PRELI (protein of relevant evolutionary and lymphoid interest) is located within an evolutionarily conserved gene cluster on chromosome 5q34–q35 and encodes a novel mitochondrial protein

Elizabeth J. FOX*, Sally A. STUBBS*, Jimmy KYAW TUN*, Jack P. LEEK†, Alexander F. MARKHAM† and Stephanie C. WRIGHT*[1]

*School of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT, U.K., and †Molecular Medicine Unit, University of Leeds, St. James's Hospital, Leeds LS9 7TF, U.K.

The characterization of mitochondrial proteins is important for the understanding of both normal cellular function and mitochondrial disease. In the present study we identify a novel mitochondrial protein, PRELI (protein of relevant evolutionary and lymphoid interest), that is encoded within the evolutionarily conserved *MAD3*/*PRELI*/*RAB24* gene cluster located at chromosome 5q34–q35. Mouse *Preli* is expressed at high levels in all settings analysed; it is co-expressed with *Rab24* from a strong bi-directional promoter, and is regulated independently from the S-phase-specific *Mad3* gene located at its 3′ end. PRELI contains a stand-alone 170 amino acid PRELI/MSF1p′ motif at its N-terminus. This domain is found in a variety of proteins from diverse eukaryotes including yeast, *Drosophila* and mammals, but its function is unknown, and the subcellular location of higher eukaryotic PRELI/MSF1P′ proteins has not been determined previously. We show here that PRELI is located in the mitochondria, and by using green-fluorescent-protein fusion proteins we identify a mitochondrial targeting signal at its N-terminus.

Key words: cell cycle, differentiation, Mad, mitochondrial-targeting signal, PRELI/MSF1p′ domain.

INTRODUCTION

Aberrations on the long arm of chromosome 5 have been implicated in a number of human malignancies; the characterization of novel genes within this region is therefore crucial for the identification of candidate disease loci. Mouse *Mad3* is located in the central region of chromosome 13 at a site that is syntenic to human 5q33–q35 [1]. It encodes a basic helix–loop–helix leucine zipper (bHLH-zip) protein that functions as a transcriptional repressor within the Myc/Max/Mad network (reviewed in [2]). *Mad3* is expressed in the S-phase of the cell cycle [3–5], and targeted disruption of the gene results in an increased sensitivity to certain apoptotic stimuli [5]. Although it has been suggested that *MAD3* acts as a tumour suppressor at 5q33–q35, there is no experimental evidence for this hypothesis; it is therefore essential to identify and characterize other genes within this region. In this report, we show that mouse *Mad3* is located in a gene cluster that contains *Rab24* and the previously uncharacterized *Preli* (protein of relevant evolutionary and lymphoid interest) gene. The organisation of the gene cluster is conserved between mouse and human, and the human genes are located at 5q34–q35. *Preli* is not co-regulated with the S-phase-specific *Mad3* gene, but is abundantly expressed in all settings analysed, and is co-ordinately transcribed with *Rab24* from a strong bi-directional promoter. Preli contains the recently described 170 amino acid PRELI/MSF1p′ motif [17] at its N-terminus (where 'MSF' is derived from 'a gene whose overexpression perturbs the intramitochondrial sorting of a certain fusion protein' [18]); the subcellular location of higher eukaryotic PRELI/MSF1p′ proteins has not previously been determined, and the biological function of this motif is unknown. We show here that mouse Preli is a novel mitochondrial protein and identify a 35-amino-acid mitochondrial-targeting signal at its N-terminus.

EXPERIMENTAL

Cell culture and transfection

C7 3T3 fibroblasts, mouse erythroleukaemia (MEL) cells and HeLa cells were grown in minimum essential medium alpha supplemented with 10 % calf serum. MEL cells were induced to differentiate by the addition of DMSO to 2 %. C7 3T3 cells were made quiescent by culturing for 28 h in serum-free medium, and induced to re-enter the cell cycle by the addition of calf serum to 10 %. They were stably transfected with 20 μg of *rab24-preli-neo*^{r} (where *neo*^{r} is the coding region of the Tn5 aminoglycoside phosphotransferase neomycin resistance gene) construct per 10-cm-diameter plate using the calcium phosphate precipitation method [6], and a pool of approx. 200 colonies was generated by selection in 1 mg/ml G418. HeLa cells were transiently transfected using 10 μg of plasmid DNA per 5-cm-diameter plate, and analysed by fluorescence microscopy 24 h later. Green fluorescent protein (GFP) fusions were made by cloning Preli cDNA fragments into pEGFP-N1 or pEGFP-C1 (ClonTech Laboratories, U.K. Limited, Basingstoke, U.K.).

RNA preparation and ribonuclease protection analysis

Total cellular RNA was isolated by the guanidinium isothiocyanate method [7] and analysed by hybridization to uniformly labelled riboprobes as described previously [8]. Hybridizations contained 10 μg RNA and were carried out at 50 °C. Markers

Abbreviations used: AML, acute myeloid leukaemia; bHLH-zip, basic helix–loop–helix leucine zipper; DAPI, 4,6-diamidino-2-phenylindole; FISH, fluorescence *in situ* hybridization; GFP, green fluorescent protein; LEA, late embryonic abundant; MDS, myelodysplastic syndrome; MEL, mouse erythroleukaemia; *neo*^{r}, coding region of the Tn5 aminoglycoside phosphotransferase neomycin resistance gene; PRELI/Preli, protein of relevant evolutionary and lymphoid interest; RACE, rapid amplification of cDNA ends; UTR, untranslated region.

[1] To whom correspondence should be addressed (e-mail s.c.wright@leeds.ac.uk).

were end-labelled *Msp*I restriction fragments of the plasmid pBR322. Plasmid constructs used for the generation of riboprobes were made by cloning the appropriate restriction fragments into pSP72 (Promega, Chilworth, Southampton, U.K.). Constructs were linearized so as to allow the generation of single-stranded probes that detected either sense or antisense transcription across *Mad3*, *Preli* or *Rab24* as indicated in the Figure legends (Figures 1 and 4A). Riboprobes described in the Figure 3 legend comprised the following sequences cloned into pSP72: *Mad3* cDNA +290 to +621; mouse *α5-tubulin* cDNA +63 to +187; *Preli* cDNA −5 to +910; *Rab24* cDNA +390 to +596 (cDNA sequences are relative to a designation of +1 for the ATG translation initiation codon).

Isolation of mouse *Mad3* and *Preli* genomic clones

Mouse *Mad3* genomic clones were isolated by screening a lambda library (B6/CBA; Stratagene, Cambridge, U.K.) with a *Mad3* cDNA probe (coding region +1 to +621) using protocols described by the manufacturer. Clones containing the 3′ end of *Mad3* were identified by hybridization to a probe corresponding to nt +555 to +621 of *Mad3* cDNA, and restriction endonuclease fragments containing *Preli*, *Rab24* and the 3′ end of *Mad3* were subcloned into pSP72 for further analysis. Automated DNA sequencing was carried out by Lark Technologies (Lark Technologies Inc., Saffron Walden, Essex, U.K.).

Isolation of the human *MAD3/PRELI/RAB24* gene cluster

Human genomic clones were isolated by screening the RPCI1 Human PAC library ([9]; UK HGMP Resource Centre) using a human *MAD3* cDNA (−5 to +240) probe. Clone RP1-128M18 was used for further analysis. The *MAD3-PRELI* intergenic region was isolated from Clone RP1-128M18 by PCR using the primer pair: 5′-GAGGGTACCGAGGGTGGATGTGGAGAGCCTGGTGT-3′ (*MAD3* sequence) and 5′-GAGGAATTCGCAGCAACAGTTTGTGTAGCCAGTCT-3′ (*PRELI* sequence). The *PRELI-RAB24* intergenic region was isolated from Clone RP1-128M18 by PCR using the primer pair: 5′-GAGGGATCCTCCCGGGGCCGCTTCAGCCACGTCAG-3′ (*RAB24* sequence) and 5′-GAGGAATTCGGCCGAAGCGGGCTCAGCACCCGCGC-3′ (*PRELI* sequence).

Rapid amplification of cDNA ends (RACE)

Mouse *Preli* cDNA was isolated from a Marathon-Ready adaptor-ligated cDNA library (mouse 17-day embryo; Clontech) by RACE. PCR amplification was carried out with the *Preli* primer 5′-GTGGGATCCTTGGGACACAGCAGAGAATAAAACAAGGAG-3′, using protocols recommended by the manufacturer. The PCR product was cloned into pSP72.

Fluorescence *in situ* hybridization (FISH)

Metaphase cells were prepared from human peripheral blood lymphocytes using standard cytogenetic techniques. The FISH procedure was used as described previously [10], with minor modifications. The genomic clone, PAC RP1-128M18, was labelled with digoxigenin-11-dUTP (Roche, Lewes, East Sussex, U.K.) by nick translation. Denatured probe was pre-annealed with human Cot-1 DNA (Invitrogen, Paisley, U.K.) and then hybridized with denatured chromosomes for 16 h at 37 °C. Subsequent post hybridization washes were with 50 % formamide, 2 × SSC (where SSC is 0.15 M NaCl/0.015 M sodium citrate) at 42 °C and 0.1 × SSC at 60 °C, following which fluorescein-conjugated anti-digoxigenin antibody (Roche) was used for the detection of hybridization sites. After counterstaining the chromosomes with DAPI (4,6-diamidino-2-phenylindole; Sigma), fluorescent signals were captured separately as greyscale images using a Zeiss Axioskop fluorescence microscope equipped with a CCD camera (Photometrics, Tuscon, Arizona, U.S.A.). These images were then merged and pseudocoloured using Smart Capture VP software, version 1.3.1. (Vysis, Richmond, Surrey, U.K.). Chromosomal identification and band assignment for the probe hybridization signal was enabled using this software. The location of the probe signal on the reverse DAPI image of chromosomes was analysed on 30 metaphase spreads.

Generation of a reporter construct containing the mouse *Preli-Rab24* bi-directional promoter

A 371 nt *Sma*I fragment containing the *Preli-Rab24* intergenic region was inserted into pSP72. A fragment comprising the *neo*^{r} gene together with the SV40 polyadenylation signal was inserted adjacent to the *Sma*I fragment such that the *Preli* promoter directed the transcription of *neo*^{r}. A cassette comprising the mouse *Mad3* cDNA coding region, together with the rabbit β-globin polyadenylation signal was then inserted such that the *Rab24* promoter drove the expression of *Mad3*.

Antibodies, Western-blot and fluorescence-microscopy analysis

Full-length mouse *Preli* cDNA was inserted into the bacterial expression vector pRSET (Invitrogen), and antibodies generated by injection of sheep with recombinant His-tagged protein (Diagnostics Scotland, Ellen's Glen Road, Edinburgh, Scotland). Antiserum was purified by affinity chromatography [11]. Whole cell extracts were prepared by lysis of 10^7 cells in 200 μl of 137 mM NaCl, 2.7 mM KCl, 1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 1 % Nonidet P-40, 10 mM NaF, 1 mM sodium orthovanadate, 5 mM sodium pyrophosphate plus protease inhibitors ('Complete' tablets; Roche) [12]. Samples (50 μg) were electrophoresed by SDS/PAGE (10 % gel) and blotted onto nitrocellulose membranes (Schleicher and Schuell) as described previously [11]. Membranes were blocked using PBS containing 5 % non-fat dried milk and incubated with anti-Preli antiserum (1:1000–1:10000 dilution). They were rinsed with PBS and then incubated with peroxidase-conjugated secondary antibody (1:20000 dilution of AffiniPure Donkey Anti-Sheep IgG; Jackson Immunoresearch Laboratories, West Grove, PA, U.S.A.). Blotted proteins were detected using the SuperSignal West Pico Chemiluminescent Substrate (Pierce Biotechnology Inc., Tattenhall, Cheshire, U.K.). Mammalian cells were immunostained using affinity-purified anti-Preli antibody (5 μg/ml) and visualized by fluorescence microscopy using the secondary antibody FITC-conjugated AffiniPure donkey anti-sheep IgG (1:200 dilution, Jackson ImmunoResearch Laboratories). Nuclei were stained using DAPI (1 μg/ml) and mitochondria were visualized using the mitochondrion-specific fluorophore MitoTracker Red CMXRos (10 nM; Molecular Probes). Fluorescence microscopy was carried out using the DeltaVision Optical Restoration Microscopy System (Applied Precision Inc., Issaquah, WA, U.S.A.). Data were collected from 25–30 optical sections of 0.2 μm thick, and three-dimensional datasets were deconvoluted using the default settings on the softWoRx deconvolution algorithm (Applied Precision Inc., Issaquah, WA, U.S.A). Deconcolved images were analysed using the Z-viewer, and selected optical sections were projected as two-dimensional images.

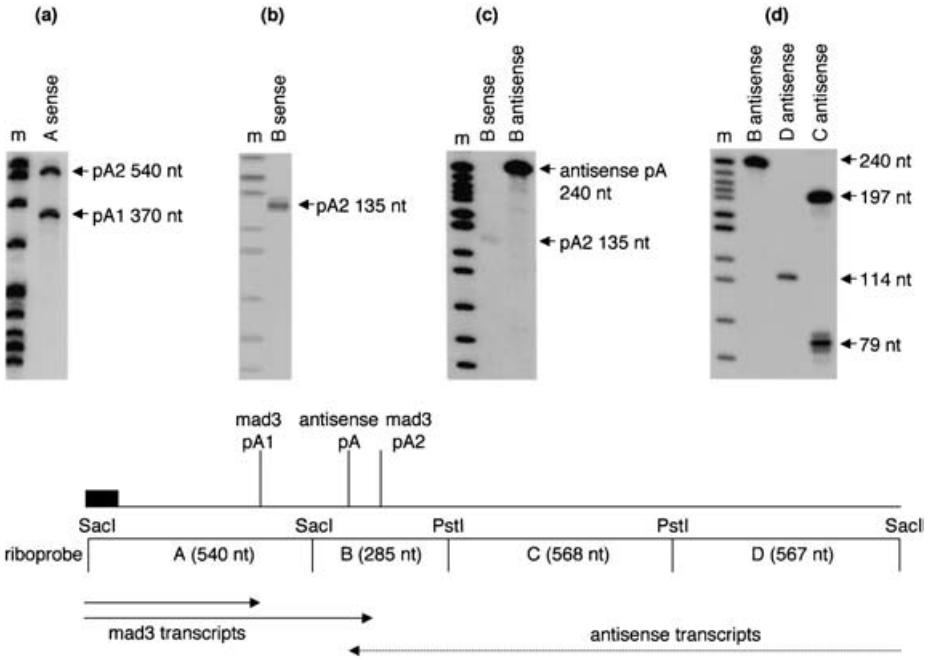

Figure 1 Detection of transcripts derived from the 3′ flanking region of mouse *Mad3*

Restriction endonuclease fragments A, B, C and D were subcloned into pSP72 and the resulting constructs were used to generate riboprobes that detected transcripts in either the sense or antisense direction with respect to *Mad3*. (**a** and **b**) The 3′ end of the mouse *Mad3* transcript was mapped by hybridization of samples of C7 3T3 RNA to riboprobes A (A sense) and B (B sense). (**c** and **d**) Antisense transcripts were detected using riboprobes B (B antisense), C (C antisense) and D (D antisense). No ribonuclease protected bands were observed upon hybridization of the riboprobes to *Escherichia coli* tRNA. m, markers. The schematic depicts the 3′ flanking region of *Mad3*. The solid box represents the 3′ end of the *Mad3* exon 6 coding region. pA1 and pA2 denote the *Mad3* polyadenylation signals, and antisense pA denotes the 3′ end of RNAs transcribed in the opposite direction to *Mad3*. The precise location of the polyadenylation signals and the exact sizes of ribonuclease protected fragments were assigned following sequence analysis of cDNA and genomic clones (see Figure 2).

Preparation of mitochondrial and cytosolic extracts

Cells were rinsed with cold PBS, pelleted by centrifugation at 500 **g** and resuspended in 3 packed-cell volumes of 20 mM Tris/HCl, pH 7.4, 0.25 M sucrose, 2 mM EGTA plus protease inhibitors ('Complete' tablets, Roche). Cells were lysed using 60 strokes of a Dounce homogeniser, and nuclei were pelleted by centrifugation at 700 **g** for 5 min at 4 °C. The supernatant was transferred to a fresh tube, and the mitochondria were pelleted by centrifugation at 10000 **g** for 30 min at 4 °C. The supernatant (cytosolic fraction) was snap-frozen in liquid nitrogen. The mitochondrial pellet was resuspended in the same cold buffer as above and snap-frozen.

RESULTS AND DISCUSSION

Characterization of the *MAD3/PRELI/RAB24* gene cluster

In order to initially characterize the mouse *Mad3* locus, a genomic *Mad3* clone was isolated from a lambda library. A 2.7 kb *Hin*dIII restriction endonuclease fragment containing the 3′ end of the gene was subcloned and used in further analysis to locate the 3′ end of the *Mad3* transcript. Ribonuclease-protection mapping was carried out using probes derived from sequences located immediately downstream from the translation termination codon; this indicated that *Mad3* has two sites of 3′-end formation (Figures 1a and 1b). The DNA sequence of the *Mad3* genomic clone revealed that these sites correspond to the use of potential polyadenylation signals located 290 nt and 588 nt downstream from the translation termination codon. During this analysis, ribonuclease protection mapping was also carried out using a probe designed to detect RNAs that are transcribed in the orientation antisense to *Mad3*. Surprisingly this probe detected an abundant RNA, which was transcribed in the opposite direction to *Mad3*, and that potentially overlapped the 3′ end of the *Mad3* transcript (Figure 1c; note relative signal intensities in lanes 1 and 2). Examination of the nucleotide sequence of the corresponding genomic region revealed that the size of the ribonuclease-protected fragment correlated with the use of a potential polyadenylation signal in the opposite orientation to that of *Mad3* (Figure 2A). Ribonuclease-protection analysis was then carried out using a series of probes designed to detect RNAs derived from the genomic region located downstream from *Mad3* in the direction antisense to *Mad3* transcription (Figures 1c and 1d). These probes also detected an abundant transcript, indicating that the region 3′ to *Mad3* may contain a genuine gene. In order to identify this novel transcript, the corresponding cDNA was cloned by 5′ RACE using a primer complementary to the 3′ end of the RNA (Figure 2A). Sequence analysis of the cloned 1 kb RACE product revealed that it corresponds to the mouse orthologue of human *PRELI* [13] and avian *Px19* [14]. The amino acid sequence alignments (Figure 2B) show a high degree of conservation between the human, mouse and avian proteins. Human *PRELI* was originally isolated from a B-lymphocyte specific cDNA library [13], whereas the avian gene was identified as a bromodeoxyuridine-sensitive transcript from an immortalised quail heart cell line [14]. *In situ* hybridization studies of developing chick embryos have suggested that Px19 plays a role in haematopoiesis during development; the gene is otherwise not well characterized and encodes a protein of unknown function.

In order to confirm that mouse *Preli* is situated adjacent to the *Mad3* gene, the DNA sequence of the genomic region located 3′ to *Mad3* was determined. Comparison of the genomic sequence

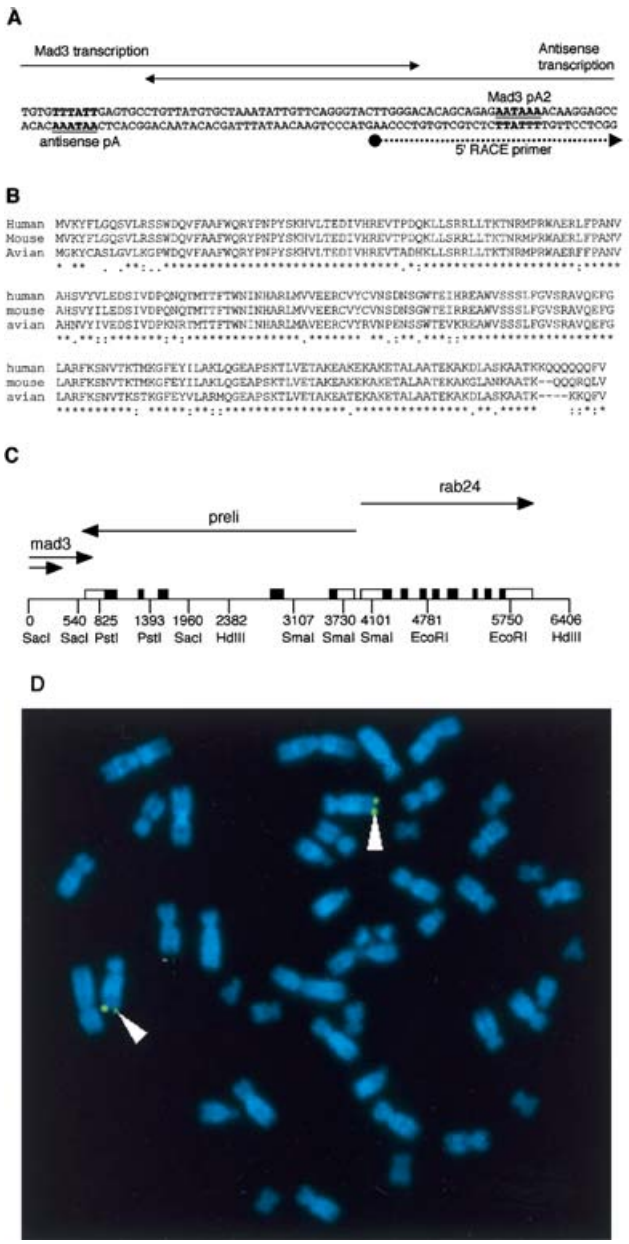

Figure 2 Characterization of the mouse *Mad3/Preli/Rab24* gene cluster

(**A**) Nucleotide sequence of the 3′ end of mouse *Mad3*. The DNA sequence of the *Mad3* 3′ flanking region is depicted. The *Mad3* pA2 site, the putative antisense pA site, and the location of the primer used for the 5′ RACE are indicated. (**B**) Amino acid alignment of human PRELI (Human; GenBank Accession Number AF201925), mouse Preli (Mouse), and avian Px19 (Avian; GenBank Accession Number GGU31977) using the ClustalW Multiple Sequence Alignment programme. Conserved and similar residues are indicated with asterisks and dots respectively. (**C**) Genomic organisation of mouse *Mad3, Preli*, and *Rab24*. The genomic organisation was deduced by comparison of cDNA and genomic sequences, and confirmed by ribonuclease protection mapping using RNA derived from C7 3T3 fibroblasts. Coding exons and untranslated regions of *Preli* and *Rab24* are depicted as filled boxes and open boxes respectively. The directions of *Mad3, Preli* and *Rab24* transcription are indicated. The positions of various restriction enzyme sites are also indicated. HdIII, *Hin*dIII. (**D**) The human *MAD3/PRELI/RAB24* gene cluster is located at chromosome 5q34–q35. Normal human metaphase spread showing localisation of the genomic PAC clone RP1-128M18 to chromosome 5q34–q35 (arrowheads).

with that of *Preli* cDNA revealed that the mouse *Preli* gene spans about 3.5 kb and is located on five exons (Figure 2C). The genomic exon sequences and 3′ untranslated region (UTR) were identical to those of the *Preli* RACE product; all exon–intron boundaries conformed to the GT-AG consensus. Ribonuclease protection mapping experiments (Figure 1 and results not shown) were consistent with the exon organisation as deduced from DNA sequence analysis. It was therefore concluded that mouse *Mad3* lies adjacent to the *Preli* gene and that the organisation of the 3′ ends of the transcripts was as predicted by the initial ribonuclease protection analyses (Figures 1 and 2A).

DNA sequence analysis of the genomic region located 3′ to mouse *Preli* revealed that it contains the gene encoding *Rab24*, as deduced by comparison with the published mouse *Rab24* cDNA sequence (Figure 2C). *Rab24* spans approx. 1.5 kb, comprises eight exons, and is transcribed in the opposite orientation to *Preli*. The ATG translation initiation codons of *Preli* and *Rab24* are about 540 nt apart, indicating that these two genes may be transcribed from a bi-directional promoter.

In order to determine whether the genomic organisation of *Mad3*, *Preli* and *Rab24* is conserved between mouse and human, human *MAD3* clones were isolated by screening the RPCI1 Human PAC library using a probe derived from the 5′ end of the human *MAD3* cDNA. Clone RP1-128M18 was used in further analysis. Southern-blot analysis showed that this clone hybridized to probes derived from the 5′ and 3′ ends of both human *MAD3* and *PRELI* cDNAs, indicating that the corresponding genes were intact. The intergenic region between human *MAD3* and *PRELI* was isolated from RP1-128M18 by PCR using primers whose sequences were derived from the most 3′ exons of published *MAD3* and *PRELI* cDNA sequences. The PCR yielded a 1 kb product whose sequence indicated that it corresponded to the *PRELI-MAD3* intergenic region. Ribonuclease protection mapping using probes derived from the cloned PCR product showed that this region contained the 3′ ends of both *MAD3* and *PRELI* (results not shown). The 3′ ends of *MAD3* and *PRELI* are thus organised similarly in human and mouse, however, in contrast to the mouse genes, human *MAD3* has one polyadenylation signal whereas *PRELI* has two. The human *PRELI-RAB24* intergenic region was similarly isolated from RP1-128M18 by PCR using primers whose sequences were derived from the 5′ ends of published human *PRELI* and *RAB24* cDNAs. DNA-sequence analysis of the cloned PCR product indicated that it corresponded to the human *PRELI-RAB24* intergenic region (see Figure 4B), thus confirming that the organisation of *Mad3, Preli* and *Rab24* is conserved between mouse and human. The mouse *Mad3* gene has been previously assigned to the central region of chromosome 13; it is tightly linked to the fibroblast-growth-factor receptor (*Fgfr4)* gene, and is located between the G-protein-coupled receptor 15 (*Gpc15)* and interleukin 9 (*Il9)* genes [1]. This region is predicted to be syntenic to human chromosome 5q33–q35. Since human chromosome 5q has been implicated in a large number of haematological malignancies, it was relevant to assign the *MAD3*/*PRELI*/*RAB24* gene cluster to this region. The human PAC RP1-128M18 was used as a probe in FISH analysis using human metaphase chromosomes. A single strong hybridization signal was revealed at 5q34–q35; no other nonrandom signals or background hybridization signals were observed (Figure 2D).

Mad3, Preli and *Rab24* are not co-ordinately regulated during the cell cycle or differentiation

Since genes that are located in clusters may serve a common biological function, we determined whether the expression of *Preli*, *Mad3* and *Rab24* is co-ordinately regulated. We, and others, have shown that *Mad3* is specifically expressed in the S-phase of the cell cycle in both proliferating and differentiating cells

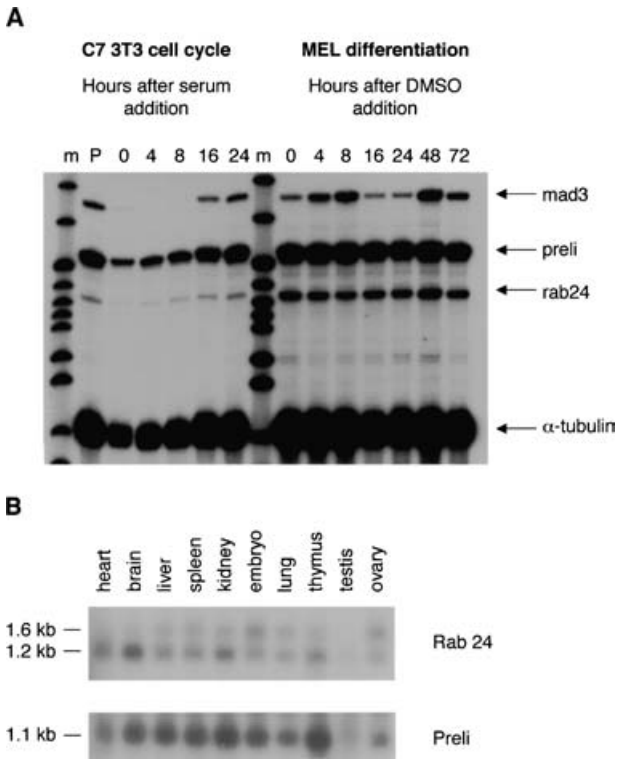

Figure 3 *Mad3, Preli* and *Rab24* are not co-ordinately regulated during the cell cycle or differentiation

(**A**) Expression of *Mad3, Preli* and *Rab24* during the cell cycle and differentiation. C7 3T3 cells were made quiescent by serum withdrawal and then stimulated to re-enter the cell cycle by the addition of serum to 10 %. MEL cells were induced to differentiate by the addition of DMSO to 2 %. Levels of expression of *Mad3* (mad3), *Preli* (preli), *Rab24* (rab24) and *α-tubulin* (α-tubulin) were measured by ribonuclease protection analysis. m, size markers. (**B**) Levels of *Preli* and *Rab24* RNA in mouse tissues. A mouse multiple tissue Northern blot (Ambion) was hybridized sequentially to riboprobes corresponding to the coding regions of *Preli* (Preli) or *Rab24* (Rab 24).

[3,4]; the expression of *Preli, Rab24* and *Mad3* was therefore monitored throughout the cell cycle in C7 3T3 fibroblasts and during differentiation in cultured MEL cells. C7 3T3 fibroblasts were synchronised in G_o by serum deprivation, and induced to re-enter the cell cycle by the addition of serum to 10 %. Expression of *Mad3, Preli, Rab24*, and *α-tubulin* were measured throughout the cell cycle by ribonuclease protection mapping (Figure 3A); progress through the cell cycle was monitored by FACS analysis and DNA synthesis assays (results not shown). The expression of *Mad3* was low in serum-starved quiescent cells and induced at the onset of S-phase (16 h) in agreement with previous observations [3]. The expression of *Preli* and *Rab24* was not cell-cycle regulated, but showed a modest increase at early stages in the cell cycle that reflected the expression pattern of the control *α-tubulin* gene. The expression of *Preli* and *Rab24* was similarly monitored during the DMSO-induced differentiation of MEL cells (Figure 3A); levels of *Mad3* showed a characteristic bi-phasic increase [3], whereas the expression of *Preli* and *Rab24* remained constant throughout. These data, therefore, indicated that *Mad3, Preli* and *Rab24* are not co-ordinately regulated during either the cell cycle or differentiation. Since *Preli* mRNA is expressed at much higher levels than *Mad3*, in all settings analysed, it can be inferred that the expression pattern of *Mad3* is not influenced by strong transcriptional control elements in the adjacent *Preli* gene.

We determined next whether *Preli* and *Rab24* have similar tissue-specific expression profiles. A commercially available mouse multiple-tissue Northern blot was hybridized to probes corresponding to *Preli* and *Rab24* (Figure 3B). Both genes were abundantly expressed in all tissues except the testis; the expression of *Rab24* was slightly higher in the brain as compared with other tissues, whereas the expression of *Preli* was slightly increased in the thymus. These results are in approximate agreement with those observed for human *PRELI*, although two mRNA species were detected on Northern blots of human *PRELI*, and the human gene is expressed at a relatively lower level in the adult liver [13]. The similar expression profiles of *Preli* and *Rab24* suggest that their transcription is, to a large extent, co-ordinately regulated.

Identification of a bi-directional promoter for *Preli* and *Rab24*

The close proximity of *Preli* and *Rab24* suggests that these genes may be expressed from a bi-directional promoter. In order to initially locate the transcription start sites for mouse *Preli* and *Rab24*, ribonuclease protection mapping was carried out using probes corresponding to a 371 nt *Sma*I fragment derived from the intergenic region (Figure 4A). Riboprobes that detected transcripts in either the *Preli* or *Rab24* direction were used. This indicated that the 371 nt region contains the transcription start sites for both *Preli* and *Rab24*, and that these are about 60 nt apart. The promoters are GC-rich and TATA-less, and the nucleotide sequence of this region is conserved between mouse and human (Figure 4B).

In order to confirm that the *Preli-Rab24* intergenic region functions as a bi-directional promoter, a eukaryotic expression vector was designed using the 371 nt *Sma*I fragment; this fragment contains approx. 105 nt of the *Preli* 5′ UTR and 205 nt of the *Rab24* 5′ UTR, but does not contain the translation initiation codon of either gene (Figure 4A). The *Sma*I fragment was cloned upstream from the coding region of the *neo*^{r} gene in an orientation such that the expression of *neo*^{r} was driven by the *Preli* promoter. A *Mad3* cDNA fragment was cloned on the opposite side of the *Sma*I fragment and the resulting construct was designated *Preli-Rab24-neo*^{r}; the candidate bi-directional promoter fragment would thus direct the expression of both *neo*^{r} (driven by *Preli*) and *Mad3* (driven by *Rab24*) in transfected mammalian cells. *Preli-Rab24-neo*^{r} was stably transfected into C7 3T3 fibroblasts, and a pool of approximately 200 neomycin resistant colonies was generated, thus indicating that the *Preli* promoter was functional in this construct. Ribonuclease protection analysis was used to confirm that the bi-directional promoter directed the expression of both *Mad3* and *neo*^{r} in stably transfected cells (Figure 4C). Transcription was initiated at the correct positions within the *Preli-Rab24* intergenic region (results not shown). It was therefore concluded that the mouse *Preli-Rab24* bi-directional promoter resides within the 371 nt *Sma*I genomic DNA fragment. The identification of this bi-directional promoter further confirms that the genomic *Mad3* 3′ flanking region contains the functional *Preli* and *Rab24* genes. It is notable that *Preli* and *Rab24* share a CpG island that is located between the genes; this arrangement is typically observed within bi-directional loci that are transcribed by RNA polymerase II, and is thought to facilitate the co-ordinate expression of the two genes (reviewed in [15]).

Preli encodes a mitochondrial protein that has an N-terminal mitochondrial-targeting signal

Preli encodes an evolutionarily conserved protein of unknown function that is abundantly expressed in all settings analysed. The

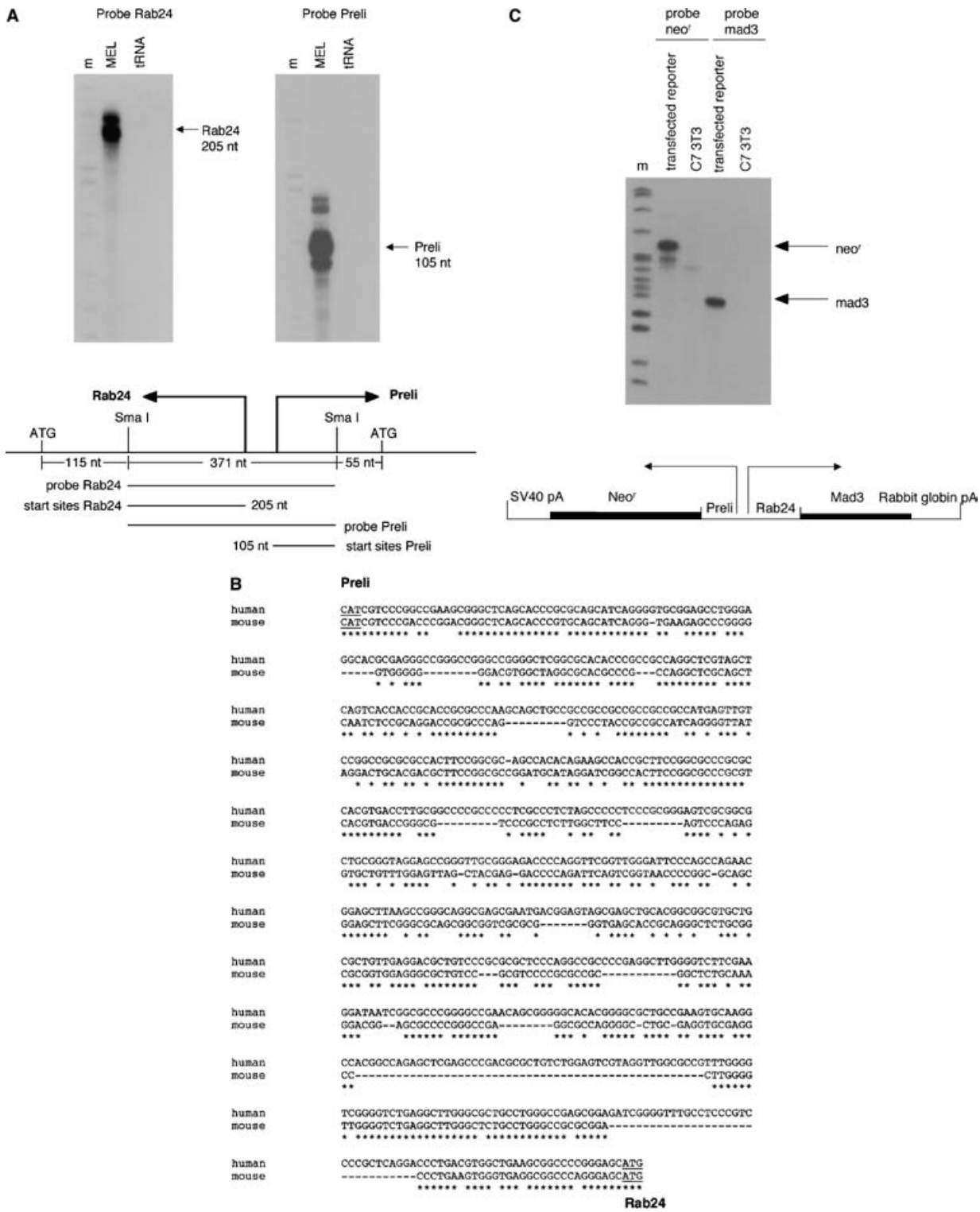

Figure 4 Characterization of the *Preli-Rab24* bi-directional promoter

(**A**) Location of the 5′ ends of mouse *Preli* and *Rab24* transcripts. A 371 nt *Sma*I fragment containing the *Preli-Rab24* intergenic region was subcloned into pSP72 and the resulting construct used to generate riboprobes that detected transcripts in either the *Preli* (**Preli**) or *Rab24* (**Rab24**) directions (see schematic). The 5′ ends of *Preli* and *Rab24* transcripts were mapped by hybridization of riboprobes to samples of MEL RNA (MEL); hybridizations to *E. coli* tRNA (tRNA) were used as negative controls. m, size markers. (**B**) Sequence alignment of the mouse and human *Preli-Rab24* intergenic regions. Nucleotide sequences were aligned using ClustalW. The translation initiation codons of Preli and Rab24 are underlined. Asterisks denote conserved residues. (**C**) Expression of a reporter gene construct containing the *Preli-Rab24* bi-directional promoter in stably transfected C7 3T3 cells. The 371 nt *Sma*I fragment containing the *Preli-Rab24* bi-directional promoter (described in **A**) was placed upstream from the coding region of the *neo*^{r} gene in the orientation such that *Preli* drives the expression of *neo*^{r} (see schematic). A *Mad3* cDNA was inserted downstream from the *Rab24* promoter, and the resulting construct introduced into C7 3T3 cells by stable transfection. RNA from untransfected C7 3T3 cells, and from a pool of approximately 200 transfected neomycin resistant colonies, was analysed by ribonuclease protection mapping using probes designed to detect the *neo*^{r} (probe neo^r) and *Mad3* (probe mad3) reporters. The *Mad3* reporter was detected using a riboprobe derived from the rabbit β-globin 3′ end; this, therefore, did not detect endogenous *Mad3* transcripts. m, size markers.

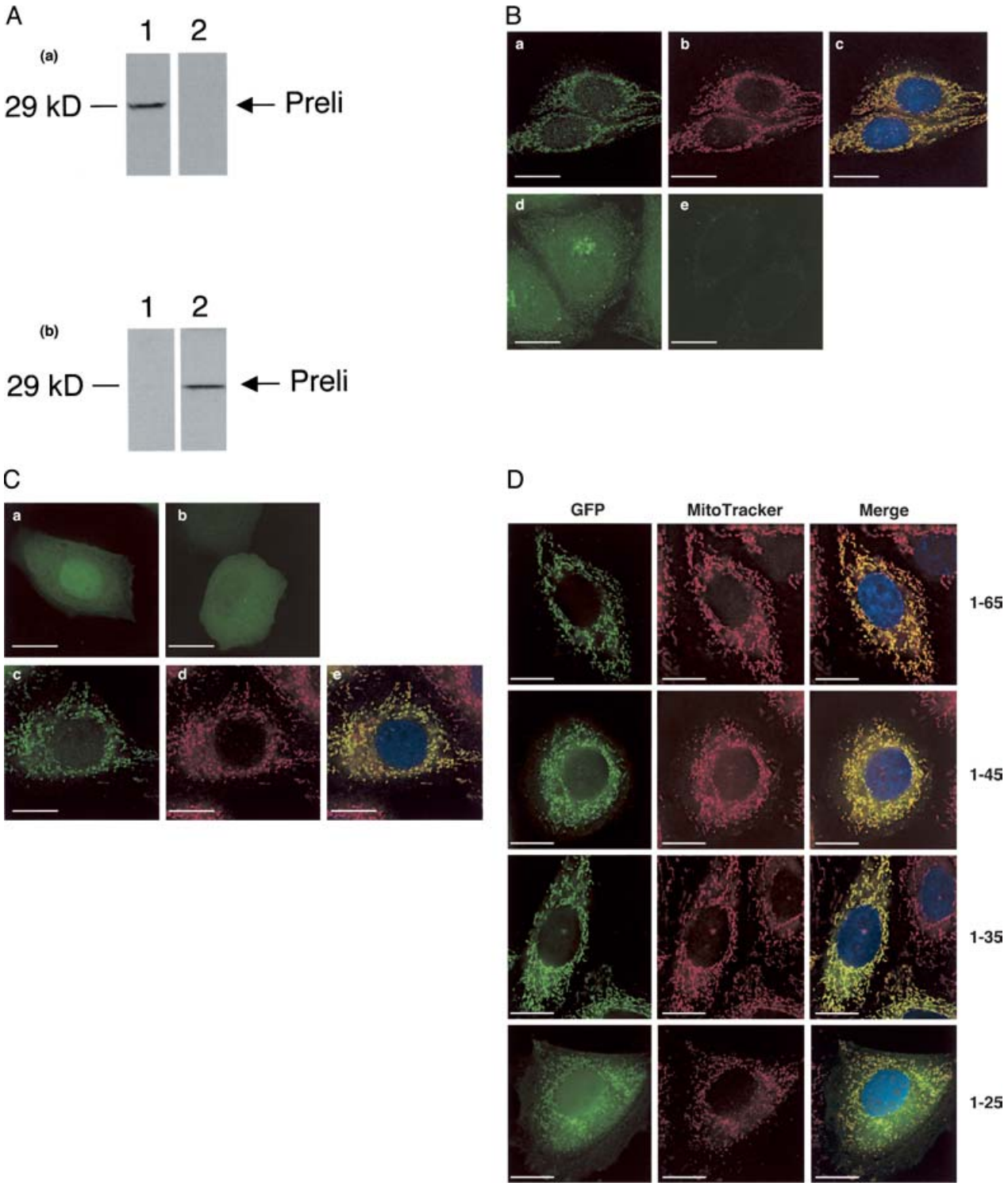

Figure 5 Preli encodes a mitochondrial protein with an N-terminal mitochondrial-targeting signal

(**A**) Western-blot analysis using anti-Preli antiserum. (**a**) Total cell extracts from HeLa cells were analysed by Western blotting using anti-Preli antiserum (1:500 dilution, lane 1) or pre-immune serum (lane 2). (**b**) Cytosolic (lane 1) and mitochondrial (lane 2) extracts from HeLa cells were analysed by Western blotting using anti-Preli antiserum (1:1000 dilution). (**B**) Immunofluorescence microscopy of HeLa cells using anti-Preli antibodies. HeLa cells were immunostained using affinity-purified anti-Preli antibodies and FITC-conjugated secondary antibody. Cells were also treated with MitoTracker Red CMXRos. (**a**) FITC fluorescence, (**b**) MitoTracker fluorescence, (**c**) merge of (**a**) and (**b**), (**d**) staining with pre-immune serum, (**e**) staining with secondary antibody alone. Scale bars represent 10 μm. (**C**) Expression of Preli–GFP fusion proteins in HeLa cells. GFP fluorescence of transiently transfected Hela cells expressing (**a**) GFP alone, (**b**) Preli–GFP fusion with GFP joined to the N-terminus of Preli, (**c**) Preli–GFP fusion with GFP joined to the C-terminus of Preli. Transfected cells in (**c**) were also stained with MitoTracker Red CMXRos and DAPI. Panel (**d**) represents MitoTracker fluorescence, and Panel (**e**) represents a merge of the GFP, MitoTracker and DAPI channels. Scale bars represent 10 μm. (**D**) Expression of Preli–GFP fusions containing N-terminal regions of Preli. HeLa cells were transiently transfected with constructs expressing the indicated amino acid ranges of Preli (1–25, 1–35, 1–45 and 1–65) fused to the N-terminus of GFP. Cells were also stained with MitoTracker Red CMXRos and DAPI. Fluorescence in the GFP and MitoTracker channels is shown. The merge panels represent fluorescence in the GFP, MitoTracker and DAPI channels. Scale bars represent 10 μm.

protein contains two conserved LEA (late embryonic abundant) motifs near the C-terminus (consensus amino acid sequence A/TAEKAK/RETKD). This motif is found in plant proteins that are involved in the protection against stress during embryonic development, and LEA proteins are thus highly expressed in seed embryos and in salt- and temperature-stressed seedlings [16]. Avian Px19 was the first vertebrate LEA protein identified [14], although the role of this domain in animal cells remains to be determined. The most notable sequence domain within PRELI is the recently described PRELI/MSF1p′ motif ([17] pfam 04707) that is found at the N-terminus of the protein. This 170 amino acid domain is predicted to form a globular $\alpha + \beta$ fold comprising four α-helices and six β-strands; it is found in stand-alone form in about 13 proteins from diverse eukaryotes, including *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans*. No function has been assigned to these proteins, and many are hypothetical, although it has been suggested that yeast MSF1p′ is involved in intramitochondrial protein sorting [18]. The domain may also be found in conjunction with a C-terminal GOLD (Golgi dynamics) domain in several proteins that play a role in Golgi dynamics and secretion [17]. Since the role of the stand-alone PRELI/MSF1p′ domain is unknown, it was particularly important to determine the subcellular location of PRELI. In addition, since *Preli* and *Rab24* are co-ordinately transcribed from a bi-directional promoter, it was relevant to determine whether the corresponding proteins share the same subcellular location or are involved in a common physiological process; Rab24 has previously been localised to the endoplasmic reticulum, although its role is unclear [19].

In order to determine initially the subcellular location of Preli, we generated an antibody against the protein by immunisation of sheep with recombinant His-tagged protein. The antiserum was shown, by Western blotting, to detect a 29 kDa protein (Figure 5A, panel a); subcellular fractionation indicated that this protein was localised in a mitochondria-enriched fraction (Figure 5A, panel b). The antiserum was purified by affinity chromatography and used in immunofluorescence microscopy to confirm the subcellular localisation of Preli in HeLa (Figure 5B) and C7 3T3 cells (results not shown). The pattern of fluorescence coincided with that observed upon staining the cells with MitoTracker Red CMXRos (Figure 5B, panels a, b and c), indicating that Preli is found in the mitochondria.

Diverse targeting signals are used to direct newly synthesised proteins to the mitochondria; these can be found at the N- or C-terminus, or internally within the protein (reviewed in [20–22]), and cannot always be reliably predicted using bioinformatics programs. The mitochondrial localisation of Preli is predicted using the program TargetP [23] but not by PSORTII or iPSORT (http://psort.nibb.ac.jp). Interestingly, none of these programs predict the mitochondrial localisation of yeast MSF1P′, although the targeting of a MSF1P′–GFP fusion to the yeast mitochondria was recently reported [24]. In order to confirm the subcellular localisation of Preli, and to define a potential mitochondrial-targeting signal, we constructed vectors that expressed fusion proteins in which GFP was joined to either the N- or C-terminus of Preli. The pattern of GFP fluorescence was studied in transiently-transfected HeLa cells. Fusion of GFP to the C-terminus of Preli resulted in a pattern of fluorescence that coincided with that of MitoTracker Red CMXRos (Figure 5C, panels c, d and e), whereas fusion to the N-terminus resulted in diffuse cytoplasmic staining (Figure 5C, Panel b). This confirmed the mitochondrial localisation of Preli, and indicated that a mitochondrial-targeting signal is most likely located at the N-terminus of the protein. Short segments derived from the N-terminus of Preli were fused to the N-terminus of GFP; the pattern of fluorescence observed in transiently transfected HeLa cells indicated that the N-terminal 35 amino acids contain a mitochondrial-targeting signal (Figure 5D). The GFP-fusion protein containing the N-terminal 25 amino acids of Preli was partially targeted to the mitochondria, although there was also significant diffuse cytoplasmic staining, indicating that the targeting signal was no longer intact within this construct (Figure 5D and results not shown).

Conclusion

During the characterization of the mouse *Mad3* genomic locus, we identified an evolutionarily conserved gene cluster that contains *Mad3*, *Preli* and *Rab24*. The human cluster is located at 5q34–q35, a region that has been implicated in a number of malignancies. Preli is a highly conserved protein that has not been characterized previously. We show here that *Preli* is co-ordinately expressed with *Rab24* from a strong bi-directional promoter, and that it is regulated independently from the S-phase-specific *Mad3* gene that is located at its 3′ end. Preli notably contains a stand-alone 170 amino acid PRELI/MSF1p′ motif at its N-terminus. The function of this domain is unknown, and the subcellular location of higher eukaryotic PRELI/MSF1p′ proteins has not been determined previously. In this study we identify Preli as a novel mitochondrial protein that has an N-terminal mitochondrial-targeting signal. Preli, therefore, most likely has the same subcellular location as yeast MSF1p′; it will be of particular interest to determine the subcellular location of other higher eukaryotic proteins that contain the stand-alone PRELI/MSF1p′ domain, and to define the role of this motif.

It remains to be determined whether the genomic organisation of *Mad3*, *Preli* and *Rab24* reflects their involvement in common physiological processes, although these genes do not have an obvious functional relationship. Aberrations on the long arm of chromosome 5 represent one of the most frequent chromosomal abnormalities in a wide range of haematological disorders including primary myelodysplastic syndrome (MDS), acute myeloid leukaemia (AML), acute lymphoblastic leukaemia, chronic myeloid leukaemia, and therapy-related (secondary) MDS and AML (for example [25,26]). 5q11–q14 and 5q31–q35 have been reported as common breakpoint sites in interstitial deletions and translocations. The latter region contains a number of genes involved in the regulation of cell proliferation and haemopoiesis (for example [27]), although the critical disease genes at many sites of chromosome aberration remain to be identified. The 5q− syndrome is a subtype of MDS-refractory anaemia that is characterized by deletions in 5q as the sole karyotypic abnormality. The critical region of deletion in the 5q− syndrome is thought to reside at 5q31–q32 [28], and several candidate disease genes have been identified [25,29]; the assignment of *MAD3*, *PRELI* and *RAB24* to 5q34–q35 indicates that these genes may importantly be excluded as candidates in the 5q− syndrome. Deletions in 5q34–q35 are also a common chromosomal aberration in testicular and ovarian germ cell tumours [30–32]; it will be of interest to determine whether the *MAD3*/*PRELI*/*RAB24* locus is involved in these malignancies.

The characterization of novel mitochondrial proteins is important for the understanding of both normal cellular function and mitochondrial disease (reviewed in [33]). Mitochondrial defects have been described in degenerative diseases, aging and cancer, with many heritable disorders being attributed to defects in nuclear-encoded mitochondrial proteins. However, only about half of the predicted mitochondrial proteins have thus far been identified, and many such diseases remain uncharacterized; the identification of novel mitochondrial proteins is therefore particularly relevant.

This work was supported by Yorkshire Cancer Research (YCR) and an MRC studentship to E. J. F. J. P. L. and A. F. M. also acknowledge The Wellcome Trust and MRC for financial support. The Leeds Bioimaging Facility is supported by a Wellcome Trust Equipment Award. We thank Dr Vas Ponnambalam, Dr John Walker, Dr Alison Baker and Dr Gareth Howell for advice on the mitochondrial-targeting experiments, and Dr Louise Kime for comments on the manuscript.

REFERENCES

1 Hurlin, P. J., Queva, C., Koskinen, P. J., Steingrimsson, E., Ayer, D. E., Copeland, N. G., Jenkins, N. A. and Eisenman, R. N. (1995) Mad3 and Mad4: novel Max-interacting transcriptional repressors that suppress c-myc-dependent transformation and are expressed during neural and epidermal differentiation. EMBO J. **14**, 5646–5659
2 Eisenman, R. N. (2001) Deconstructing myc. Genes Dev. **15**, 2023–2030
3 Fox, E. J. and Wright, S. C. (2001) S-phase-specific expression of the *Mad3* gene in proliferating and differentiating cells. Biochem. J. **359**, 361–367
4 Pulverer, B., Sommer, A., McArthur, G. A., Eisenman, R. N. and Luscher, B. (2000) Analysis of Myc/Max/Mad network members in adipogenesis: inhibition of the proliferative burst and differentiation by ectopically expressed Mad1. J. Cell. Physiol. **183**, 399–410
5 Queva, C., McArthur, G. A., Iritani, B. M. and Eisenman, R. N. (2001) Targeted deletion of the S-phase-specific Myc antagonist Mad3 sensitizes neuronal and lymphoid cells to radiation-induced apoptosis. Mol. Cell. Biol. **21**, 703–712
6 Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R. (1979) Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell (Cambridge, Mass.) **16**, 777–785
7 Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**, 156–159
8 Wright, S. and Bishop, J. M. (1989) DNA sequences that mediate attenuation of transcription from the mouse protooncogene *myc* . Proc. Natl. Acad. Sci. U.S.A. **86**, 505–509
9 Ioannou, P. and De Jong, P. (1996) Construction of bacterial artificial chromosome libraries using the modified P1 (PAC) system. In Current Protocols in Human Genetics (Dracopoli, N. C., Haines, J. L., Korf, B. R., Moir, D. T., Morton, C. C., Seidman, C. E., Seidman, J. G. and Smith, D. R. eds.). John Wiley and Sons, NY
10 Lu-Kuo, J., Ward, D. C. and Spritz, R. A. (1993) Fluorescence in situ hybridization mapping of 25 markers on distal human chromosome 2q surrounding the human Waardenburg syndrome, type I (WS1) locus (*PAX3* gene). Genomics **16**, 173–179
11 Harlow, E. and Lane, D. (1988) Antibodies: a laboratory manual. Cold Spring Harbor Laboratory Press
12 Fan, S., Wang, J. A., Yuan, R. Q., Rockwell, S., Andres, J., Zlatapolskiy, A., Goldberg, I. D. and Rosen, E. M. (1998) Scatter factor protects epithelial and carcinoma cells against apoptosis induced by DNA-damaging agents. Oncogene **17**, 131–141
13 Guzman-Rojas, L., Sims, J. C., Rangel, R., Guret, C., Sun, Y., Alcocer, J. M. and Martinez-Valdez, H. (2000) PRELI, the human homologue of the avian px19, is expressed by germinal center B lymphocytes. Int. Immunol. **12**, 607–612
14 Niu, S., Antin, P. B. and Morkin, E. (1996) Cloning and sequencing of a developmentally regulated avian mRNA containing the LEA motif found in plant seed proteins. Gene **175**, 187–191
15 Adachi, N. and Lieber, M. R. (2002) Bidirectional gene organization: a common architectural feature of the human genome. Cell (Cambridge, Mass.) **109**, 807–809
16 Dure, L., 3rd. (1993) A repeating 11-mer amino acid motif and plant desiccation. Plant. J. **3**, 363–369
17 Anantharaman, V. and Aravind, L. (2002) The GOLD domain, a novel protein module involved in Golgi function and secretion. Genome Biol. **3**, research0023.1–0023.7
18 Nakai, M., Takada, T. and Endo, T. (1993) Cloning of the YAP19 gene encoding a putative yeast homolog of AP19, the mammalian small chain of the clathrin-assembly proteins. Biochim. Biophys. Acta **1174**, 282–284
19 Olkkonen, V. M., Dupree, P., Killisch, I., Lutcke, A., Zerial, M. and Simons, K. (1993) Molecular cloning and subcellular localization of three GTP-binding proteins of the rab subfamily. J. Cell Sci. **106**, 1249–1261
20 Chacinska, A., Pfanner, N. and Meisinger, C. (2002) How mitochondria import hydrophilic and hydrophobic proteins. Trends Cell Biol. **12**, 299–303
21 Truscott, K. N., Brandner, K. and Pfanner, N. (2003) Mechanisms of protein import into mitochondria. Curr. Biol. **13**, R326–R337
22 Neupert, W. and Brunner, M. (2002) The protein import motor of mitochondria. Nat. Rev. Mol. Cell Biol. **3**, 555–565
23 Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. **300**, 1005–1016
24 Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K. (2003) Global analysis of protein localization in budding yeast. Nature (London) **425**, 686–691
25 Boultwood, J., Fidler, C., Soularue, P., Strickson, A. J., Kostrzewa, M., Jaju, R. J., Cotter, F. E., Fairweather, N., Monaco, A. P., Muller, U. et al. (1997) Novel genes mapping to the critical region of the 5q- syndrome. Genomics **45**, 88–96
26 Le Beau, M. M. (1992) Deletions of chromosome 5 in malignant myeloid disorders. Cancer Surv. **15**, 143–159
27 Le Beau, M. M., Lemons, R. S., Espinosa, R., 3rd, Larson, R. A., Arai, N. and Rowley, J. D. (1989) Interleukin-4 and interleukin-5 map to human chromosome 5 in a region encoding growth factors and receptors and are deleted in myeloid leukemias with a del(5q). Blood **73**, 647–650
28 Jaju, R. J., Boultwood, J., Oliver, F. J., Kostrzewa, M., Fidler, C., Parker, N., McPherson, J. D., Morris, S. W., Muller, U., Wainscoat, J. S. and Kearney, L. (1998) Molecular cytogenetic delineation of the critical deleted region in the 5q- syndrome. Genes Chromosomes Cancer **22**, 251–256
29 Boultwood, J., Fidler, C., Strickson, A. J., Watkins, F., Kostrzewa, M., Jaju, R. J., Muller, U. and Wainscoat, J. S. (2000) Transcription mapping of the 5q- syndrome critical region: cloning of two novel genes and sequencing, expression, and mapping of a further six novel cDNAs. Genomics **66**, 26–34
30 Murty, V. V., Reuter, V. E., Bosl, G. J. and Chaganti, R. S. (1996) Deletion mapping identifies loss of heterozygosity at 5p15.1–15.2, 5q11 and 5q34–5q35 in human male germ cell tumors. Oncogene **12**, 2719–2723
31 Peng, H. Q., Liu, L., Goss, P. E., Bailey, D. and Hogg, D. (1999) Chromosomal deletions occur in restricted regions of 5q in testicular germ cell cancer. Oncogene **18**, 3277–3283
32 Faulkner, S. W. and Friedlander, M. L. (2000) Molecular genetic analysis of malignant ovarian germ cell tumors. Gynecol. Oncol. **77**, 283–288
33 Wallace, D. C. (1999) Mitochondrial diseases in man and mouse. Science (Washington, D.C.) **283**, 1482–1488

Received 1 October 2003/13 November 2003; accepted 26 November 2003
Published as BJ Immediate Publication 26 November 2003, DOI 10.1042/BJ20031504