Nesprins: a novel family of spectrin-repeat-containing proteins that localize to the nuclear membrane in multiple tissues

Qiuping Zhang[1], Jeremy N. Skepper[2], Fangtang Yang[3], John D. Davies[1], Laszlo Hegyi[1,2], Roland G. Roberts[5], Peter L. Weissberg[1], Juliet A. Ellis[4] and Catherine M. Shanahan[1,*]

[1]Department of Medicine, Division of Cardiovascular Medicine, University of Cambridge, Box 110, Addenbrooke's Hospital, Hills Road, Cambridge, CB2 2QQ, UK
[2]Department of Anatomy, Multi-Imaging Centre, Tennis Court Rd, Cambridge, CB2 3DY, UK
[3]Centre for Veterinary Science, University of Cambridge, Madingley Road, Cambridge, CB3 0ES, UK
[4]Randall Centre for the Molecular Mechanism of Cell Function, Kings College, New Hunts House, Guy's Campus, London, SE1 1UL, UK
[5]Division of Medical & Molecular Genetics, GKT Medical School, 8th Floor, Guy's Tower, Guy's Hospital, London SE1 9RT, UK
*Author for correspondence (e-mail: cs131@mole.bio.cam.ac.uk)

Accepted 13 September 2001
Journal of Cell Science 114, 4485-4498 (2001) © The Company of Biologists Ltd

SUMMARY

In search of vascular smooth muscle cell differentiation markers, we identified two genes encoding members of a new family of type II integral membrane proteins. Both are ubiquitously expressed, and tissue-specific alternative mRNA initiation and splicing generate at least two major isoforms of each protein, with the smaller isoforms being truncated at the N-terminus. We have named these proteins nesprin-1 and -2 for nuclear envelope spectrin repeat, as they are characterized by the presence of multiple, clustered spectrin repeats, bipartite nuclear localization sequences and a conserved C-terminal, single transmembrane domain. Transient transfection of EGFP-fusion expression constructs demonstrated their localization to the nuclear membrane with a novel C-terminal, TM-domain-containing sequence essential for perinuclear localization. Using antibodies to nesprin-1, we documented its colocalization with LAP1, emerin and lamins at the nuclear envelope, and immunogold labeling confirmed its presence at the nuclear envelope and in the nucleus where it colocalized with heterochromatin. Nesprin-1 is developmentally regulated in both smooth and skeletal muscle and is re-localized from the nuclear envelope to the nucleus and cytoplasm during C2C12 myoblast differentiation. These data and structural analogies with other proteins suggest that nesprins may function as 'dystrophins of the nucleus' to maintain nuclear organization and structural integrity.

Key words: Nuclear envelope, Nucleus, Spectrin-repeats, Differentation, Muscle

INTRODUCTION

Vascular smooth muscle cells (VSMCs) are the major cellular component of the blood vessel wall where they normally exist in a contractile, differentiated state, and they function to regulate vascular tone. However, in association with a number of diseases. VSMCs dedifferentiate, lose many of their contractile properties and become functionally impaired. Studies focusing on the identification of markers of the contractile phenotype have identified a number of smooth-muscle-specific genes encoding proteins that regulate contractile function, such as α-SM actin and SM22α, whose expression is downregulated in dedifferentiated cells (Schwartz et al., 1995; Shanahan and Weissberg, 1998).

In a previous experiment, we used differential cDNA screening to isolate gene markers of VSMC differentiation, and we identified a partial rat cDNA clone, 1RA1, which on Northern blots hybridized to an mRNA transcript of approximately 10.0 kb (Shanahan et al., 1993). Isolation and full length sequencing of the human homologue of 1RA1 identified a gene that encodes the founding member of a novel gene family. We demonstrate that, unlike other VSMC differentiation markers, the protein encoded by this gene, renamed nesprin-1, does not form part of the contractile apparatus but localizes primarily to the nuclear membrane.

The nuclear envelope (NE) consists of two lipid bilayers, the inner and outer nuclear membranes, which are joined at nuclear pores (Gerace and Burke, 1988). The inner membrane is supported by the lamina, a network of intermediate filament proteins including lamins B and A/C, whereas the outer membrane is continuous with the endoplasmic reticulum (ER) (Bergmann and Singer, 1983; Stuurman et al., 1998). Only a few proteins are known to localize to the NE. Of particular interest are the lamin-associated proteins, which include emerin, lamina-associated polypeptides (LAP) 1 and 2 and the lamin B receptor (LBR). These proteins share little sequence homology but are structurally related, in that they all have a long N-terminal nucleoplasmic domain, a common targeting mechanism to the NE, variable numbers of transmembrane (TM) domains and the ability to bind lamins (Furukawa et al.,

1995; Lin et al., 2000; Rolls et al., 1999). The importance of NE proteins to cellular function has been highlighted by genetic studies that have revealed that mutations in emerin and lamins A/C cause Emery-Dreifuss muscular dystrophy (EDMD) (Bione et al., 1994; Bonne et al., 1999). Moreover, different mutations in lamins A/C are associated with dilated cardiomyopathy, limb girdle muscular dystrophy 1B and a phenotypically unrelated syndrome Dunnigan-type familial partial lipodystrophy (Cao and Hegele, 2000; Fatkin et al., 1999; Muchir et al., 2000). It is unclear why mutations in these ubiquitously expressed proteins should lead to tissue-specific pathologies, hence the importance of gaining further insight into the function and interactions of NE proteins (Wilson, 2000).

In this study, we have identified and characterized two members of a new family of nuclear-membrane-associated proteins, called nesprins, which contain variable numbers of spectrin-repeat domains with homology to similar domains in the cytoskeletal proteins α- and β-spectrin and dystrophin. In addition, bipartite nuclear localization signals (NLS) and a single TM domain near the C-terminus make them structurally similar to lamin-binding proteins of the NE. Nesprin-1 is on human chromosome 6q25 and nesprin-2 on chromosome 14q23. Nesprin mRNAs undergo extensive alternative initiation and splicing to produce isoforms markedly different in size, with the smaller α isoforms being truncated at the N_2-terminal and therefore lacking a number of spectrin repeats. Although nesprins are widely expressed, high expression of both nesprin-1 and -2 is observed in skeletal, cardiac and vascular smooth muscle. Our studies suggest that nesprins may have a specific function in muscle cell differentiation. However, high expression of nesprin-1 in peripheral blood leukocytes and the spleen, and of nesprin-2 in pancreas and testis, suggests a broader range of functions. We propose that nesprins may be membrane-anchored 'dystrophins' of the nucleus that may potentially have a role in dystrophic muscular syndromes in man.

MATERIALS AND METHODS

Library screening/rapid amplification of cDNA ends (RACE)

Full-length rat nesprin-1 was generated from rat aortic cDNA using a combination of RACE and library screening. 5×10^6 plaques of a rat aortic cDNA library were screened using a 2.1 kb rat 1RA1 cDNA clone as a probe (Shanahan et al., 1993). Positive phage were rescued into Bluescript SK$^-$ (Stratagene) and sequenced. 5′ RACE was performed using touchdown PCR and Marathon cDNA Amplification and Advantage-GC2 PCR kit (Clontech). PCR fragments were cloned into the pCRII vector (Invitrogen) for automated sequencing using BigDye™ sequencing kit. Human nesprin-1 and -2 were generated using Marathon-ready cDNAs from human spleen, heart, placenta, testis (Clontech) and skeletal muscle (Origene).

RT-PCR and northern analysis

Total cytoplasmic RNA was isolated from rat aortic tissue and (2 μg) reverse-transcribed as previously described (Shanahan et al., 1999). Control reactions without reverse transcriptase/RNA were performed. PCR amplification used standard reaction conditions and the cycling parameters: 94°C for 0.5 minutes, 55°C for 1 minute and extension at 72°C for 1.5 minutes for 30 cycles. Northern blots were performed as previously described (Shanahan et al., 1993). Multiple human tissue Northerns (Clontech) were hybridised according to the manufacturer's instructions using [^{32}P]-labeled human nesprin cDNA probes generated by PCR for nesprin-1 and from IMAGE clone DKFZp434G173Q1 (http://www.rzpd.de) for nesprin-2.

Database analysis

Sequence information was assembled and analysed using the Staden program. Database homology searches were performed using the NetBlast program, Wisconsin genetics Group (GCG) package, version 9.1. Protein homology domains were identified using SMART, Profilescan and PIX. The spectrin repeats were aligned using Clustal_X (Thompson et al., 1997). Intron/exon boundaries were mapped using the NCBI Human Genome BLAST server and confirmed using RT-PCR.

Plasmid constructs and site-directed mutagenesis

Human cDNAs for enhanced green fluorescent protein (EGFP) constructs were amplified using high fidelity GC-rich PCR (Roche) and inserted in frame into *Bgl*II/*BspE*I and *Sal*I sites of the pEGFP-C1 vector (Clontech). The TM deletion (LIGLAC) constructs were generated using QuikChange™ XL site-directed mutagenesis kit (Stratagene).

Cell culture and transfection analysis

Human VSMCs were cultured as described previously (Shanahan et al., 1999). COS-7 green monkey fibroblasts and mouse C2C12 myoblasts were cultured at 37°C/5% CO_2 in M-199/10% fetal calf serum (FCS) and Dulbecco's modified Eagle's medium (DMEM)/15% FCS (Sigma), respectively. To generate myotubes, confluent C2C12 myoblasts were changed to 2% horse serum and 10% CO_2 for three to five days. For transient transfection, cells were plated onto chamber slides at 1×10^5 cells/ml and transfected using Superfect™ (Qiagen). EGFP was visualised using an Olympus IX70 inverted fluorescence microscope with a Chroma GFP-optimised (495 nm) filter, and live images were captured using an Hamamatsu CCD digital camera and Image proplus photographic system (Media Cybernetics).

Antibody production and western blotting

Rabbit antibodies to human nesprin-1α and -1β were generated against five synthetic polypeptides: N1: KLAFLLKDWEKC (residues 2191-2202); N2: KGIADSLEKLRTFKC (residues 2204-2217); N3: SKASEIEYKLGKVNDRC (residues 2374-2390); C1: GRSTPNRQ-KTPRGKC (residues 3212-3226); C2: CALSNNFARSFHPMLR (residues 3297-3313) (Immune Systems Ltd, Paignton, UK). Each polypeptide was conjugated to keyhole limpet hemocyanin, and the conjugates were injected into rabbits to produce polyclonal antibodies, which were subsequently affinity purified and ELISA tested. Western blots were performed according to standard procedures. Nesprin-1 was detected using N3, C1 and C2 antibodies (diluted 1:800-1:2000), followed by incubation with a horseradish-peroxidase-conjugated anti-rabbit IgG secondary antibody diluted 1:2000 (NA934, Amersham Pharmacia Biotech). Antibody-specificity controls were performed by pre-incubation with peptides. The ECL+Plus chemiluminescent kit (Amersham Pharmacia Biotech) was used for signal detection.

In vitro transcription/translation and immunoprecipitation

Nesprin-1α cDNA was inserted into pcDNA 3.1 and transcribed from the T7 promoter and translated in a TNT^R T7-coupled reticulocyte lysate system (Promega, UK) according to the manufacturer's instructions. Incorporation into microsomal membranes was assessed by the inclusion of canine pancreatic microsomes (Promega), added either during translation for two hours (co-translational studies) or subsequent to translation for a further 30 minutes (post-translational studies). To study post-translational membrane insertion, the ribosomes were removed by sedimentation for 30 minutes at 4°C and

100,000 rpm (356,000 ***g***) (rotor TLA 100) in a Beckman table-top ultracentrifuge prior to the addition of the microsomes. To determine if resulting proteins were translocated into microsomes, 0.1 mg/ml Proteinase K and 1% Triton-X100 were used for proteolysis, and sedimentation was used to determine the presence of membrane-inserted proteins as described previously (Ellis et al., 1998).

C2C12 cells were cultured for 24 hours in methionine-free DMEM in the presence of 500 μCi/ml [^{35}S]-methionine, and in vitro translation of nesprin-1α was performed in the presence of [^{35}S]-methionine. Immunoprecipitation [IP] was performed by adding the labelled nesprin-1α translation product or C2C12 cells to lysis IP buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA-Na_2, 1% Triton X-100) and incubation with nesprin-1 (N3 or C1) antibodies at 4°C overnight. This was followed by the addition of Protein G-agarose beads (P4691, Sigma) and a further incubation for one hour at 4°C. After washing with IP buffer and sedimentation of the protein complexes, the pellets were subjected to SDS-PAGE electrophoresis and autoradiography.

Immunohistochemistry and confocal microscopy

Frozen sections of fresh human tissue biopsies were placed onto Superfrost plus microscope slides, fixed in acetone for 10 minutes at –20°C and permeablized with 0.5% NP-40/TBS (Tris buffered saline) for 3.5 minutes at room temperature. After blocking with 20% normal goat serum/TBS at RT for 30 minutes and incubation overnight with N3, C1 or C2 antibodies (diluted 1:100 in TBS), the sections were incubated in the dark for 30 minutes with a Alexa Fluor™ 488 conjugated goat anti-rabbit IgG secondary antibody diluted 1:200 (A-11008, Vector Laboratories, Inc). Rabbit IgG (I5006, Sigma) was used as a negative control. Cultured cells or myotubes on chamber slides were stained as described above. Other markers examined using monoclonal antibodies include: α-smooth muscle actin (A2547, Sigma), α-sarcomeric actin (A2172, Sigma), myosin (clone A4. 1025, Alexis Corporation), calnexin (C45520, Transduction Laboratories), lamin A/C (JOL2, abCAM, UK), emerin (1:100) and LAP1 (1:100) (gift of J. Ellis) were incubated with Alexa Fluor™ 568 conjugated goat anti-mouse IgG secondary antibody diluted 1:200 (A11004, Vector Laboratories, Inc). All samples were mounted in Vectashield™ mounting medium containing DAPI (Vector Laboratories, Inc) and images were captured using a Leica TCS-NT-UV laser scanning confocal system.

Immunogold labelling

Cells were fixed in 4% (v/v) formaldehyde in 0.1 mol/L PIPES buffer (Sigma), cryo-protected in 25% (v/v) polypropylene glycol and frozen in melting propane. They were freeze-substituted in dry methanol at –90°C containing 0.1% uranyl acetate and embedded in Lowicryl HM-20 (Taab) at –50°C. Thin sections (50 nm) were mounted on Formvar film grids and incubated in TBS at pH 7.4 containing 0.5% FCS and 10% normal goat serum for five minutes at room temperature to block non-specific binding. This was followed by incubation with primary nesprin-1 antibodies, diluted 1:100 in blocking buffer, for 16 hours at room temperature. After rinsing with TBS, they were incubated with secondary goat anti-rabbit immunoglobulins conjugated to 10 nm gold particles (British Biocell, Cardiff, UK) diluted 1:100 in TBS at pH 8.2 for 60 minutes at room temperature. After extensive rinsing in TBS, followed by dH_2O, they were stained with uranyl acetate and lead citrate and viewed in a CM-100 transmission EM (Philips, Eindhoven, Netherlands). Non-specific binding was tested by omission of the primary antibody.

Fluorescence in situ hybridisation

Human metaphase spreads were prepared from a phytohemagglutinin (PHA, Gibco BRL, UK)-stimulated blood cultures, and the human nesprin-1 and -2 cDNA clones were mapped to human chromosomes by fluorescence in situ hybridisation (FISH) as previously described (Yang et al., 2000).

RESULTS

Isolation of a novel cDNA encoding nesprin-1, a differentiation marker for VSMCs

Differential cDNA screening of a rat VSMC library identified a 2.1 kb cDNA clone, 1RA1, that hybridized to a VSMC mRNA of approximately 10.0 kb and that was more highly expressed in differentiated aortic VSMCs than in cultured, dedifferentiated VSMCs (Fig. 1A). RT-PCR demonstrated that rat 1RA1 was upregulated during vascular development with an expression profile typical of known contractile protein

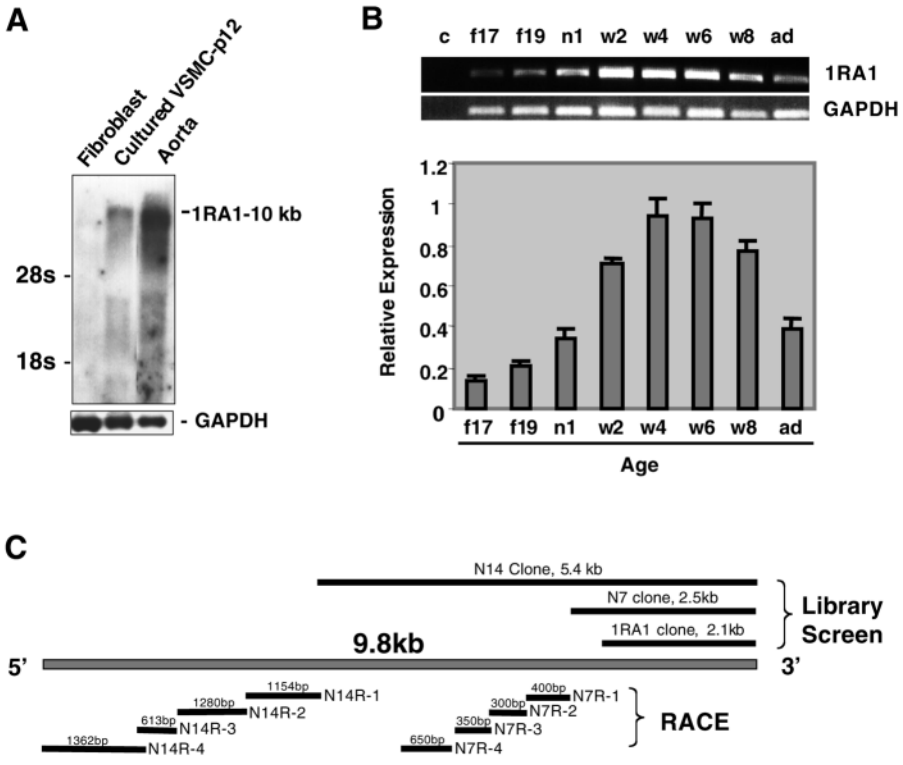


Fig. 1. Cloning and expression of rat nesprin-1. (A) Northern blot showing high expression of a 10 kb transcript of rat nesprin-1 (1RA1) in aortic tissue compared with cultured, passage 12, dedifferentiated VSMCs and fibroblasts. (B) Semi-quantitative RT-PCR analysis of rat nesprin-1 (1RA1), showing an increase in expression during developmental differentiation of the aorta. f17, foetal day 17; f19, foetal day 19; n1, neonatal day 1; w2-w8, weeks 2 to 8; ad, adult and c, negative control. The graph shows relative expression of 1RA1 at each stage normalized to a GAPDH control. (C) Schematic representation of isolation of rat nesprin-1 cDNA sequence.

differentiation markers (Shanahan and Weissberg, 1998) (Fig. 1B). Re-screening of the rat VSMC library and RACE identified a 9.8 kb cDNA sequence containing a single open reading frame predicting a 2,838 amino acid (aa) protein, which was named nesprin-1 (Fig. 1C).

Comparison of the full-length rat sequence with human databases identified two partial cDNA clones, lymphocyte membrane-associated protein and KIAA0796 from brain with >82% identity to rat nesprin-1, suggesting that they encoded the human orthologue (Nagase et al., 2000). Using these homologies to design PCR and RACE primers, the full-length human cDNA of 10,742 bases, containing a single open reading frame and with 83% identity to the rat sequence, was isolated from spleen.

Northern blots demonstrated that nesprin-1 was ubiquitously expressed in human tissues, but had the highest expression in spleen, peripheral blood leukocytes and heart. However, two major transcripts at 10.7 kb and 3.8 kb were identified that were differentially expressed between tissues (Fig. 2A). Using RACE and heart cDNA, the smaller transcript was identified as a 5′-truncation of the longer transcript. The two transcripts both contained conserved Kozak consensus sequences and single open reading frames, predicting proteins of 982 and 3,321 residues, named nesprin-1α and nesprin-1β respectively (Fig. 2B,C).

Human nesprin-1 was mapped to a contig on chromosome 6q24-25 identified in the Human Genome BLAST server database. This information, combined with PCR analysis, was used to map the 64 exons of nesprin-1 distributed across approximately 200 kb of DNA (Fig. 2B). The nesprin-1α transcript is initiated in an alternate exon (45) containing only 5′ untranslated (UTR) sequence and had an additional exon 55. A rare variant nesprin-1α_2, identified in heart, skeletal muscle and spleen, is initiated in the alternatively spliced exon 47, which encodes an additional, N-terminal 31 residues.

Isolation of nesprin-2

Further searching of public databases revealed two partial cDNA clones (IMAGE:DKFZp434G173.1 and KIAA1011) isolated from testis and brain that shared approximately 64% homology with nesprin-1. Northern blot analysis of multiple human tissue mRNAs using the 4.3 kb IMAGE clone as a probe revealed high expression of a 2.7 kb transcript in pancreas, heart and skeletal muscle and lower expression of numerous other transcripts, ranging in size from 3.1-11.0 kb, in other tissues including testis, pancreas and kidney (Fig. 3A). RACE analysis revealed that the largest 11.1 kb transcript, which was isolated from the testis, encoded nesprin-2γ, a protein of 3,270 residues. The major transcript at 2.7 kb, nesprin-2α, and another at 3.4 kb, nesprin-2β, were truncated versions of the longer transcript. These sequences contained single open reading frames that predicted proteins of 542 and 763 residues respectively (Fig. 3).

The nesprin-2 sequence was homologous to a contig on human chromosome 14, and intron/exon boundary mapping identified 65 exons spanning 150 kb of genomic DNA (Fig. 3C). Nesprin-2 shows a number of different transcripts owing to alternate initiation occurring in exons 51, 55 and 57 that generate transcripts encoding nesprin-2β_2 -2α, and -2α_2. Nesprin-2β is generated by alternate initiation in exon 49, which contains only 5′-UTR sequence. Competing 5′-splice sites (cryptic splice site) in exon 61 result in an insertion of 14 residues in nesprin-2α_2. Additionally, exon 56, which is only utilized in nesprin-2α, encodes 23 residues with 65% identity to those encoded by exon 55 in nesprin-1, which is also alternatively spliced.

FISH analysis

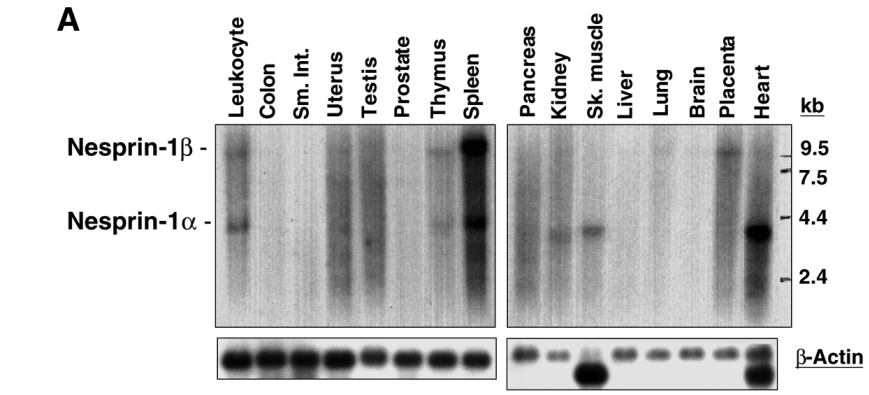
FISH identified regions of hybridization on human chromosome 6q25 for nesprin-1 and 14q23 for nesprin-2 (not shown). These localizations correspond to the stated origin of the genomic DNA contigs used to determine gene structure.

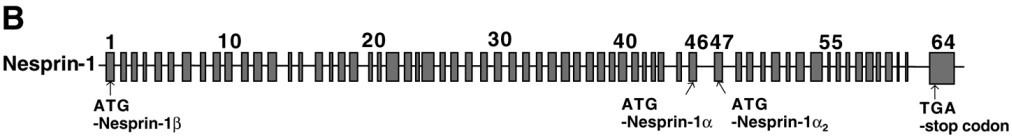
Predicted protein structure of nesprins

Database searches identified a partial cDNA, syne-1B, encoding part of the mouse orthologue of nesprin-1 and sharing >85% sequence homology to human and rat, and a shorter mouse cDNA, syne-1A, that corresponded to the rare nesprin-1α_2 splice variant in man (Apel et al., 2000). Nesprin-1 also shares weak homology (21-28%) with human α- and β-spectrin and dystrophin, all large cytoskeletal proteins, and to Kakapo/shot, a *Drosophila* protein that localizes to muscle-tendon junctions (Gregory and Brown, 1998; Grum et al., 1999; Strumpf and Volk, 1998; Winder et al., 1995). However, these homologies were restricted to multiple spectrin-repeat domains present in all these sequences.

Profilescan and SMART programmes predicted that both nesprin-1 and -2 contain a large number of conserved spectrin repeats (Fig. 4A). These are arranged with a group of two repeats located close to the C-terminal end, separated from a variable number of clustered repeats located centrally. Using Clustal_X to align the 21 repeats in nesprin-1β, we saw that the three helix domains in the repeats are more degenerate than that observed in dystrophin (Fig. 4B). In members of the dystrophin family, adherence to the repeat consensus is generally more degenerate than in the spectrins themselves, which may be related to constraints associated with spectrin's need to dimerise (Winder et al., 1995). In invertebrate dystrophin rod domains, many of the repeats are too divergent to identify by conventional pattern recognition, yet they show consistent off-diagonal signals in dotplots (both against themselves and against human dystrophin) characteristic of an internally repeated structure (Greener and Roberts, 2000). Similarly, when the rod domains of nesprin-1 and nesprin-2 were plotted against themselves or against human dystrophin an distinct 'box' of off-diagonal signals was obtained. In addition the PHD_sec program yields strong predictions for a series of long alpha-helices interrupted by short turns throughout the nesprin rod domains. Also, like dystrophin, COILS shows that nesprin rod domains have a periodic propensity for the formation of coiled structures, leading us to suggest that between the canonical 'spectin repeats' the general structure of the nesprins is maintained, leading to a fairly uniform rod-like entity.

TMpred analysis predicted a single TM domain located 30 residues from the C-terminus of each protein (Fig. 4C). This 21-residue, leucine-rich TM sequence is included within 60 residues of the C-terminus that shares 58% similarity with the C-terminal amino acids of the *Drosophila* Klarsicht protein and a predicted *C. elegans* protein with homology to basic helix-loop-helix transcription factors (Jackle and Jahn, 1998; Mosley-Bishop et al., 1999; Welte et al., 1998) (Fig. 4D). This





C

```
                                                                                SR1→
MELDAAVQKF LEQNGQLGKP LAKKIGKLTE LHQQTIRQAE NRLSKLNQAT SHLEEYNEML ELILKWIEKA KVLAHGTIAW    80
                                                                                      SR2→
NSASQLRKQY ILHQTLLEES KEIDSELEAM TEKLQYLTSV YCTEKMSQQV AELGRETEEL RQMIKIRLQN LQDAAKDMKK   160
FEAELKKLQA ALEQAQATLT SPEVGRLSLK EQLSHRQHLL SEMESLKPKV QAVQLCQSAL RIPEDVVASL PLCHAALRLQ   240
                                  SR3→
EEASRLQHTA IQQCNIMQEA VVQYEQYEQE MKHLQQLIEG AHREIEDKPV ATSNIQELQA QISRHEELAQ KIKGYQEQIA   320
SLNSKCKMLT MKAKHATMLL TVTEVEGLAE GTEDLDGELL PTPSAHPSVV MMTAGRCHTL LSPVTEESGE EGTNSEISSP   400
PACRSPSPVA NTDASVNQDI AYYQALSAER LQTDAAKIHP STSASQEFYE PGLEPSATAK LGDLQRSWET LKNVISEKQR   480
SR4→
TLYEALERQQ KYQDSLQSIS TKMEAIELKL SESPEPGRSP ESQMAEHQAL MDEILMLQDE INELQSSLAE ELVSESCEAD   560
                                                                  SR5→
PAEQLALQST LTVLAERMST IRMKASGKRQ LLEEKLNDQL EEQRQEQALQ RYRCEADELD SWLLSTKATL DTALSPPKEP   640
MDMEAQLMDC QNMLVEIEQK VVALSELSVH NENLLLEGKA HTKDEAEQLA GKLRRLKGSL LELQRALHDK QLNMQGTAQE   720
KEESDVDLTA TQSPGVQEWL AQARTTWTQQ RQSSLQQQKE LEQELAEQKS LLRSVASRGE EILIQHSAAE TSGDAGEKPD   800
VLSQELGMEG EKSSAEDQMR MKWESLHQEF STKQKLLQNV LEQEQEQVLY SRPNRLLSGV PLYKGDVPTQ DKSAVTSLLD   880
GLNQAFEEVS SQSGGAKRQS IHLEQKLYDG VSATSTWLDD VEERLFVATA LLPEETETCL FNQEILAKDI KEMSEEMDKN   960
KNLFSQAFPE NGDNRDVIED TLGCLLGRLS LLDSVVNQRC HQMKERLQQI LNFQNDLKVL FTSLADNKYI ILQKLANVFE  1040
QPVAEQIEAI QQAEDGLKEF DAGIIELKRR GDELQVEQPS MQELSKLQDM YDELMMIIGS RRSGLNQNLT LKSQYERALQ  1120
DLADLLETGQ EKMAGDQKII VSSKEEIQQP LDKHKEYFQG LESHMILTVT LFRKIISFAV QKETQFHTEL MAQASAVLKR  1200
                  SR6→
AHKRGVELEY ILETWSHLDE DQQELSRQLE VVESSIPSVG LVEENEDRLI DRITLYQHLK SSLNEYQPKL YQVLDDGKRL  1280
                                                  SR7→
LISISCSDLE SQLNQLGECW LSNTNKMSKE LHRLETILKH WTRYQSESAD LIHWLQSAKD RLEFWTQQSV TVPQELEMVR  1360
                                                                        SR8→
DHLNAFLEFS KEVDAQSSLK SSVLSTGNQL LRLKKVDTAT LRSELSRIDS QWTDLLTNIP AVQEKLHQLQ MDKLPSRHAI  1440
SEVMSWTSLM ENAIQKDEDN IKNSIGYKAI HEYLQKYKGF KIDINCKQLT VDFVNQSVLQ ISSQDVESKR SDKTDFAEQL  1520
                  SR9→                                                  SR10→
GAMNKSWQIL QGLVTEKIQL LEGLLESWSE YENNVQCLKT WFETQEKRLK QQHRIGDQAS VQNALKDCQD LEDLIKAKDK  1600
EVEKIEQNGL ALIQTKKEDV SSIVMSTLRE LGQTWANLDH MVGQLKILLK SVLDQWSSHK VAFDKINSYL MEARYSLSRF  1680
                                                                  SR11→
RLLTGSLEAV QVQVDNLQNL QDDLEKQERS LQKFGSITNQ LLKECHPPVT ETLTNTLKEV NMRWNNLLEE IAEQLQSSKA  1760
LLQLWQRYKD YSKQCASTVQ QQEDRTNELL KAATNKDIAD DEVATWIQDC NDLLKGLGTV KDSLFVLHEL GEQLKQQVDA  1840
                                                  SR12→
SAASAIQSDQ LSLSQHLCAL EQALCKQQTS LQAGVLDYET FAKSLEALEA WIVEAEEILQ GQDPSHSSDL STIQERMEEL  1920
                                                                        SR13→
KGQMLKFSSM APDLDRLNEL GYRLPLNDKE IKRMQNLNRH WSLISSQTTE RFSKLQSFLL QHQTFLEKCE TWMEFLVQTE  2000
QKLAVEISGN YQHLLEQQRA HELFQAEMFS RQQILHSIII DGQRLLEQGQ VDDRDEFNLK LTLLSNQWQG VIRRAQQRRG  2080
SR14→
IIDSQIRQWQ RYREMAEKLR KWLVEVSYLP MSGLGSVPIP LQQARTLFDE VQFKEKVFLR QQGSYILTVE AGKQLLLSAD  2160
                                                  SR15→
SGAEAALQAE LAEIQEKWKS ASMRLEEQKK KLAFLLKDWE KCEKGIADSL EKLRTFKKKL SQSLPDHHEE LHAEQMRCKE  2240
                                                                        SR16→
LENAVGSWTD DLTQLSLLKD TLSAYISADD ISILNERVEL LQRQWEELCH QLSLRRQQIG ERLNEWAVFS EKNKELCEWL  2320
TQMESKVSQN GDILIEEMIE KLKKDYQEEI AIAQENKIQL QQMGERLAKA SHESKASEIE YKLGKVNDRW QHLLDLIAAR  2400  (38)
                  SR17→                                                                 ^
VKKLKETLVA VQQLDKNMSS LRTWLAHIES ELAKPIVYDS CNSEEIQRKL NEQQELQRDI EKHSTGVASV LNLCEVLLHD  2480  (118)
                                                  SR18→
CDACATDAEC DSIQQATRNL DRRWRNICAM SMERRLKIEE TWRLWQKFLD DYSRFEDWLK SSERTAAFPS SSGVIYTVAK  2560  (198)
                                                                        SR19→
EELKKFEAFQ RQVHECLTQL ELINKQYRRL ARENRTDSAC SLKQMVHEGN QRWDNLQKRV TSILRRLKHF IGQREEFETA  2640  (278)
RDSILVWLTE MDLQLTNIEH FSECDVQAKI KQLKAFQQEI SLNHNKIEQI IAQGEQLIEK SEPLDAAIIE EELDELRKYC  2720  (358)
QEVFGRVERY HKKLIRLPLP DDEHDLSDRE LELEDSAALS DLHWHDRSAD SLLSPQPSSN LSLSLAQPLR SERSGRDTPA  2800  (438)
SVDSIPLEWD HDYDLSRDLE SAMSRALPSE DEEGQDDKDF YLRGAVALS* ********** ********** **GDHSALES  2857
---------- ---------- ---------- ---------- ---------D VMIPESPEAY VKLTENAIKN TS--------        (518)
QIRQLGKALD DSRFQIQQTE NIIRSKTPTG PELDTSYKGY MKLLGECSSS IDSVKRLEHK LKEEEESLPG FVNLHSTETQ  2937  (598)
                                  SR20→
TAGVIDRWEL LQAQALSKEL RMKQNLQKWQ QFNSDLNSIW AWLGDTEEEL EQLQRLELST DIQTIELQIK KLKELQKAVD  3017  (678)
                                                                  SR21→
HRKAIILSIN LCSPEFTQAD SKESRDLQDR LSQMNGRWDR VCSLLEEWRG LLQDALMQCQ GFHEMSHGLL LMLENIDRRK  3097  (758)
NEIVPIDSNL DAEILQDHHK QLMQIKHELL ESQLRVASLQ DMSCQLLVNA EGTDCLEAKE KVHVIGNRLK LLLKEVSRHI  3177  (838)
KELEKLLDVS SSQQDLSSWS SADELDTSGS VSPTSGRSTP NRQKTPRGKC SLSQPGPSVS SPHSRSTKGG SDSSLSEPGP  3257  (918)
GRSGRGFLFR VLRALPLQL LLLLLIGLAC LVPMSEEDYS CALSNNFARS FHPMLRYTNG PPPL                    3321  (982)
```

Fig. 2. Human nesprin-1 expression and gene analysis. (A) Northern blot of nesprin-1 expression in human tissues showing transcripts at 10.7 kb and 3.8 kb. The β-actin loading control is shown. (B) Gene structure of human nesprin-1. The numbered boxes indicate exons, the horizontal lines indicate introns. Nesprin-1α comprises 19 exons, starting at exon 45 and excluding exon 47; nesprin-1β comprises 61 exons (excluding exons 45, 47 and 55); nesprin-1α_2 comprises18 exons and is initiated in exon 47, which contains an additional 31 N-terminal residues. (C) The deduced amino acid sequence of human nesprin-1α and -1β. The first methionine of nesprin-1α and -1β are shown in bold and underlined, the spectrin repeats (SR) are shown by numbered grey boxes. The open box indicates the TM region and the dashed open boxes indicate NLS. The amino acids encoded by exon 55 are shown in bold, whereas the open triangle indicates the insertion site of the 31 residues encoded by exon 47 (not shown). GenBank accession numbers for human nesprin-1α and 1β are AY061756 and AY061755, respectively.

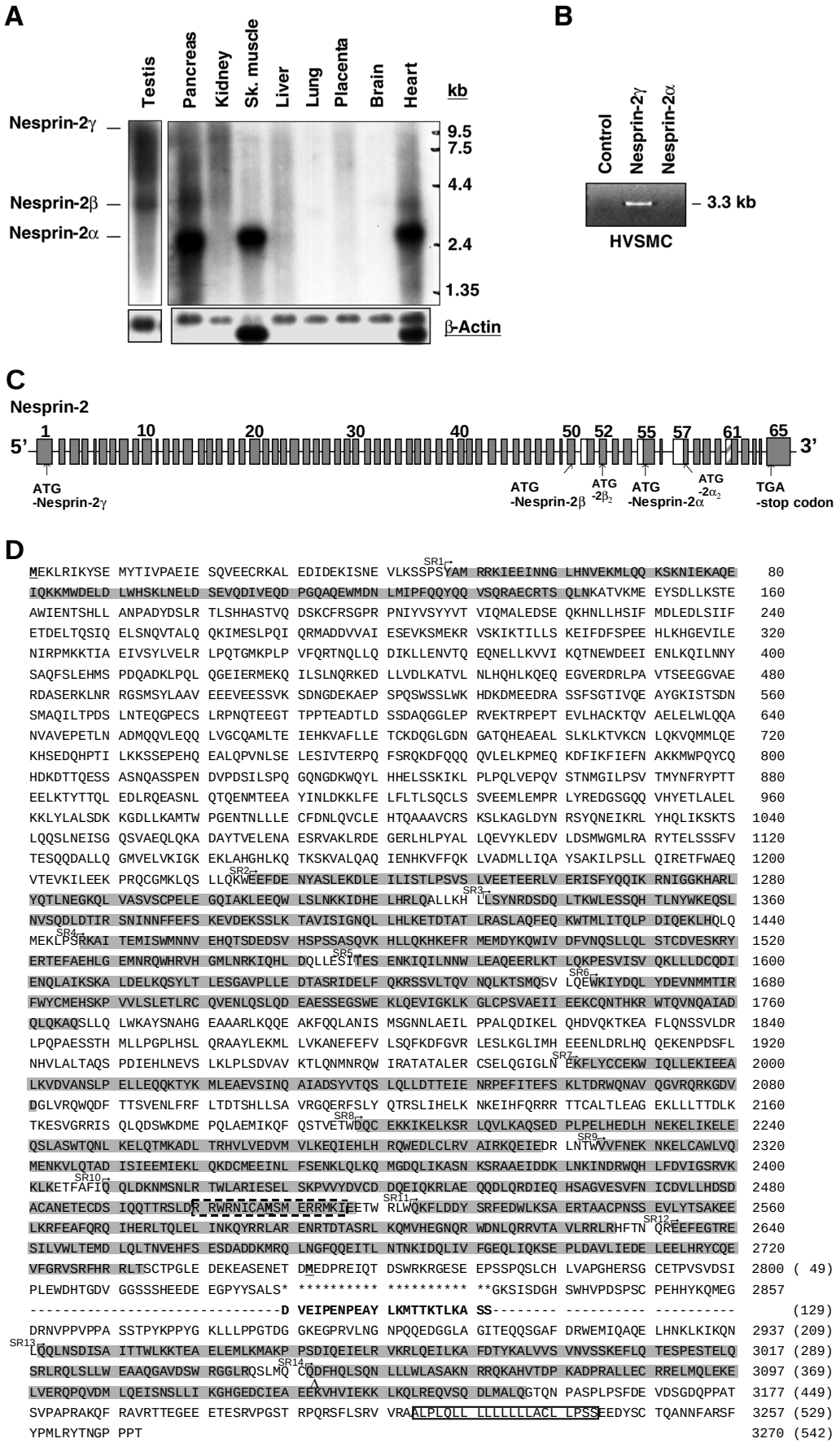
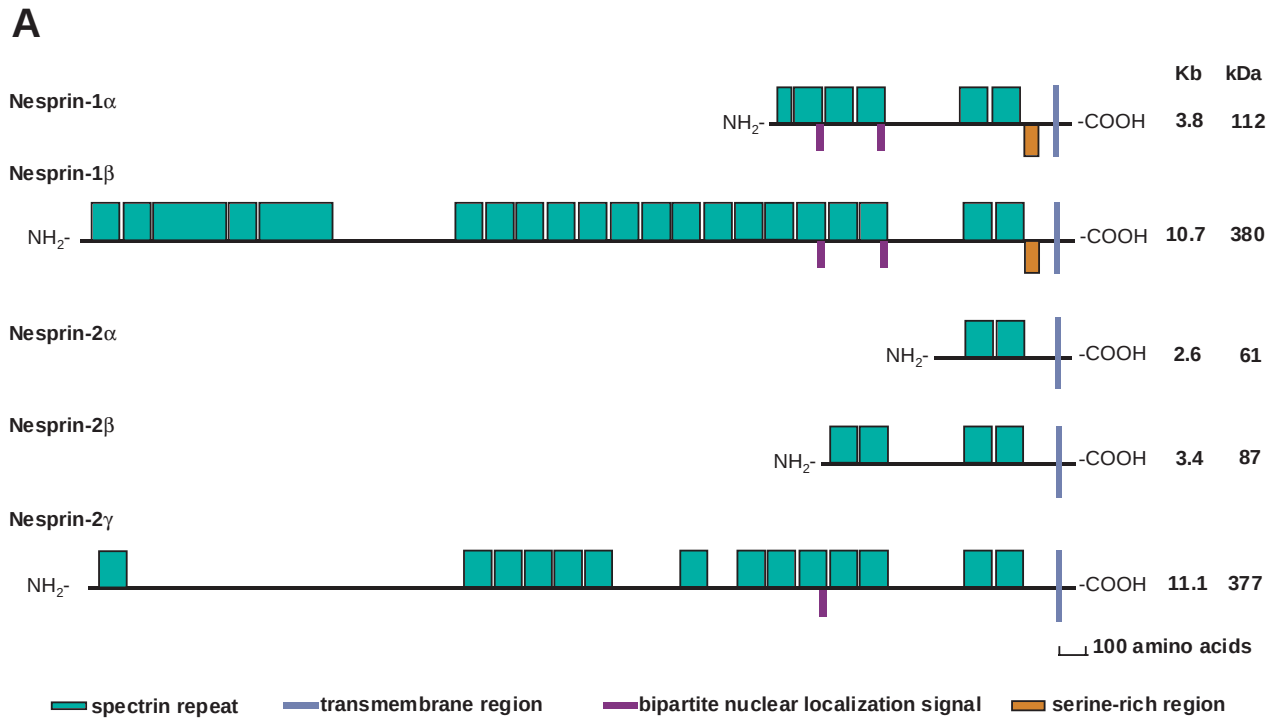
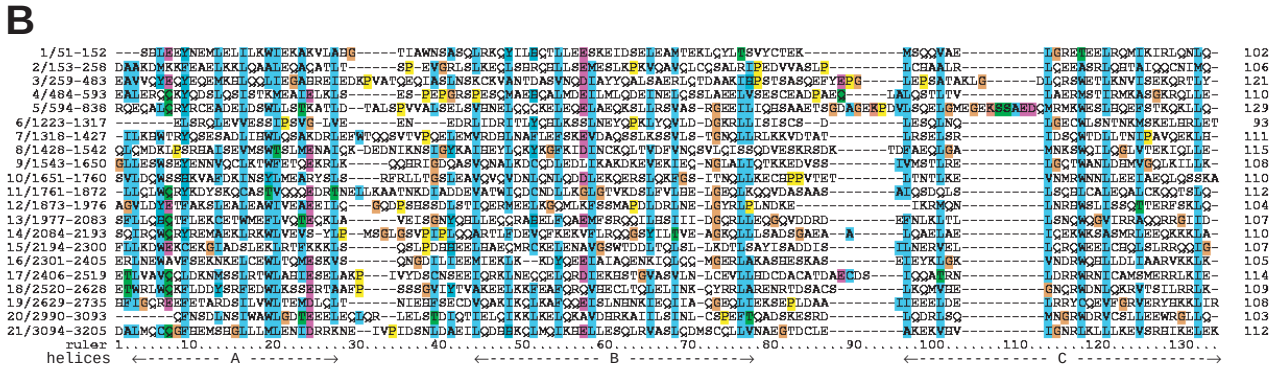
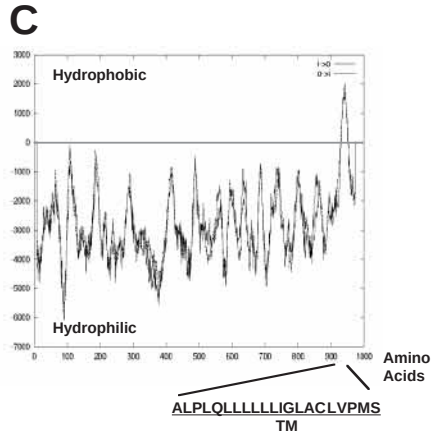


Fig. 3. Characterization and expression analysis of human nesprin-2. (A) Northern blot analysis of nesprin-2 expression in human tissues. Major transcripts at 11.1 kb, 3.4 kb and 2.7 kb are indicated. The β-actin loading control is shown. (B) RT-PCR of human *nesprin-2* gene expression in aorta indicated that the larger 2γ transcript was present in these cells. (C) Gene structure of human nesprin-2. Nesprin 2α comprises 11 exons (55-65), nesprin-2β is initiated in exon 49, which contains only 5′ UTR and comprises 16 exons (excluding exon 56). Nesprin-2γ comprises 63 exons (excluding exons 49 and 56). Alternatively spliced exons using cryptic splice sites are shown as open boxes. Alternate splicing at these sites produces transcripts $2\alpha_2$ and $2\beta_2$. (D) The deduced amino acid sequence of human nesprin-2α, -2β and -2γ. The first methionines of nesprin-2α, -2β and -2γ are shown in bold and underlined. The 23 residues encoded by exon 56 and specific to nesprin-2α are shown in bold. The open triangle shows the insertion site for 14 residues exclusive to $2\alpha_2$ and encoded by exon 61 spliced at a cryptic splice site (hatched box in 3C). The spectrin repeats (SR) are shown by numbered grey boxes, the open box indicates the TM region and the dashed open box indicates NLS. GenBank accession numbers for human nesprin-2α, -2β and -2γ are AY061758, AY061757 and AY061759, respectively.

```
MEKLRIKYSE MYTIVPAEIE SQVEECRKAL EDIDEKISNE VLKSSPSYAM RRKIEEINNG LHNVEKMLQQ KSKNIEKAQE    80
IQKKMWDELD LWHSKLNELD SEVQDIVEQD PGQAQEWMDN LMIPFQQYQQ VSQRAECRTS QLNKATVKME EYSDLLKSTE   160
AWIENTSHLL ANPADYDSLR TLSHHASTVQ DSKCFRSGPR PNIYVSYYVT VIQMALEDSE QKHNLLHSIF MDLEDLSIIF   240
ETDELTQSIQ ELSNQVTALQ QKIMESLPQI QRMADDVVAI ESEVKSMEKR VSKIKTILLS KEIFDFSPEE HLKHGEVILE   320
NIRPMKKTIA EIVSYLVELR LPQTGMKPLP VFQRTNQLLQ DIKLLENVTQ EQNELLKVVI KQTNEWDEEI ENLKQILNNY   400
SAQFSLEHMS PDQADKLPQL QGEIERMEKQ ILSLNQRKED LLVDLKATVL NLHQHLKQEQ EGVERDRLPA VTSEEGGVAE   480
RDASERKLNR RGSMSYLAAV EEEVEESSVK SDNGDEKAEP SPQSWSSLWK HDKDMEEDRA SSFSGTIVQE AYGKISTSDN   560
SMAQILTPDS LNTEQGPECS LRPNQTEEGT TPPTEADTLD SSDAQGGLEP RVEKTRPEPT EVLHACKTQV AELELWLQQA   640
NVAVEPETLN ADMQQVLEQQ LVGCQAMLTE IEHKVAFLLE TCKDQGLGDN GATQHEAEAL SLKLKTVKCN LQKVQMMLQE   720
KHSEDQHPTI LKKSSEPEHQ EALQPVNLSE LESIVTERPQ FSRQKDFQQQ QVLELKPMEQ KDFIKFIEFN AKKMWPQYCQ   800
HDKDTTQESS ASNQASSPEN DVPDSILSPQ GQNGDKWQYL HHELSSKIKL PLPQLVEPQV STNMGILPSV TMYNFRYPTT   880
EELKTYTTQL EDLRQEASNL QTQENMTEEA YINLDKKLFE LFLTLSQCLS SVEEMLEMPR LYREDGSGQQ VHYETLALEL   960
KKLYLALSDK KGDLLKAMTW PGENTNLLLE CFDNLQVCLE HTQAAAVCRS KSLKAGLDYN RSYQNEIKRL YHQLIKSKTS  1040
LQQSLNEISG QSVAEQLQKA DAYTVELENA ESRVAKLRDE GERLHLPYAL LQEVYKLEDV LDSMWGMLRA RYTELSSSFV  1120
TESQQDALLQ GMVELVKIGK EKLAHGHLKQ TKSKVALQAQ IENHKVFFQK LVADMLLIQA YSAKILPSLL QIRETFWAEQ  1200
VTEVKILEEK PRQCGMKLQS LLQKWEEFDE NYASLEKDLE ILISTLPSVS LVEETEERLV ERISFYQQIK RNIGGKHARL  1280
YQTLNEGKQL VASVSCPELE GQIAKLEEQW LSLNKKIDHE LHRLQALLKH LLSYNRDSDQ LTKWLESSQH TLNYWKEQSL  1360
NVSQDLDTIR SNINNFFEFS KEVDEKSSLK TAVISIGNQL LHLKETDTAT LRASLAQFEQ KWTMLITQLP DIQEKLHQLQ  1440
MEKLPSRKAI TEMISWMNNV EHQTSDEDSV HSPSSASQVK HLLQKHKEFR MEMDYKQWIV DFVNQSLLQL STCDVESKRY  1520
ERTEFAEHLG EMNRQWHRVH GMLNRKIQHL DQLLESITES ENKIQILNNW LEAQEERLKT LQKPESVISV QKLLLDCQDI  1600
ENQLAIKSKA LDELKQSYLT LESGAVPLLE DTASRIDELF QKRSSVLTQV NQLKTSMQSV LQEWKIYDQL YDEVNMMTIR  1680
FWYCMEHSKP VVLSLETLRC QVENLQSLQD EAESSEGSWE KLQEVIGKLK GLCPSVAEII EEKCQNTHKR WTQVNQAIAD  1760
QLQKAQSLLQ LWKAYSNAHG EAAARLKQQE AKFQQLANIS MSGNNLAEIL PPALQDIKEL QHDVQKTKEA FLQNSSVLDR  1840
LPQPAESSTH MLLPGPLHSL QRAAYLEKML LVKANEFEFV LSQFKDFGVR LESLKGLIMH EEENLDRLHQ QEKENPDSFL  1920
NHVLALTAQS PDIEHLNEVS LKLPLSDVAV KTLQNMNRQW IRATATALER CSELQGIGLN EKFLYCCEKW IQLLEKIEEA  2000
LKVDVANSLP ELLEQQKTYK MLEAEVSINQ AIADSYVTQS LQLLDTTEIE NRPEFITEFS KLTDRWQNAV QGVRQRKGDV  2080
DGLVRQWQDF TTSVENLFRF LTDTSHLLSA VRGQERFSLY QTRSLIHELK NKEIHFQRRR TTCALTLEAG EKLLLTTDLK  2160
TKESVGRRIS QLQDSWKDME PQLAEMIKQF QSTVETWDQC EKKIKELKSR LQVLKAQSED PLPELHEDLH NEKELIKELE  2240
QSLASWTQNL KELQTMKADL TRHVLVEDVM VLKEQIEHLH RQWEDLCLRV AIRKQEIEDR LNTWVVFNEK NKELCAWLVQ  2320
MENKVLQTAD ISIEEMIEKL QKDCMEEINL FSENKLQLKQ MGDQLIKASN KSRAAEIDDK LNKINDRWQH LFDVIGSRVK  2400
KLKETFAFIQ QLDKNMSNLR TWLARIESEL SKPVVYDVCD DQEIQKRLAE QQDLQRDIEQ HSAGVESVFN ICDVLLHDSD  2480
ACANETECDS IQQTTRSLDR RWRNICAMSM ERRMKIEETW RLWQKFLDDY SRFEDWLKSA ERTAACPNSS EVLYTSAKEE  2560
LKRFEAFQRQ IHERLTQLEL INKQYRRLAR ENRTDTASRL KQMVHEGNQR WDNLQRRVTA VLRRLRHFTN QREEFEGTRE  2640
SILVWLTEMD LQLTNVEHFS ESDADDKMRQ LNGFQQEITL NTNKIDQLIV FGEQLIQKSE PLDAVLIEDE LEELHRYCQE  2720
VFGRVSRFHR RLTSCTPGLE DEKEASENET DMEDPREIQT DSWRKRGESE EPSSPQSLCH LVAPGHERSG CETPVSVDSI  2800 ( 49)
PLEWDHTGDV GGSSSHEEDE EGPYYSALS* ********** ********** **GKSISDGH SWHVPDSPSC PEHHYKQMEG  2857
------------------------------D VEIPENPEAY LKMTTKTLKA SS-------- ---------- ----------       (129)
DRNVPPVPPA SSTPYKPPYG KLLLPPGTDG GKEGPRVLNG NPQQEDGGLA GITEQQSGAF DRWEMIQAQE LHNKLKIKQN  2937 (209)
LQQLNSDISA ITTWLKKTEA ELEMLKMAKP PSDIQEIELR VKRLQEILKA FDTYKALVVS VNVSSKEFLQ TESPESTELQ  3017 (289)
SRLRQLSLLW EAAQGAVDSW RGGLRQSLMQ CQDFHQLSQN LLLWLASAKN RRQKAHVTDP KADPRALLEC RRELMQLEKE  3097 (369)
LVERQPQVDM LQEISNSLLI KGHGEDCIEA EEKVHVIEKK LKQLREQVSQ DLMALQGTQN PASPLPSFDE VDSGDQPPAT  3177 (449)
SVPAPRAKQF RAVRTTEGEE ETESRVPGST RPQRSFLSRV VRALPLQLL LLLLLLLACL LPSSEEDYSC TQANNFARSF  3257 (529)
YPMLRYTNGP PPT                                                                           3270 (542)
```







D

```
klarsicht          2227 RGWAWRIARAAVPMQVALFTIFCAACLMQPN----CCDNLNNLSMSFTPQLRYIRGPPPI 2282
C.elegans           887 RSRWRRVLRTALPLQALLVLLMGAACLVPHCDDEYCCQLLNNFAKSFDPSLEFVNGPPPF  946
Human-Nesprin-1α    923 RGFLFRVLRAALPLQLLLLLLIGLACLVPMSEEDYSCALSNNFARSFHPMLRYTNGPPPL  982
Human-Nesprin-2α    483 RSFLSRVVRAALPLQLLLLLLLLLACLLPSSEEDYSCTQANNFARSFYPMLRYTNGPPPT  542
Rat-Nesprin-1      2779 RDFLFRVLRAALPIQLLLLLLIGLTCRVPISEEDYSCALSNNFARSFHPMLRYTNGPPPL 2838
Mouse-syn-1A        890 RAFLFRILRAALPFQLLLLLLIGLTCLVPMSEKDYSCALSNNFARSFHPMLRYTNGPPPL  949

Dro-Otefin          371 AFNRWLNSLEQKYHIKSKLFIVLVVLLLIGVYYIFY                          406
Human-Emerin        216 GLGQDRQVPLWGQLLLFLVFVIVLFFIYHFMQAEEGNPF                      254
Rat-emerin          218 ALGQDRQVPLWGQLLLFLAFATFLLFVYYSIQAQEGNPFWMDP                 260
Mouse-Emerin        217 ALGQDRQVPLWGQLLLFLVFAAFLLFVYYSIQAEEGNPFWMDP                 259
Rat-lap2            403 KTKKGRSVPMWIKMLLFALVAGFLFLVYQAMETNQGNPFTNFLQDTKISN         452
Xenops-lap2         465 KVKPGRSLPAWIKILLCIILVVFCFLVYQAMETNDGIAFSKLFLGITESNKTEN       518
```

Fig. 4. Predicted protein structure of nesprins-1 and -2. (A) Human nesprin protein domains generated by SMART and ProfileScan programmes. (B) Alignment of the 21 spectrin repeats of the nesprin-1 protein, showing homology domains within the three helices of each repeat. All G (orange) and P (yellow) residues are coloured. Other colouring is by conserved property in >55% of any column: uncoloured residues lack a sufficiently conserved property. Blue, hydrophobic residues; purple, negative residues, green, hydrophilic residues. (C) TMpred analysis showing the predicted hydrophobic TM domain in nesprin-1. (D) Alignment of the 60 C-terminal residues of nesprins showing homology with *Drosophila* Klarsicht and *C.elegans* (unknown). Note the lack of sequence homology with C-terminal TM domains (shown in bold) found in other NE proteins, including emerin and myoferlin.

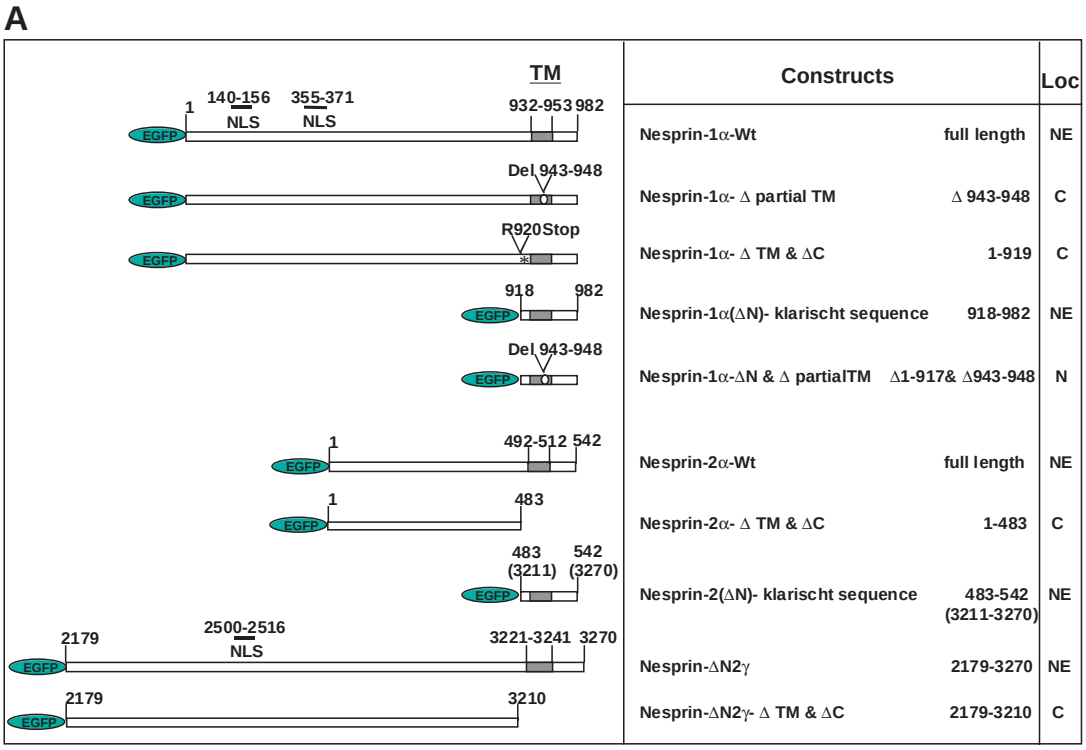
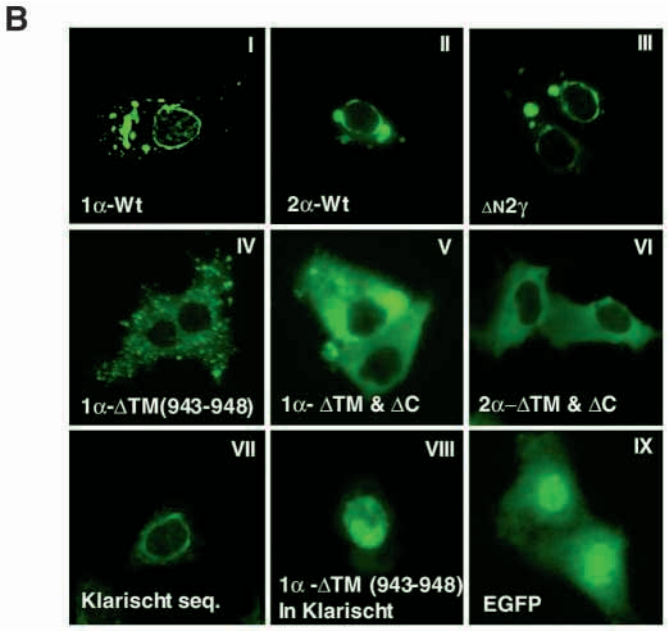


Fig. 5. (A) Nesprin-EGFP fusion constructs and (B) their subcellular localization in transfected COS-7 cells as indicated. (IX) EGFP vector alone.



structural arrangement, of a TM domain less than 60 residues from the C-terminus, is similar to that of lamin-binding integral membrane proteins of the NE. The NH-terminal regions of these proteins project into the nucleoplasm with the C-terminus anchored at the NE via variable numbers of membrane-spanning domains. However, this 60-residue nesprin sequence is not homologous to those in other NE proteins nor do nesprins contain a conserved LEM domain (Dechat et al., 2000) (Fig. 4D).

The N-terminal domains of nesprins both have conserved bipartite NLSs, although the truncated nesprin-2α and -2β lack these signals (Jans et al., 2000). Additionally, a putative protein-protein interaction motif (leucine zipper in nesprin-2γ) and a protein-DNA binding motif (helix-turn-helix motif in nesprin-1β) can be identified, both not present in the shorter α isoforms. Finally, a number of potential N-glycosylation sites and putative sites for phosphorylation were identified in each protein, suggesting they may be post-translationally modified. Close to the TM domain in nesprin-1 is a serine-rich motif, which represents a potential O-glycosylation site that is also found in emerin (Ellis et al., 1998).

Subcellular localization of nesprins

To determine the subcellular localization of the nesprins, nesprin-1 and -2 were fused in frame to the C-terminus of EGFP, and these expression constructs were transfected into COS-7 cells and C2C12 myoblasts (Fig. 5). Full-length nesprin-1α, nesprin-2α and N-terminally truncated 2γ-EGFP fusion constructs localized to the NE (Fig. 5B, I-III). In addition, there was nuclear staining and cytoplasmic inclusions in some cells, which may reflect either overexpression or true cytoplasmic localization.

To determine the signals required for NE targeting, deletion constructs were made (Fig. 5A). Deletion of the 60 C-terminal amino acids in nesprin-2α and -ΔN2γ, or mutation of the TM domain by the deletion of 6 residues (LIGLAC) in the context of nesprin-1α, resulted in only cytoplasmic localization of the EGFP fusion protein (Fig. 5B, IV-VI). Thus, constructs containing conserved bipartite NLSs but without the TM domain were not localized to the nucleus. This suggested that these consensus NLSs are inoperative, at least in COS-7 and C2C12 cells. Therefore, to determine if this 60-residue sequence was sufficient to direct NE localization, it was fused to EGFP, and in transfected cells it

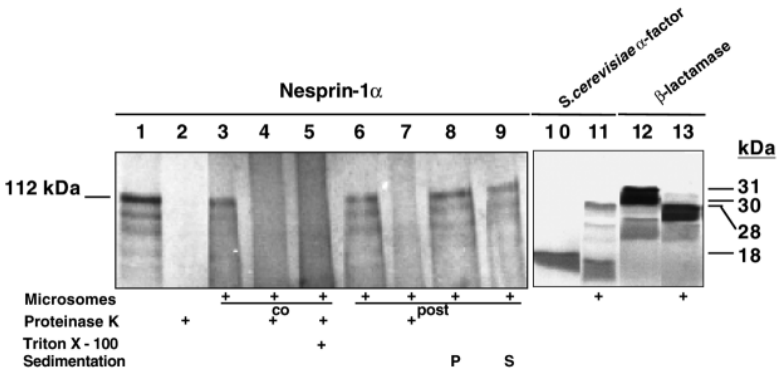


Fig. 6. In vitro transcription/translation of human nesprin-1α cDNA resulted in a 112 kDa product (lane 1). Translation reactions were as follows: nesprin-1α cDNA (lanes 1-9) and controls *S. cerevisiae* α-factor (10-11) and β-lactamase cDNA (12-13); either in the absence (1-2, 10,12) or in the presence of canine pancreatic microsome (3-9, 11,13); added either co-translationally (3-5, 11,13) or post-translationally (6-9). Reactions were digested with Proteinase-K alone (2, 4 and 7) or with 1% Triton X-100 (5). Translation products were subjected to sedimentation and separated into pellet (lane 8) and supernatant (lane 9) fractions. In these experiments, two controls were used to confirm microsome functionality. *S. cerevisiae* α-factor exhibits an increase in molecular weight due to being N-glycosylated, and pro-β-lactamase undergoes signal peptide cleavage to produce the lower molecular weight mature form.

was found to direct NE localization (Fig. 5B, VII). The introduction of the six-residue deletion into the TM domain of this 60-residue sequence resulted in only nuclear localization of the EGFP-fusion protein (Fig. 5B, VIII). Although this fusion protein is small enough to diffuse into the nucleus, its retention there suggests that this 60 residue sequence contains nuclear targeting signals, with the TM domain required for anchorage to the NE.

Nesprin-1 is a NE protein anchored by a C-terminal TM domain

Nesprin-1 shares features with a number of proteins of the inner NE that are trafficked to the nucleus via the ER, in particular the location of a single TM domain near the C-terminus (Kutay et al., 1993). To determine whether this was a true TM domain, we used in vitro transcription/translation in the presence of microsomes to determine if the protein was post-translationally modified and inserted into the microsomal membrane. In vitro translation of nesprin-1α in the presence of intact microsomes resulted in a protein of 112 kDa, suggesting that it was not glycosylated or proteolytically cleaved. Digestion of the nesprin-1α protein with proteinase-K in the presence of microsomes resulted in degradation of the N-terminal portion of the protein. This suggested that nesprin-1α inserts in a type II orientation with the C-terminus within the microsomal lumen. However, we were not able to confirm this, as the remaining amount of protein after digestion was too small to detect on a protein gel. Therefore, we separated the microsome and supernatant fractions using ultracentrifugation and found that approximately 60% of the nesprin-1 protein recovered was retained in the microsomal fraction. Although not definitive, this ratio is consistent with previous observations of other integral membrane proteins and is therefore consistent with nesprin-1α being attached to the membrane (Ellis et al., 1998) (Fig. 6).

Antibody characterization and subcellular localization of endogenous nesprin-1

Polyclonal antibodies were made against five peptides (N-terminus 1, 2 and 3 and C-terminus 1 and 2), covering both nesprin-1α and -1β amino acid sequences (Fig. 7A). These antibodies were affinity purified, and three (N3, C1 and C2) were shown to recognise nesprin-1α protein obtained from in vitro translation. This band migrated at 112 kDa on SDS-PAGE, which matched the size of the protein predicted by the open reading frame of the cDNA sequence (Fig. 7B). Moreover, these antibodies were able to immunoprecipitate a band of 112 kDa from an in vitro translation of nesprin-1α cDNA performed in the presence of [^{35}S]-methionine (Fig. 7C). On western blots of lysates from nesprin-1-EGFP fusion protein transfected COS-7 cells, these three antibodies recognised a specific protein at the predicted size of 139 kDa (112 kDa nesprin-1α plus 27 kDa GFP) (Fig. 7D).

Western blotting of tissue lysates from human VSMCs and peripheral blood leukocytes identified a number of proteins. Both N3 and C1 bound to protein doublets at approximately 95 kDa, 160 kDa and 360-380 kDa, the largest band corresponding to the predicted size of nesprin-1β. Additionally, N3 identified a band at 112 kDa in VSMCs, the predicted size of nesprin-1α (Fig. 7E). IP of [^{35}S]-labelled C2C12 myoblast cell lysates using antibodies N3 and C1 further confirmed that these antibodies recognise proteins at 95 kDa, 112 kDa and 160 kDa. Western blotting of C2C12 myoblasts also showed a similar pattern of multiple bands to that observed in VSMCs and leukocytes (Fig. 7F). Thus, these two antibodies weakly crossreact but they consistently identify, with different specificities, bands at 112 kDa and ~380 kDa corresponding to predicted protein products and additional bands at ~95 kDa and ~160 kDa. All of these bands were lost on pre-binding of the antibody with peptide (data not shown). These multiple bands may represent as yet unidentified splice variants, homologous protein domains or post-translationally modified products.

Immunolocalization studies, using antibodies N3, C1 and C2, performed on human VSMCs and C2C12 myoblasts confirmed the NE localization of nesprin-1. However, in some cells, particularly VSMCs, there was also significant nuclear localization of the protein and some staining consistent with ER localization (Fig. 7G). Therefore, we performed colocalization studies using antibodies to other NE proteins, the ER and the cytoskeleton. These analyses showed that nesprin-1 colocalised with nuclear lamins A/C, emerin and LAP1 at the NE. In most cells there was no consistent localization with either the ER, cytoplasm or cytoskeleton, although there was consistent nuclear staining of some cells (Fig. 8). Although we cannot rule out a role for nesprin-1 in these various cellular locations, these analyses led us to predict that nesprin-1 is predominantly a protein of the inner NE.

To confirm the subcellular localization, we performed

immunogold labeling analysis of the distribution of nesprin-1 protein in C2C12 myoblasts using antibodies N3 and C1. We found that nesprin-1 localized to the NE but was also present within the nucleus (Fig. 9). Here it co-localized with regions of heterochromatin and was absent in euchromatic regions, whereas in other cells prominent localization to the nucleolus was observed.

Tissue distribution and cellular localization of nesprin-1

Immunohistochemistry was used to determine the tissue distribution and localization of nesprin-1 in a variety of human tissues. In all tissues examined, perinuclear localization of nesprin-1 was observed as well as some nuclear staining, which was particularly evident in aortic smooth muscle (Fig. 10A). Only a subset of skeletal muscle cells was positive for nesprin-1, suggesting it may have a specific function in a subset of muscle cells.

Given that nesprin-1 was originally isolated from contractile VSMCs, an explanation for the different cellular localizations observed for nesprin-1 is that protein localization is determined by the differentiation state of the cell. Moreover, the heterogeneity of nesprin-1 distribution in skeletal muscle and its similarities with other NE proteins suggest that nesprin-1 may have a specific function in skeletal muscle. Therefore, we examined the localization of nesprin-1 during differentiation of C2C12 myoblasts into myotubes, a recognised model of skeletal muscle differentiation. This analysis showed that nesprin-1 was localized to the NE in myoblasts but redistributed, mainly to the nucleus and some to the cytoplasm, upon differentiation. The cytoplasmic localization did not appear to correlate with that of cytoskeletal components, instead nesprin-1 was diffusely distributed throughout the cell. This change in localization was observed with all three antibodies (N3, C1 and C2) generated against nesprin-1 and contrasts with that of emerin, which is present at the NE in both myoblasts and differentiated myotubes (Fig. 10B).

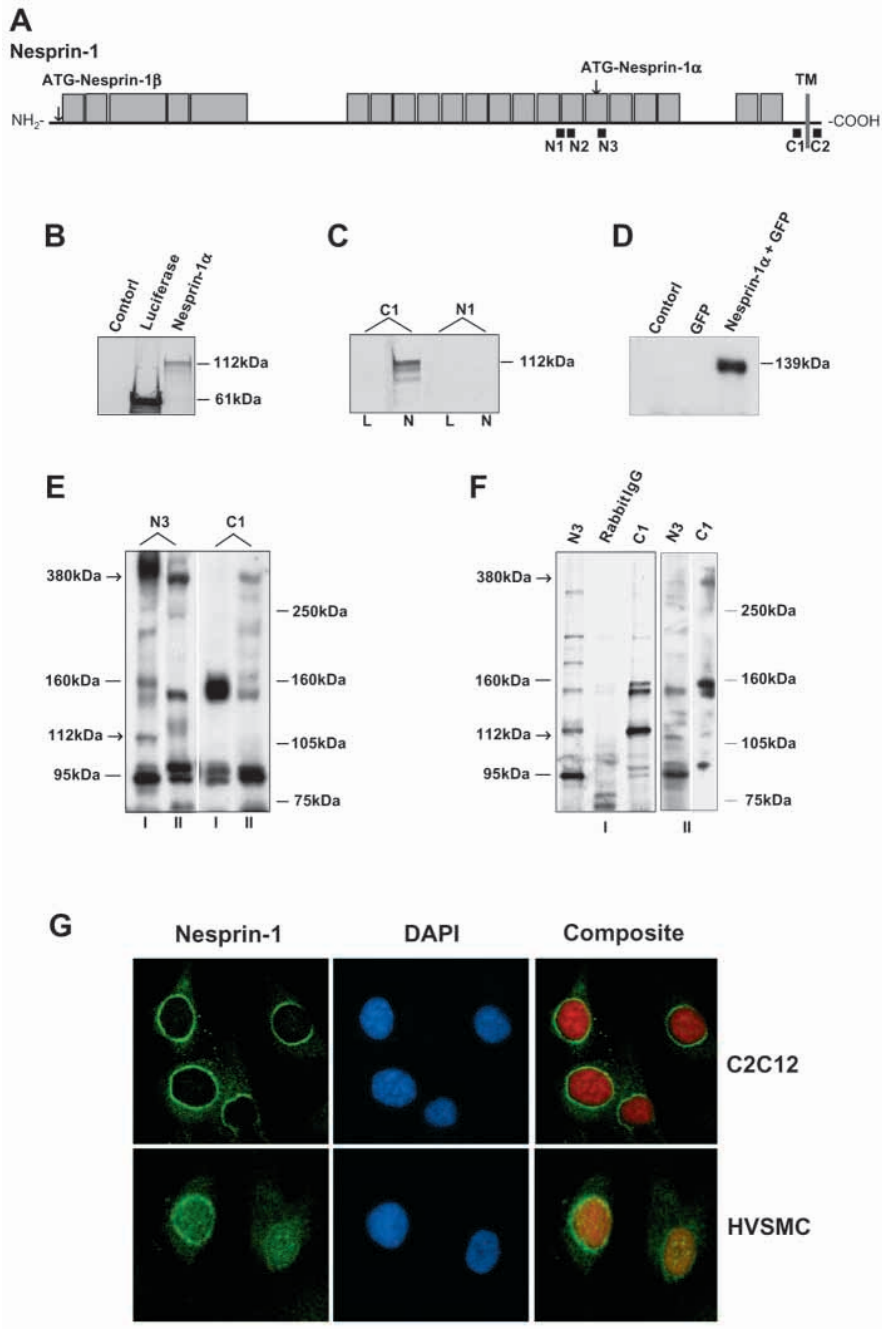


Fig. 7. (A) Schematic representation of the positions of the five peptides used for generating polyclonal antibodies against nesprin-1. The peptides N1 and N2 were specific to nesprin-1β but did not produce functional antibodies. The small black boxes indicate peptides. (B) In vitro transcription and translation of human nesprin-1α cDNA. The expression vector pcDNA3.1 containing the full-length cDNA of human nesprin-1α (112 kDa), the luciferase (61 kDa) T7 control cDNA as a positive control, the TNT lysate reaction without DNA as negative control. (C) Immunoprecipitation of human nesprin-1α generated from radioactive in vitro transcription/translation using antibodies C1 that recognised a 112kDa protein and N1, a non-specific negative control. (D) Western blot of human nesprin-1α-EGFP fusion protein in transfected COS-7 cells. Control; untransfected COS-7 cells, GFP: pEGFP-C1 vector alone and nesprin 1α-pEGFP-C1. (E) Western blot of human VSMC (I) and peripheral blood leukocyte (II) cell lysates using antibodies N3 and C1. The 112 kDa and 380 kDa bands are indicated by arrows. (F) Immunoprecipitation of C2C12 cells (I) using antibodies N3 and C1. Rabbit IgG as negative control. Western blot of C2C12 whole cell lysates (II) indicating that many of the IP products are also identified on western blots. (G) Immunofluorescent staining of nesprin-1 in mouse C2C12 myoblasts and human VSMCs. Endogenous nesprin-1 (green) was detected by nesprin-1 antibodies (N3, C1 and C2) and visualised by confocal microscope. DAPI is shown as false colour red in composite.

DISCUSSION

Nesprins; a new family of NE proteins

In this study, we have identified and characterized the protein structure and tissue distribution of nesprins, a novel family of ubiquitously expressed proteins localized primarily at the NE. Their general structure is that of type II integral membrane proteins, with a long N-terminal nucleoplasmic domain and a TM domain close to the C-terminus (Hartmann et al., 1989; Kutay et al., 1993). They are related to the lamin-binding NE proteins, but their major structural feature is the presence of multiple spectrin repeats, which are predicted to form a dystrophin-like rod domain. Indeed the arrangement of varying lengths of a spectrin-repeat rod domain followed by a means of membrane attachment is highly reminiscent of dystrophin, which, like nesprins, uses multiple promoters to generate at least five isoforms that differ in the length of their rod domain. Although the specific functions of dystrophin remain unclear, there is evidence for major mechanical and signalling roles (Roberts, 2001). Thus, the presence of a large dystrophin-like family of proteins at the NE points to the possibility that nesprins may regulate nuclear functions similar to those regulated by their cytoplasmic counterparts. These may include the maintenance of nuclear structural integrity, spatial orientation of nuclear contents and regulation of nuclear signalling.

Potential NE/nuclear binding partners for nesprins

The large rod proteins of the cytoskeleton, including spectrin-α and -β, dystrophin, utrophin and α-actinin, all form large complexes with multiple binding partners that function to maintain the structural and membrane stability of the cell. The formation of complexes enables dystrophin to link the cytoskeleton with the cell matrix, which is critical for cell signalling (Winder et al., 1995). Spectrin confers elasticity to the cell, and it has been shown to organise organelle structure and spatial orientation within the cell, thereby regulating intracellular transport (De Matteis and Morrow, 2000; Grum et al., 1999). The identification of protein-binding partners of nesprins, both in the nucleus and at the NE, may provide important clues as to their function. Cytoskeletal components known to bind spectrin repeats include actin, adducin and PDZ-containing proteins (Xia et al., 1997). Some of these interactions occur at 'specialized' regions within the rod domains of spectrin-repeat containing proteins. For

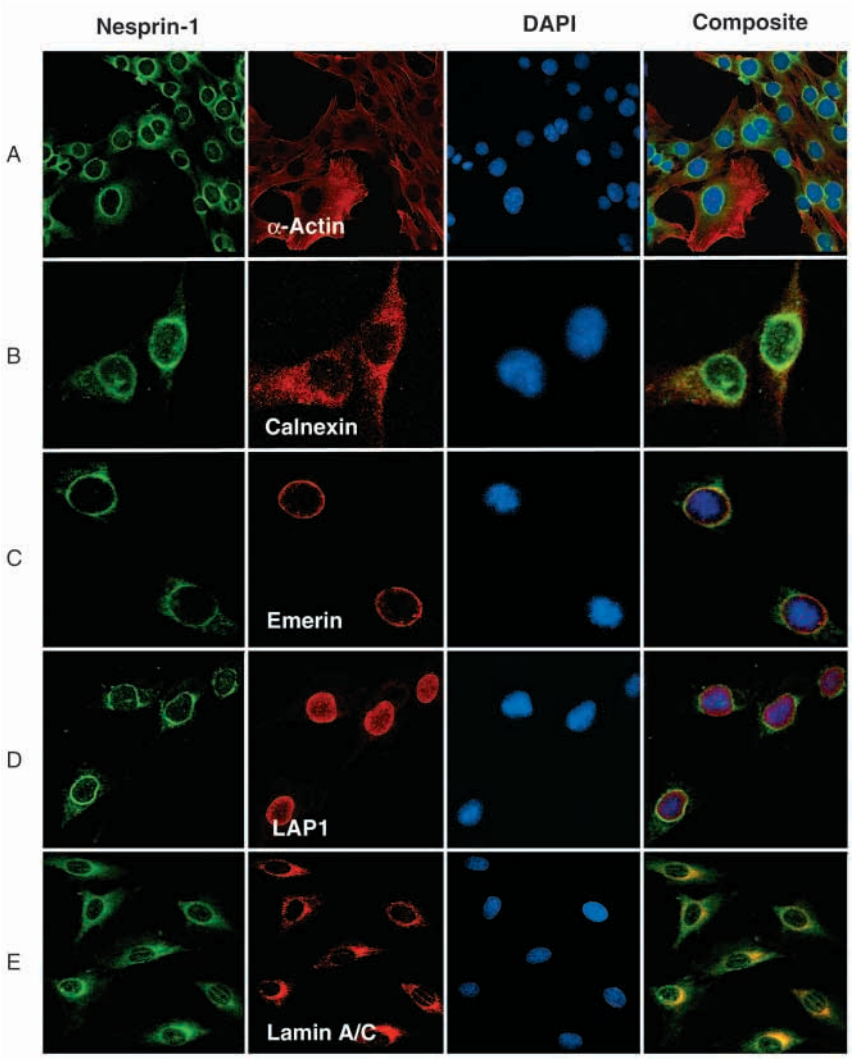


Fig. 8. Subcellular colocalization of endogenous nesprin-1 in C2C12 myoblasts (A,C) and human VMSCs (B,D,E). Cytoskeletal α-actin (A), ER protein calnexin (B), and NE proteins, emerin (C), LAP1 (D), lamin A/C (E). Nesprins are shown in green, other markers in red (as indicated) and the nucleus in blue. Yellow in merged images indicates regions of colocalization.

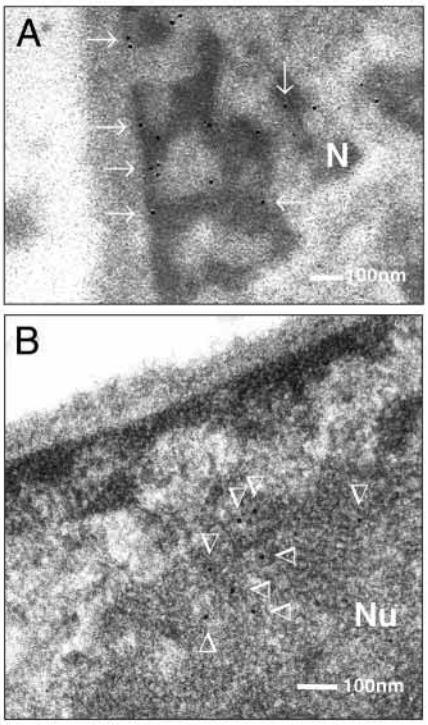


Fig. 9. Immunogold localization of nesprin-1 in C2C12 myoblasts. Gold particles (shown as black) localized along the NE are arrowed. Gold particles are also present (arrows) in heterochromatic regions of the nucleus (N) and are absent from euchromatic zones. In some cells, prominent localization to the nucleolus (Nu) was observed (B), as indicated by arrowheads.

example, a basic region within the rod domain of dystrophin has been shown to bind F-actin, a characteristic that is not conserved in the closely related protein utrophin (Amann et al., 1999). A search in nesprins for similar specialized repeats revealed only weak similarities to repeats in other proteins such as dystrophin and kakapo, making it difficult to make any specific functional extrapolations.

To date, no nuclear protein has yet been identified as a spectrin-repeat binding partner. Our studies showed that nesprin-1 colocalized at the NE with lamins, emerin and LAP1, suggesting that these may be potential binding partners. However, we found that nesprin-1 correctly localized to the NE in skeletal muscle cells from patients with EDMD who were null for emerin, suggesting that if an interaction between nesprin-1 and emerin occurs, it is not required to retain nesprin-1 at the NE (data not shown). However, a recent finding of potential importance was the identification of actin-binding and -modifying proteins in the nucleus, suggesting that actin may be part of the nuclear matrix and involved in processes such as chromatin remodelling and transcriptional control (Rando et al., 2000). Importantly, EM immunogold localization of nesprin-1 within the interphase nucleus suggested that nesprin-1 may be associated preferentially with heterochromatin and/or the nucleolus. Clearly, further in vitro studies are now required to determine whether nesprins can bind to actin, lamins and chromatin.

Identification of a new nuclear targeting signal

Proteins that reside in the NE must be specifically targeted to it and retained there. NE proteins exhibit a broad range of often multiple, hierarchical, targeting mechanisms (Jans et al., 2000; Soullam and Worman, 1995). For example, lamins are targeted to the nucleus via NLSs and enter via the nuclear pore complex (Stuurman et al., 1998). Targeting signals for the lamin-binding proteins LBR, LAP2 and emerin reside in their nucleoplasmic and TM domains and include conserved bipartite NLSs (Cartegni et al., 1997; Fairley et al., 1999). Additional signals for lamin and/or chromatin binding are also important for their correct localization and retention at the NE (Foisner and Gerace, 1993; Furukawa et al., 1998; Vlcek et al., 1999). Similarly, myoferlin and its related proteins target to the NE via a C-terminal TM domain-containing region (Ashery-Padan et al., 1997; Davis et al., 2000). In contrast, nurim, a novel protein of unknown function, has nuclear-targeting signals that reside in multiple membrane-spanning domains, whereas mAKAP, an A-kinase anchoring protein present in differentiated myocytes, has a central protein domain composed of three spectrin repeats in which its nuclear targeting signals reside (Kapiloff et al., 1999; Rolls et al., 1999).

We have shown that the NE-targeting signals for nesprins reside in their conserved hydrophobic C-terminal region that is devoid of spectrin-repeat domains. This 60-residue region shares no homology with other known mammalian NE proteins but is homologous to a predicted *C. elegans* helix-loop-helix-

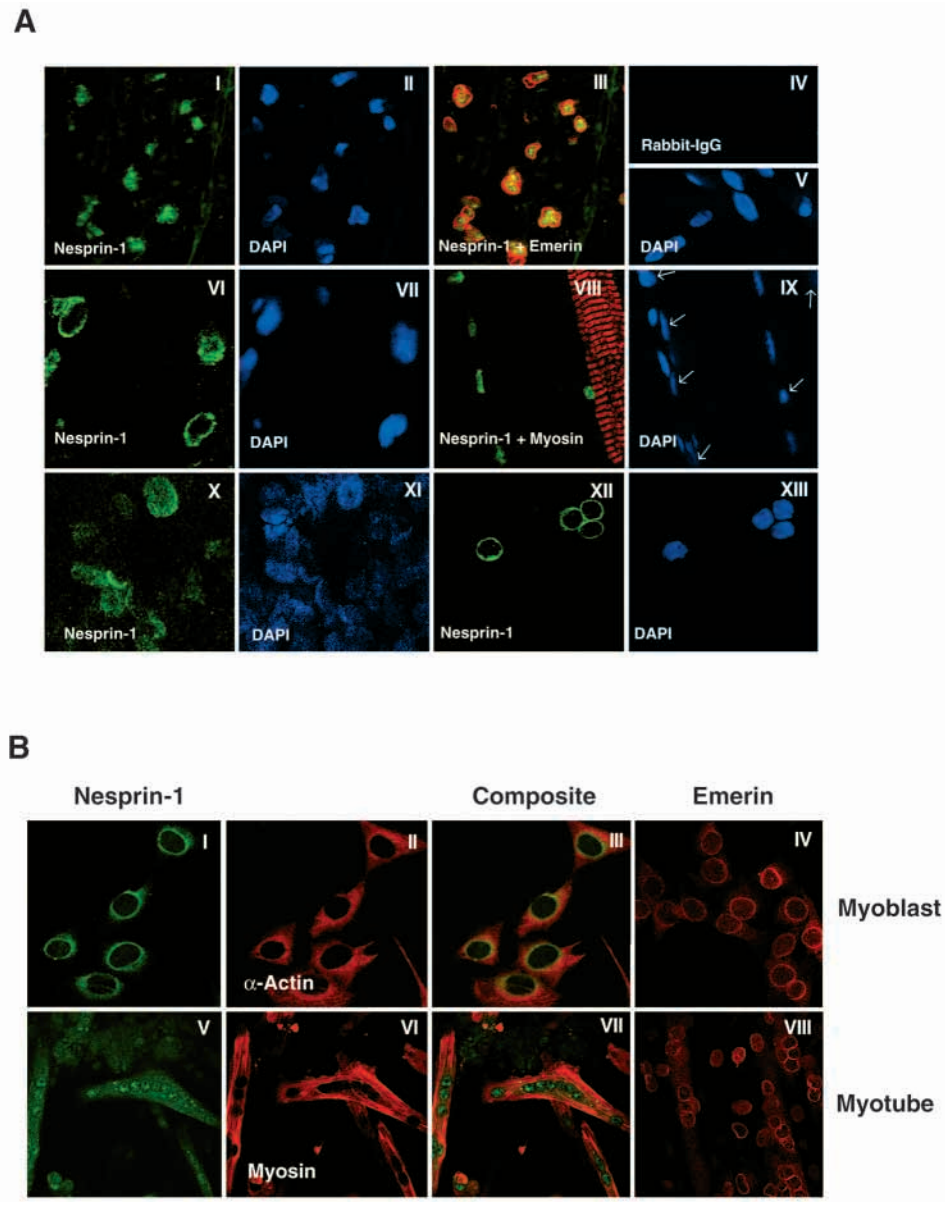


Fig. 10. Subcellular colocalization of Nesprin-1 in human tissues and in C2C12 differentiation. (A) Immunofluorescence of nesprin-1 in human tissues; aorta (I-III), heart (VI-VII), skeletal muscle (VIII-IX), spleen (X-XI) and leukocytes from peripheral blood (XII-XIII). Nesprins are shown in green, emerin and myosin in red, the nucleus in blue and rabbit IgG is a negative control (IV-V). (B) Immunolocalization of nesprin-1 in mouse C2C12 myoblasts (I-III) and myotubes (V-VII). Nesprin-1 is shown in green, emerin (IV and VIII), α-actin and myosin are shown in red.

domain-containing protein, which might reside in the nucleus, and to the hydrophobic tail of Klarsicht, a large protein (2,282 amino acids) in *Drosophila* that has been shown to exhibit a perinuclear localization in larval eye disc cells and has been postulated to target to lipid vesicles in the embryo (Mosley-Bishop et al., 1999; Welte et al., 1998). Klarischt functions as a regulator of lipid vesicle and nuclear migrations during development; however, the localization signals that target it to different subcellular domains have not been identified, although our studies would suggest that the hydrophobic tail is involved.

We demonstrated that mutation or deletion of the 60-residue Klarsicht homology domain of nesprins, in the context of a nesprin-EGFP fusion construct, resulted in the incorrect localization of the construct to the cytoplasm. The deletion that was introduced into the nesprin TM domain sequence was similar to a naturally occurring mutation in the emerin gene (VIVLFF, residues 236-241) of a patient with EDMD that had been shown, in vitro and in vivo, to ablate NE localization of emerin and result in low levels of cytoplasmic emerin (Fairley et al., 1999). Mutation of the TM domain in the context of the short terminal 60-residue domain resulted in only nuclear localization of the EGFP fusion protein. This fusion protein is small enough (approximately 32 kDa) to be able to diffuse via the nuclear pores into the nucleus (60 kDa cut off); however, its retention in the nucleus suggests that this domain contains additional nuclear targeting signals. A number of NE proteins are thought to post-translationally insert into the ER, diffuse through the ER network and become anchored to the inner nuclear membrane via their C-terminal TM domain (Kutay et al., 1993; Soullam and Worman, 1995). Indeed we often observed ER localization of nesprin-1 protein, consistent with this being the mechanism via which nesprins enter the nucleus with the TM domain required for targeting and anchorage to the NE. However, unlike emerin and LAPs, which in the absence of their TM domains have functional bipartite NLS, the NLS in the larger nesprin isoforms, which are well conserved, appear not to be functional at least in the cells we tested (Cartegni et al., 1997). However, there is a possibility that nesprin NLS may function at different developmental stages. For example, the NLS of L-periaxin, a neural filament protein, appear to function only in a specific subset of neural cells and during development; in other cell types L-periaxin is localized to the plasma membrane (Sherman and Brophy, 2000). Similarly, nesprin-1 may use a hierarchical, tissue-specific targeting system, enabling it to target to the nucleus during skeletal muscle differentiation.

A role for nesprins in muscle differentiation?

Nesprin-1 was originally isolated from differentiated, contractile VSMCs, and its subcellular localization was altered during in vitro differentiation of C2C12 myoblasts into myotubes, suggesting that nesprins may perform specific functions in muscle differentiation and have different nuclear and cytoplasmic roles. Multiple functions have been described for other muscle proteins. For example, MLP (muscle-specific LIM protein) enhances myogenic differentiation by associating with MyoD in the nucleus, although it binds to β-spectrin in the cytoplasm and contributes to structural integrity of the cell (Flick and Konieczny, 2000). The mechanisms underlying its alternate localizations are unknown. Interestingly, a mouse cDNA (named syne-1A for synaptic nuclear envelope), homologous to the shortest human nesprin-1α_2 isoform, was identified in a yeast two-hybrid assay by its binding to MuSK (muscle-specific tyrosine kinase), a protein concentrated in post-synaptic plasma membranes. In mouse, using immunohistochemistry, syne-1 protein was localized only to the NE of smooth, cardiac and skeletal muscle but not in other tissues. In skeletal muscle, staining was heterogeneous, as we also observed in human muscle, with greatest staining in myofibres at the neuromuscular junction (NMJ). This data led the authors to speculate that syne-1 had a specific function at the NMJ (hence its name). However, Apel et al. were unable to demonstrate the absence of syne-1 staining in denervated animals or in transgenic animals with defects in NMJ function, nor were they able to demonstrate colocalization of syne-1 with MuSK in Sol8 skeletal muscle cells, thus making it difficult to conclude that the protein has a specific role at the NMJ (Apel et al., 2000). However, it is worth noting that our data on C2C12 cell differentiation would support the hypothesis that in specific circumstances, such as during skeletal muscle differentiation, nesprin-1 may localize to the cytoplasm. Thus, there is some evidence for a tissue-specific role for some nesprin-1 isoforms in skeletal muscle; however our studies have shown that both nesprin-1 and -2 are ubiquitously expressed. Therefore to imply by name 'synaptic specificity' for both these proteins is misleading; this is why we propose a new name which does not imply functional specificity.

Role in muscular dystrophy/human disease?

Nesprins share characteristics with both dystrophin, a cytoskeletal protein implicated in Duchenne and Becker muscular dystrophy, and emerin and lamins A/C, the proteins mutated in EDMD. Thus, it is tempting to speculate that nesprins may be involved in some forms of muscular dystrophy. Nesprins-1 and -2 are located on chromosomes 6q25 and 14q23, respectively, and so far no muscular dystrophy syndrome has been mapped to 6q25. However, an Italian kindred of arrythmogenic right ventricular dystrophy (ARVD) has been mapped to chromosome 14q23, and it may be worthwhile determining if *nesprin-2* is the gene involved in this disorder, which, interestingly, involves substitution of heart-muscle tissue with adipocyte-like tissue (Rampazzo et al., 1994). Future work on nesprins to identify their role in cell function will rely on the identification of binding partners in both the nuclear membrane and the nucleus and the generation of knockout animals to determine their in vivo function.

The authors would like to thank Simon Hughes and Stephanie Bayol for the gift of the C2C12 myoblasts, Andrew Sutherland-Smith for Clustal X alignments and Adrian Newman for photographic assistance. This work was supported by grants from the British Heart Foundation (RG/200004) to C.M.S.

REFERENCES

Amann, K. J., Guo, A. W. and Ervasti, J. M. (1999). Utrophin lacks the rod domain actin binding activity of dystrophin. *J. Biol. Chem.* **274**, 35375-35380.

Apel, E. D., Lewis, R. M., Grady, R. M. and Sanes, J. R. (2000). Syne-1, A dystrophin- and klarsicht-related protein associated with synaptic nuclei at the neuromuscular junction. *J. Biol. Chem.* **275**, 31986-31995.

Ashery-Padan, R., Weiss, A. M., Feinstein, N. and Gruenbaum, Y. (1997). Distinct regions specify the targeting of otefin to the nucleoplasmic side of the nuclear envelope. *J. Biol. Chem.* **272**, 2493-2499.

Bergmann, J. E. and Singer, S. J. (1983). Immunoelectron microscopic studies of the intracellular transport of the membrane glycoprotein (G) of vesicular stomatitis virus in infected Chinese hamster ovary cells. *J. Cell Biol*. **97**, 1777-1787.

Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G. and Toniolo, D. (1994). Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **8**, 323-327.

Bonne, G., Di Barletta, M. R., Varnous, S., Becane, H. M., Hammouda, E. H., Merlini, L., Muntoni, F., Greenberg, C. R., Gary, F., Urtizberea, J. A. et al. (1999). Mutations in the gene encoding lamin A/C cause autosomal dominant Emery-Dreifuss muscular dystrophy. *Nat. Genet.* **21**, 285-288.

Cao, H. and Hegele, R. A. (2000). Nuclear lamin A/C R482Q mutation in canadian kindreds with Dunnigan-type familial partial lipodystrophy. *Hum. Mol. Genet.* **9**, 109-112.

Cartegni, L., di Barletta, M. R., Barresi, R., Squarzoni, S., Sabatelli, P., Maraldi, N., Mora, M., Di Blasi, C., Cornelio, F., Merlini, L. et al. (1997). Heart-specific localization of emerin: new insights into Emery-Dreifuss muscular dystrophy. *Hum. Mol. Genet.* **6**, 2257-2264.

Davis, D. B., Delmonte, A. J., Ly, C. T. and McNally, E. M. (2000). Myoferlin, a candidate gene and potential modifier of muscular dystrophy. *Hum. Mol. Genet*. **9**, 217-226.

De Matteis, M. A. and Morrow, J. S. (2000). Spectrin tethers and mesh in the biosynthetic pathway. *J. Cell Sci.* **113**, 2331-2343.

Dechat, T., Vlcek, S. and Foisner, R. (2000). Lamina-associated polypeptide 2 isoforms and related proteins in cell cycle-dependent nuclear structure dynamics. *J. Struct. Biol.* **129**, 335-345.

Ellis, J. A., Craxton, M., Yates, J. R. and Kendrick-Jones, J. (1998). Aberrant intracellular targeting and cell cycle-dependent phosphorylation of emerin contribute to the Emery-Dreifuss muscular dystrophy phenotype. *J. Cell Sci.* **111**, 781-792.

Fairley, E. A., Kendrick-Jones, J. and Ellis, J. A. (1999). The Emery-Dreifuss muscular dystrophy phenotype arises from aberrant targeting and binding of emerin at the inner nuclear membrane. *J. Cell Sci.* **112**, 2571-2582.

Fatkin, D., MacRae, C., Sasaki, T., Wolff, M. R., Porcu, M., Frenneaux, M., Atherton, J., Vidaillet, H. J., Spudich, S., De Girolami, U. et al. (1999). Missense mutations in the rod domain of the lamin A/C gene as causes of dilated cardiomyopathy and conduction-system disease. *N. Engl. J. Med.* **341**, 1715-1724.

Flick, M. J. and Konieczny, S. F. (2000). The muscle regulatory and structural protein MLP is a cytoskeletal binding partner of betaI-spectrin. *J. Cell Sci*. **113**, 1553-1564.

Foisner, R. and Gerace, L. (1993). Integral membrane proteins of the nuclear envelope interact with lamins and chromosomes, and binding is modulated by mitotic phosphorylation. *Cell* **73**, 1267-1279.

Furukawa, K., Fritze, C. E. and Gerace, L. (1998). The major nuclear envelope targeting domain of LAP2 coincides with its lamin binding region but is distinct from its chromatin interaction domain. *J. Biol. Chem.* **273**, 4213-4219.

Furukawa, K., Pante, N., Aebi, U. and Gerace, L. (1995). Cloning of a cDNA for lamina-associated polypeptide 2 (LAP2) and identification of regions that specify targeting to the nuclear envelope. *EMBO J.* **14**, 1626-1636.

Gerace, L. and Burke, B. (1988). Functional organization of the nuclear envelope. *Annu. Rev. Cell Biol.* **4**, 335-374.

Greener, M. J. and Roberts, R. G. (2000). Conservation of components of the dystrophin complex in *Drosophila*. *FEBS Lett.* **482**, 13-18.

Gregory, S. L. and Brown, N. H. (1998). kakapo, a gene required for adhesion between and within cell layers in Drosophila, encodes a large cytoskeletal linker protein related to plectin and dystrophin. *J. Cell Biol.* **143**, 1271-1282.

Grum, V. L., Li, D., MacDonald, R. I. and Mondragon, A. (1999). Structures of two repeats of spectrin suggest models of flexibility. *Cell* **98**, 523-535.

Hartmann, E., Rapoport, T. A. and Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc. Natl. Acad. Sci.USA* **86**, 5786-5790.

Jackle, H. and Jahn, R. (1998). Vesicle transport: klarsicht clears up the matter. *Curr. Biol.* **8**, R542-544.

Jans, D. A., Xiao, C.-Y. and Lam, M. H. C. (2000). Nuclear targeting signal recognition: a key control point in nuclear transport? *BioEssays* **22**, 532-544.

Kapiloff, M. S., Schillace, R. V., Westphal, A. M. and Scott, J. D. (1999). mAKAP: an A-kinase anchoring protein targeted to the nuclear membrane of differentiated myocytes. *J. Cell Sci.* **112**, 2725-2736.

Kutay, U., Hartmann, E. and Rapoport, T. (1993). A class of membrane proteins with a C-terminal anchor. *Trends Cell Biol*, **3**, 72-75.

Lin, F., Blake, D. L., Callebaut, I., Skerjanc, I. S., Holmer, L., McBurney, M. W., Paulin-Levasseur, M. and Worman, H. J. (2000). MAN1, an inner nuclear membrane protein that shares the LEM domain with lamina-associated polypeptide 2 and emerin. *J. Biol. Chem.* **275**, 4840-4847.

Mosley-Bishop, K. L., Li, Q., Patterson, L. and Fischer, J. A. (1999). Molecular analysis of the klarsicht gene and its role in nuclear migration within differentiating cells of the *Drosophila* eye. *Curr. Biol*. **9**, 1211-1220.

Muchir, A., Bonne, G., van der Kooi, A. J., van Meegen, M., Baas, F., Bolhuis, P. A., de Visser, M. and Schwartz, K. (2000). Identification of mutations in the gene encoding lamins A/C in autosomal dominant limb girdle muscular dystrophy with atrioventricular conduction disturbances (LGMD1B). *Hum. Mol. Genet.* **9**, 1453-1459.

Nagase, T., Kikuno, R., Ishikawa, K., Hirosawa, M. and Ohara, O. (2000). Prediction of the coding sequences of unidentified human genes. XVII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res.* **7**, 143-150.

Rampazzo, A., Nava, A., Danieli, G. A., Buja, G., Daliento, L., Fasoli, G., Scognamiglio, R., Corrado, D. and Thiene, G. (1994). The gene for arrhythmogenic right ventricular cardiomyopathy maps to chromosome 14q23-q24. *Hum. Mol. Genet.* **3**, 959-962.

Rando, O. J., Zhao, K. and Crabtree, G. R. (2000). Searching for a function for nuclear actin. *Trends Cell Biol.* **10**, 92-97.

Roberts, R. (2001). Dystrophins and dystrobrevins. *Genome Biol.* **2**, 3001-3007.

Rolls, M. M., Stein, P. A., Taylor, S. S., Ha, E., McKeon, F. and Rapoport, T. A. (1999). A visual screen of a GFP-fusion library identifies a new type of nuclear envelope membrane protein. *J. Cell Biol.* **146**, 29-44.

Schwartz, S., deBlois, D. and O'Brien, E. (1995). The intima: soil for atherosclerosis and restenosis. *Circ. Res.* **77**, 445-465.

Shanahan, C. and Weissberg, P. (1998). Smooth muscle cell heterogeneity. Patterns of gene expression in vascular smooth muscle cells in vitro and in vivo. *Arterioscler. Thromb. Vasc. Biol.* **18**, 333-338.

Shanahan, C., Weissberg, P. and Metcalfe, J. (1993). Isolation of markers of differentiated and proliferating vascular smooth muscle cells. *Circ. Res.* **73**, 193-204.

Shanahan, C. M., Cary, N. R., Salisbury, J. R., Proudfoot, D., Weissberg, P. L. and Edmonds, M. E. (1999). Medial localization of mineralization-regulating proteins in association with Monckeberg's sclerosis: evidence for smooth muscle cell-mediated vascular calcification. *Circulation* **100**, 2168-2176.

Sherman, D. L. and Brophy, P. J. (2000). A tripartite nuclear localization signal in the PDZ-domain protein L-periaxin. *J. Biol. Chem.* **275**, 4537-4540.

Soullam, B. and Worman, H. J. (1995). Signals and structural features involved in integral membrane protein targeting to the inner nuclear membrane. *J. Cell Biol.* **130**, 15-27.

Strumpf, D. and Volk, T. (1998). Kakapo, a novel cytoskeletal-associated protein is essential for the restricted localization of the neuregulin-like factor, vein, at the muscle-tendon junction site. *J. Cell Biol.* **143**, 1259-1270.

Stuurman, N., Heins, S. and Aebi, U. (1998). Nuclear lamins: their structure, assembly, and interactions. *J. Struct. Biol.* **122**, 42-66.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.

Vlcek, S., Just, H., Dechat, T. and Foisner, R. (1999). Functional diversity of LAP2alpha and LAP2beta in postmitotic chromosome association is caused by an alpha-specific nuclear targeting domain. *EMBO J.* **18**, 6370-6384.

Welte, M. A., Gross, S. P., Postner, M., Block, S. M. and Wieschaus, E. F. (1998). Developmental regulation of vesicle transport in Drosophila embryos: forces and kinetics. *Cell* **92**, 547-557.

Wilson, K. L. (2000). The nuclear envelope, muscular dystrophy and gene expression. *Trends Cell Biol.* **10**, 125-129.

Winder, S. J., Gibson, T. J. and Kendrick-Jones, J. (1995). Dystrophin and utrophin: the missing links! *FEBS Lett.* **369**, 27-33.

Xia, H., Winokur, S. T., Kuo, W. L., Altherr, M. R. and Bredt, D. S. (1997). Actinin-associated LIM protein: identification of a domain interaction between PDZ and spectrin-like repeat motifs. *J. Cell Biol*. **139**, 507-515.

Yang, F., Graphodatsky, A. S., O'Brien, P. C., Colabella, A., Solanky, N., Squire, M., Sargan, D. R. and Ferguson-Smith, M. A. (2000). Reciprocal chromosome painting illuminates the history of genome evolution of the domestic cat, dog and human. *Chromosome Res.* **8**, 393-404.