofman, K. (1996) Nucleic Acids Res. **24**, 189–196
9 Köhler, G. and Milstein, C. (1975) Nature (London) **256**, 495–497
10 Dechesne, C. A., Léger, J., Bouvagnet, P., Mairhofer, H. and Léger, J. J. (1985) Circ. Res. **57**, 767–775
11 Lee, C. C., Pons, F., Jones, P. G., Bies, R. D., Schlang, A. M., Leger, J. J. and Caskey, C. T. (1993) Hum. Gene Ther. **4**, 273–281
12 Leung, A. T., Imagawa, T., Block, B., Franzini-Armstrong, C. and Campbell, K. P. (1988) J. Biol. Chem. **263**, 994–1001
13 Knudson, C. M., Stang, K. K., Jorgensen, A. O. and Campbell, K. P. (1993) J. Biol. Chem. **268**, 12637–12645
14 Ohlendieck, K., Ervasti, J. M., Snook, J. B. and Campbell, K. P. (1991) J. Cell Biol. **112**, 135–148
15 Guo, W., Jorgensen, A. O. and Campbell, K. P. (1994) J. Biol. Chem. **269**, 28359–28365
16 Costello, B., Chadwick, C., Saito, A., Chu, A., Maurer, A. and Fleischer, S. (1986) J. Cell Biol. **103**, 741–753
17 Walter, M. A., Spillett, D. J., Thomas, P., Weissenbach, J. and Goodfellow, P. N. (1994) Nat. Genet. **7**, 22–28
18 Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillett, D., Muselet, D., Prud'homme, J. F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J. and Goodfellow, P. N. (1996) Hum. Mol. Genet. **5**, 339–346
19 Schuler, G. D., Boguski, M. S., Stewart, E. A., Stein, L. D., Gyapay, G., Rice, K., White, R. E., Rodriguez-Tomé, P. and Aggarwal, A. (1996) Science **274**, 540–546
20 Stewart, E. A., McKusick, K. B., Aggarwal, A., Bajorek, E., Brady, A., Chu, A., Fang, N., Hadley, D., Harris, M., Hussain, S. et al. (1997) Genome Res. **7**, 422–433
21 Kozak, M. (1991) J. Biol. Chem. **266**, 19867–19870
22 Kyte, J. and Doolittle, R. F. (1982) J. Mol. Biol. **157**, 105–132
23 Hidalgo, C., Gonzalez, M. E. and Lagos, R. (1983) J. Biol. Chem. **258**, 13937–13945
24 Saito, A., Seiler, S., Chu, A. and Fleischer, S. (1984) J. Cell Biol. **99**, 875–885
25 Caswell, A.H., Brandt, N.R., Brunschwig, J.-P. and Kawamoto, R. M. (1988) Methods Enzymol. **157**, 68–84
26 Ikemoto, N., Kim, D. H. and Antoniu, B. (1988) Methods Enzymol. **157**, 469–480
27 Schott, J. J., Charpentier, F., Peltier, S., Foley, P., Drouin, E., Bouhour, J.B, Donnelly, P., Vergnaud, G., Bachner, L., Moisan, J. P. et al. (1995) Am. J. Hum. Genet. **57**, 1114–1122
28 Jorgensen, A., Arnold, W., Shen, A., Yuan, S., Gaver, M. and Campbell, K. (1990) J. Cell Biol. **110**, 1173–1185
29 Leung, A. T., Imagawa, T. and Campbell, K. (1987) J. Biol. Chem. **262**, 7943–7946
30 Brandt, N. R., Caswell, A. H. and Brunschwig, J. P. (1980) J. Biol. Chem. **255**, 6290–6298
31 Caswell, A. H., Brandt, N. R., Brunschwig, J. P. and Purkerson, S. (1991) Biochemistry **30**, 7507–7513
32 Guo, W. and Campbell, K. P. (1995) J. Biol. Chem. **270**, 9027–9030
33 Flucher, B. E., Andrews, S. B., Fleischer, S., Marks, A. R., Caswell, A. and Powell, J. A. (1993) J. Cell Biol. **123**, 1161–1174
34 McPherson, P. S. and Campbell, K. P. (1993) J. Biol. Chem. **268**, 13765–13768
35 Kolmerer, B., Olivieri, N., Witt, C. C., Herrmann, B. G. and Labeit, S. (1996) J. Mol. Biol. **256**, 556–563
36 Guo, W., Jorgensen, A. O., Jones, L. R. and Campbell, K. P. (1996) J. Biol. Chem. **271**, 458–465

Received 18 May 1998/27 July 1998; accepted 21 August 1998