Nucleic Acids Research, 2005, Vol. 33, No. 4 ***1309–1322***
doi:10.1093/nar/gki269

A family of RS domain proteins with novel subcellular localization and trafficking

Steven J. Kavanagh[1,2], Thomas C. Schulz[1,2], Philippa Davey[1,2], Charles Claudianos[3], Carrie Russell[1] and Peter D. Rathjen[1,2,4,*]

[1]School of Molecular and Biomedical Science and [2]Australian Research Council Special Research Centre in Molecular Genetics, University of Adelaide, Adelaide 5005, Australia, [3]Molecular Genetics and Evolution, Research School of Biological Sciences, Australian National University, ACT 2601, Australia and [4]National Stem Cell Centre, Notting Hill, VIC 3168, Australia

Received September 7, 2004; Revised December 23, 2004; Accepted February 8, 2005 DDBJ/EMBL/GenBank accession no.[+]

ABSTRACT

We report the sequence, conservation and cell biology of a novel protein, Psc1, which is expressed and regulated within the embryonic pluripotent cell population of the mouse. The Psc1 sequence includes an RS domain and an RNA recognition motif (RRM), and a sequential arrangement of protein motifs that has not been demonstrated for other RS domain proteins. This arrangement was conserved in a second mouse protein (BAC34721). The identification of Psc1 and BAC34721 homologues in vertebrates and related proteins, more widely throughout evolution, defines a new family of RS domain proteins termed acidic rich RS (ARRS) domain proteins. Psc1 incorporated into the nuclear speckles, but demonstrated novel aspects of subcellular distribution including localization to speckles proximal to the nuclear periphery and localization to punctate structures in the cytoplasm termed cytospeckles. Integration of Psc1 into cytospeckles was dependent on the RRM. Cytospeckles were dynamic within the cytoplasm and appeared to traffic into the nucleus. These observations suggest a novel role in RNA metabolism for ARRS proteins.

INTRODUCTION

Repeated and/or interspersed arginine/serine dipeptide repeats are a feature of many nuclear proteins with diverse roles including regulation of splicing, transcription, RNA Pol II binding, actin binding, kinase and phosphatase activity, and cell cycle regulation (1). Over 240 RS domain proteins have been identified, the best characterized being the SR and SR-related families, which facilitate spliceosome formation and orchestrate splice site selection (2–5). SR proteins are characterized by an RS domain, one or two RNA recognition motifs (RRMs) and subcellular localization to discrete regions in the nucleus, termed nuclear speckles (6). Nuclear speckles are 20–40 irregularly shaped subnuclear structures (7), which are rich in splicing related factors and recognized by a monoclonal antibody to SC35 (7) that recognizes a range of splicing factors. Localization to nuclear speckles is believed to be diagnostic for proteins involved in mRNA processing (8). These structures do not correlate with regions of active transcription (9,10) and are considered to act as storage sites from which splicing factors are recruited to regulate RNA splicing. Over 140 proteins are known to localize to nuclear speckles including known splicing factors from SR and SR-related families, small nuclear ribonucleoproteins (snRNPs) and other diverse factors such as RNA Pol II (11), the eukaryotic initiation factor eIF4E (12) and the regulators of actin-binding proteins (13).

The RS domain has been shown to mediate protein–protein (14) and protein–RNA interactions (15), to function in nuclear import (16–18) and to play a role in the targeting of proteins such as SC35 and Transformer (19) to nuclear speckles. RS domains from SR proteins, non-SR proteins and synthetic RS domains have also been shown to activate splicing (20). However, the RS domain does not appear to facilitate nuclear import and localization for all RS domain proteins, as SF2/ASF and SRp40 are capable of localization to nuclear speckles in the absence of this domain (21). Where nuclear/cytoplasmic shuttling of RS domain proteins such as SF2/ASF, U2AF and 9G8 has been demonstrated, the RS domain is required, but not sufficient for cytoplasmic localization (22). Nuclear import can be dependent on RS domain phosphorylation and is mediated by SR transportins (TRN-SR) in both mammals

*To whom correspondence should be addressed. Tel: +61 8 8303 5650; Fax: +61 8 8303 4348; Email: peter.rathjen@adelaide.edu.au

[+]AY461716

(17,18) and *Drosophila* (16). The export pathways for SR proteins have not been defined, but can also be influenced by phosphorylation status (23,24). It is now emerging that RS domain phosphorylation also functions in mRNA export (25) and RNA binding specificity (26).

Peri-implantation stem cell 1 (*Psc1*) was identified on the basis of differential expression between mouse embryonic stem (ES) cells and early primitive ectoderm-like (EPL) cells, an *in vitro* equivalent of primitive ectoderm (27). In the early embryo, *Psc1* expression is restricted to the inner cell mass (ICM) of the blastocyst and down regulated on the formation of the primitive ectoderm between 5.0 and 5.75 days post coitum. In this paper, we describe the Psc1 sequence, identify related proteins in vertebrates and invertebrates that define a new class of RS domain proteins termed acidic rich RS (ARRS) domain proteins and demonstrate a novel subcellular distribution that includes localization to punctate sites within the nucleus (nuclear speckles) and cytoplasm (cytospeckles), and the transport between the two compartments. We show by mutational analyses that the RRM is critical for the integration of Psc1 into cytospeckles, the RS domain functions in nuclear import, and both the RS domain and the RRM are necessary for subnuclear localization. A conserved C-terminal domain associates with microtubules and may be required for trafficking of cytospeckles into the nucleus. Taken together these observations suggest a novel role for this new family of RS domain proteins in RNA metabolism.

MATERIALS AND METHODS

cDNA isolation, sequencing and analysis

A λ ZAP II library (Clontech Inc.), prepared from D3 ES cell RNA (28), was screened using a 381 bp *Psc1* cDNA fragment (nt 1660–2040) identified by differential display PCR (27). A third round positive plaque containing *Psc1* nt 901–3512 was zapped into pBluescript SK (clone 8.1; Stratagene) and sequenced. RACE PCR to isolate the 5′ end of the transcript was carried out by ampliFINDER RACE/PCR (Clontech) according to the manufacturer's instructions. D3 ES cell RNA was reverse transcribed using primer 1736 (5′-TTTACTTTGATTGTTGTTCC-3′) and amplified using the 5′ anchor primer and primer 1064 (5′-TAGAATTCGGCA-GAGCAACTTCATCAACAACAACTA-3′). First round RACE/PCR product was cloned into pBluescript KS and sequenced to generate *Psc1* nt 478–1088 bp. For second round RACE/PCR, D3 ES cell RNA was reverse transcribed using primer 1275 (5′-TGGGAAGCACAACAGAAGGT-3′) and amplified using the 5′ anchor primer and primer 582 (5′-TAGAATTCGTACCGCTCATAGTCTCTCCAC-3′), cloned into pBlusecript KS and sequenced to generate *Psc1* nt 1–603. The open reading frame (ORF) was identified by the presence of an in-frame ATG start codon preceded by two in-frame stop codons. All plasmids were sequenced in both directions. Protein homologies were identified with the aid of SIM Software (http://kr.expasy.org/tools/sim-prot.html) from the Expert Protein Analysis System (ExPASy) of the Swiss Institute of Bioinformatics (http://kr.expasy.org) and BLASTP server software (www.ncbi.nlm.nih.gov). DNA to protein translations used the ExPASy 'Translate Tool' (http://kr.expasy.org/tools/dna.html). DNA sequence homologies were identified through BLASTN and 'ALIGN' software from the GENESTREAM network server (http://www2.igh.cnrs.fr/bin/align-guess.cgi). Default parameters were applied for all server applications. Sequence analysis of the KIAA1311 cDNA revealed a probable frame shift which did not allow the identification of the start codon. The frame shift, a single nucleotide insertion at position 566, was identified by comparison with the *Psc1* cDNA and corrected to identify the probable start codon. The complete ORF of BAC34721 was derived from BC049360, which spans the BAC34721 sequence. Phylogenic analysis was derived through multiple protein alignment using 'CLUSTAL W' and the neighbour-joining method with standard distances and mean character differences (29).

Plasmid vectors

Psc1-HA: Three copies of the haemagglutinin epitope tag followed by a stop codon were cloned 3′ of Psc1 nt 18–3171, encompassing the full-length *Psc1* ORF (nt 157–3171) in pXMT2 (30). *GFP–Psc1*: Psc1 nt 103–3512, including full-length *Psc1* ORF (nt 157–3171) was cloned in-frame 3′ of green fluorescent protein (GFP) into pEGFP-C2 (Clontech). *GFP–Psc1ΔRS*: Constructed using Quikchange site-directed mutagenesis (Stratagene) on GFP–Psc1 to delete *Psc1* nt 577–876. *GFP–RS*: The RS domain of Psc1 was generated by PCR amplification of *Psc1* nt 577–876 and cloned in-frame 3′ of GFP into pEGFP-C2. *GFP–Psc1ΔRRM*: Constructed using Quikchange site-directed mutagenesis on GFP–Psc1 to delete *Psc1* nt 1738–2022. *GFP–RRM*: The RRM of Psc1 was generated by PCR amplification of *Psc1* nt 1579–2211 and cloned in-frame 3′ of GFP into pEGFP-C2. *GFP–Psc1ΔCD*: Constructed using Quikchange site-directed mutagenesis on GFP–Psc1 to delete *Psc1* nt 2368–2835. *GFP–CD*: The C domain and adjacent RG repeat sequence of Psc1 was generated by PCR amplification of *Psc1* nt 2347–3039 and cloned in-frame 3′ of GFP into pEGFP-C2. *GFP–SC35*: The *SC35* ORF was amplified by PCR on pCGSC35 (gift from Dr A. Krainer, Cold Spring Harbor Laboratory, NY) and cloned in-frame 3′ of GFP into pEGFP-C2. *GFP–SF2/ASF*: The *SF2/ASF* ORF was amplified by PCR on pCG-SF2/ASF (gift from Dr A. Krainer, Cold Spring Harbor Laboratory, NY) and cloned in frame into pEGFP-C2. *His–Psc1–FLAG*: The *Psc1* ORF (nt 157–3171) was PCR amplified using primers 5′Psc1His (5′-AGAATTCCACCATGCATCATCATCAT-CATCATCATCATCTCATAGAAGATGTGGATGCCC-3′) and 3′Psc1FLAG (5′-TCACTTGTCATCGTCGTCCTTG-TAGTCTCTTCGCCACGAACGAGACTC-3′, which contained sequences encoding eight 5′ histidine repeats and a 3′ FLAG sequence, respectively and was cloned into EcoR1 digested pcDNA3.1 (Invitrogen). *pGEX2T-RRM*: The Psc1 RRM was generated by PCR amplification of *Psc1* nt 1579–2211 and was cloned in-frame 3′ of GST into pGEX2T (Pharmacia). *pGEX2T-Ab*: *Psc1* nt 2059–2295 were amplified by PCR and cloned in-frame 3′ of GST in pGEX2T. All plasmids were sequenced by automated DNA sequencing (PE Biosystems). Details of plasmid construction are available on request.

Cell culture, transfection and cell counts

COS-1 cells were maintained in DMEM (Gibco, 12430-054)/10% fetal calf serum (JRH Biosciences). ES and EPL cultures were maintained as described previously (31). COS-1 cells

were seeded on to cover and transfected at 50–60% confluence using FuGene^{TM}6 (Roche Molecular Biochemicals) according to the manufacturer's instructions. Cells were analysed 10–12 h post transfection. Cell counts for the percentage of cells expressing nuclear, cytoplasmic or nuclear and cytoplasmic protein were derived from scoring 300 transfected cells from each of the three separate GFP–Psc1 transfection assays.

Indirect immunofluorescence and microscopy

Cells on coverslips were fixed with methanol and rehydrated in phosphate-buffered saline (PBS). All primary antibodies were applied for 1 h at room temperature in PBS with 0.1% Triton X-100 (PBT) containing 3% BSA (Sigma). Affinity purified rat anti-haemagglutinin antibody (Boehringer) was used at a dilution of 1:1000. Monoclonal mouse anti-SC35 (gift from Prof. T. Maniatis, Harvard University) and purified polyclonal rabbit anti-Psc1 were used at a dilution of 1:500. Cells were washed 3 × 5 min followed by one wash of 30 min in PBT between antibody applications. Secondary antibodies: sheep anti-rabbit (IgG) TRITC conjugate (Sigma), goat anti-mouse (IgG) TRITC conjugate (Sigma) and goat anti-rat IgG fluorescein conjugate (Sigma) were applied at a dilution of 1:1000 in PBT containing 3% BSA for 30 min at room temperature. For double labelling of Psc1-HA and SC35, goat anti-rat IgG fluorescein conjugate and goat anti-mouse (IgG) TRITC conjugate were adsorbed in 1% mouse serum and 0.2% rat serum, respectively for 1 h prior to use and applied sequentially with 3 × 5 min washes followed by a 30 min wash in PBT between applications. This wash regime was repeated following the application of secondary antibody and the coverslips were mounted for analysis. Hoechst 33258 (200 µl) at 5 µg/ml was applied to cells for 2 min prior to the final wash. For real time imaging, cells were stained with Hoechst 33342, a vital nuclear stain (32,33), at 100 ng/ml for 5 min and the medium was replaced immediately prior to analysis and maintained at 37°C using a stage mounted warming plate. Conventional images were viewed on a Zeiss Axioplan microscope with 100× lens in oil immersion and captured on Olympus UTV1X-2/CMAD3 coolsnap fx camera to V++ 4.0 (Digital Optics) or Photoshop 6.0 (Adobe). All confocal images were captured using Bio-Rad MRC-1000UV Confocal Laser Scanning System with a Nikon Diaphot 300 inverted microscope equipped with 60× Water/NA1.4 or 40× Water/NA1.25 (real time) lens and imaged using Photoshop 6.0.

RNA binding assay

^{32}P-labelled adenovirus major late transcript RNA was generated by an *in vitro* transcription reaction (Roche) using 1 µg linearized pBSAD1 cDNA template (gift from Dr M. Little, University of Queensland) and 100 µCi [^{32}P]UTP (Perkin Elmer). The reaction was treated with 10 U DNase 1 and purified through a Sephadex G-50 column (Amersham Biosciences). pGEX2T and pGEX2T-RRM were transformed into BL21 *Escherichia coli* for recombinant protein production. GST-containing proteins were purified using glutathione–Sepharose 4B (Pharmacia), as described by the manufacturer. Approximately 2 mg of purified GST–RRM and 6 mg of GST were dialysed overnight against 20 mM HEPES, pH 8.0, 100 mM KCl, 5% glycerol (v/v), 0.2 mM EDTA and 1 mM dithiothreitol, then the buffer was renewed and exchange continued for further 4 h. An aliquot of 1 µg of each of GST and GST–RRM were used for the RNA binding assay as described by Krainer *et al*. (34) using ~150 fmol of *in vitro* transcribed [^{32}P]UTP labelled RNA. Cold competitor ES cell RNA (0 ng, 10 ng, 100 ng or 1 µg) was added as indicated. Samples were fractionated on a 12.5% SDS–PAGE gel and visualized by autoradiography.

Production of affinity purified polyclonal Psc1 antibodies

PGEX2T-Ab was transformed into BL21 *E.coli* for recombinant protein production. Cells were lysed by sonication (4 × 30 s) and GST-Psc1Ab was purified using glutathione–Sepharose 4B (Pharmacia), according to the manufacturer's instructions. Two male Semi lop rabbits were injected subcutaneously with 100 µg of purified GST-Psc1Ab in 1 ml Freund's complete adjuvant on day 0, and again on days 21 and 42. A final injection of 50 µg of purified GST-Psc1Ab in 1 ml of Freund's incomplete adjuvant was administered on day 70, and the serum was collected after 10 days. Preimmune samples were taken prior to immunization. Psc1 antibodies were purified from 20 ml serum by incubating overnight at 4°C with agitation with 2 ml glutathione–agarose (Sigma) coupled to GST (gift from Dr G. Booker, Adelaide University) according to the manufacturer's instructions. The next day, the slurry was centrifuged and the supernatant was removed and gravity fed (0.4 ml/min) three times through a 2 ml Affiprep 10 (BioRad) GST-Psc1 column constructed by cyanogen bromide coupling of purified GST-Psc1Ab protein to Affiprep 10 according to the manufacturer's instructions.

Western blot assay

Whole cell extracts were prepared from 10^7 COS-1 cells by NP-40 detergent lysis. Pelleted samples were resuspended in 50 µl of SDS load buffer. *In vitro* transcription translation from plasmid DNA was carried out using the TNT Quick coupled transcription/translation system (Promega) in 50 µl reaction volumes. Proteins were fractionated by SDS–PAGE and transferred to nitrocellulose (Protran). Primary antibodies diluted 1:1000 in PBT were applied to the membrane and incubated for 1 h at room temperature followed by incubation with HRP-conjugated secondary antibodies diluted 1:2000 and applied for 1 h. Blots were developed in ECL (SuperSignal Substrates, Pearce) according to the manufacturer's instructions.

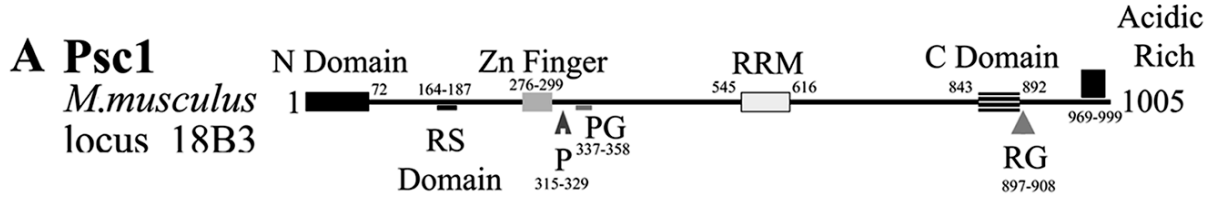
RESULTS

Psc1 cDNA isolation and characterization

Partial *Psc1* cDNA clones were isolated from a D3 ES cell library and the 5′ end of the transcript was cloned by 5′ RACE PCR using D3 ES cell RNA to generate a *Psc1* cDNA of 3521 bp with an incomplete 3′-UTR (accession no. AY461716), consistent with the longest transcript size of 5.5 kb identified by northern analysis (27). BLAST searches revealed 93% identity with KIAA1311 cDNA, an EST isolated from human brain tissue with no described function (35).

The ORF of 1005 amino acids was confirmed by the presence of in-frame stop codons within the 5′-UTR. BLASTP database analysis was used to identify conserved domains within Psc1 (Figure 1A). The N domain (amino acids 1–78),

shared 30% identity with the N-terminal region of the 77 kDa human protein Hprp3p (Figure 1A, panel i), which binds via its C-terminus to the prespliceosomal U4/U6 snRNP complex. This region of Hprp3p has no known role in subcellular localization and is proposed to be involved in protein–protein interactions (36). A short RS-rich sequence located between residues 164 and 187 (Figure 1D) was identified as containing an RS domain on the basis of four consecutive SR dipeptide repeats. RS domains are inconsistent in size with these motifs defined in proteins with as few as two consecutive SR dipeptide repeats (37). Psc1 amino acids 276–299 and 545–616 were identified as containing a $C(X)_8C(X)_5C(X)_3H$ Zn finger motif (Figure 1A, panel ii) and a predicted RRM (Figure 1A, panel iii), with 43 and 31% identity to the respective consensus sequences derived from the NCBI conserved domain database (38). The organization of Psc1 was therefore different from, and more complex than the common arrangement in SR proteins of one or two RRMs followed by a C-terminal RS domain (2,39).



i **N Domain**

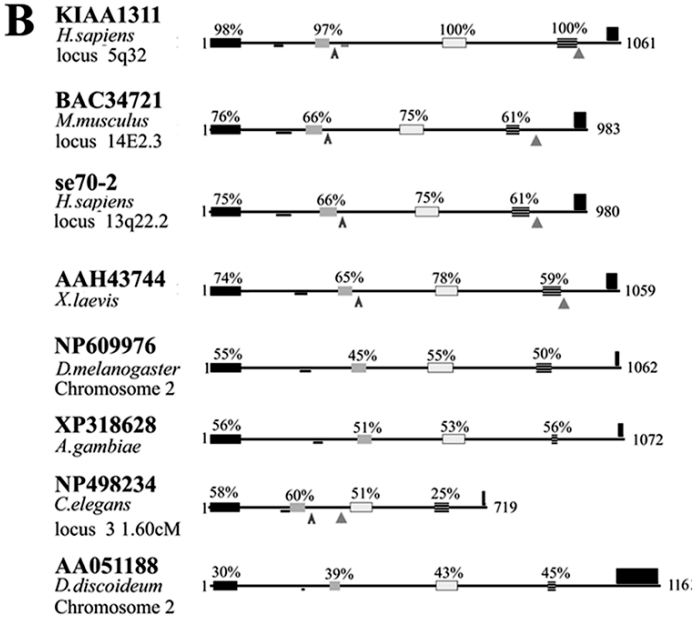
```
Psc1    1 MLI--EDVDALKSWLAKLLEPICDADPSALANYVVALVKKDKPEKELKAFCADQLDVFLQKETSGFVDKLFESL 72
Hprp3p  1 MALSKRELDELKPWIEKTVKRVLGFSEPTVVTAALN-VGKGMDKK--KA--ADHLKPFLDDSTLRFVDKLFEAV 70
```

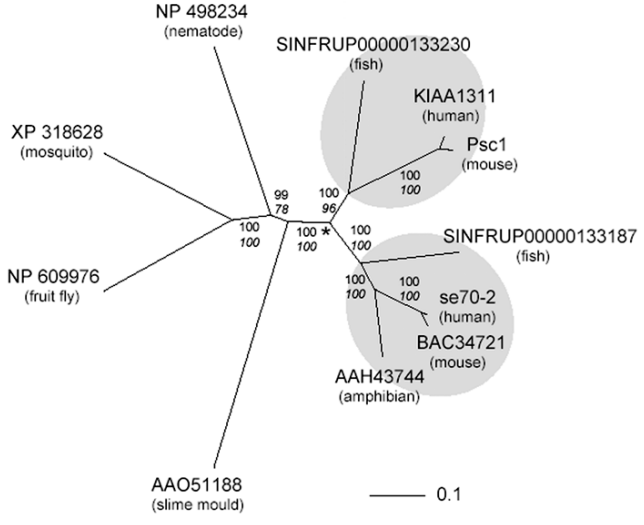
ii **Zn Finger**

```
Psc1      276 RRCRDYDERGFCVLGDLCQFDHG 299
Consensus     ELCRFFLRTGYCKYGDRCKFAHG
```

iii **RRM**

```
Psc1   545 KLEVKKIPQELNNITKLNEHFSKFGTIVNIQVAFKGDPEA---ALIQYLTNEEARKAISS-TEAVLNNRFIRVLW 616
Consensus  TLFVGNLPPDTTE-EDLRELFSKFGKVVSVRIVRDKDTKSKGFAFVEFEDEEDAEKAIEALNGKELGGRPLRVKR
```





D

RS Domain

```
Psc1      RSKSRSKSRGLSRSRSRSRGRS
KIAA1311  RSKSRSKSRGLSRSRSRSRGRS
BAC34721  RSRSYSRSRSRSWSKERLRDRDRDRSRTRSRSRTRSR
se70-2    RSRSYSRSRSRSWSKERLRERDRDRSRTRSRSRTRSR
AAH43744  RSRSYSRSRSRSWSKERQR..DRDR....SRSRTRSR
NP609976  SRRRRASLRSRSRSRSRSRSNERAFRRSRSR
XP318628  RSRRTQNRSRSRSRSRSRSFE
NP498234  SRSGRTSSSDRDRHSKRRSRSRSRSSSGSRTPPYKSR
AA051188  RKRSRSRSNSP
```

Proline Rich

```
Psc1      PPPPPGLPPPPPP
KIAA1311  PPPPPGLPPPPPP
BAC34721  PPPPGLPPPPP
se70-2    PPPPGLPPPPP
AAH43744  PPPPGLPPPP
NP609976  DNE
XP318628  DNE
NP498234  PPPGYNPLNPPPP
AA051188  DNE
```

Acidic Rich Region

```
Psc1      EEEEVKEEETETSDLFLH...DDDDEDEDEYESRSWRR
KIAA1311  EEEEVKEEETETSDLFLP...DDDDEDEDEYESRSWRR
BAC34721  DEEEFQEE....SLVDDSLL.QDDDEEEEDNESRSWRR
se70-2    DEEEFQEE....SLVDDSLL.QDDDEEEEDNESRSWRR
AAH43744  EEEEFHED....SIVDDSLLQDDDEEEEDDNESRSWRR
NP609976  ......ESESLLGSDTLPELRLEDEEEDEESEDRSWR
XP318628  .....ETLTETDELIEMPIMVEEEEEEDDIEPERSWRR
NP498234  NLESDEEDDLLND
AA051188  DEEEQDYEEIPTETI*NDDDDDDDEDDYDEDDKPSWKH
```

Psc1 contained two additional repeat sequences of unknown significance, eight consecutive glycine/arginine dipeptide repeats (positions 897–908) and 11 consecutive proline–glycine dipeptide repeats (positions 337–358). Psc1 also contained a region rich in proline (40%) between amino acids 315 and 329, and an acidic rich region containing 70% aspartic acid and glutamic acid residues located towards the C-terminus (amino acids 969–999), (Figure 1D). The arrangement of all sequence elements is depicted in Figure 1A.

Evolutionary conservation of Psc1 and ARRS family members

The translated Psc1 protein was used in BLAST database analyses. We identified a number of highly conserved homologues including a second mouse protein, BAC34721 (40) (Figure 1B). Phylogenetic analyses indicated two human genes KIAA1311 and se70-2 (accession no. AAH41655) that were orthologues of Psc1 and BAC34721, respectively. The two proteins are most likely the result of a putative gene duplication event having occurred in a common vertebrate ancestor. We consistently identify two orthologous genes from completed fish, chicken, mouse and human genome projects (Figure 1C). The vertebrate genes share a common ancestry with a single copy gene present in insects, nematodes and slime-moulds. Together these proteins distinguish a highly conserved gene family that encode the ARRS domain containing proteins.

Other than structurally inferred function there is little known concerning the role of ARRS proteins. Serological screening identified se70-2 as a tumour antigen, the transcript of which is upregulated in cutaneous T-cell lymphomas and leukemia cell lines (41). The schematic representation of a selection of ARRS proteins encoded by the genes found on the NCBI database is shown in Figure 1B. Human KIAA1311, mouse BAC34721, amphibian (*Xenopus laevis*, AAH43744), fruit fly (*Drosophila melanogaster*, NP_609976), mosquito (*Anopheles gambiae*, XP_318628), nematode worm (*Caenorhabditis elegans*, NP_498234) and slime-mould (*Dictyostelium discoideum*, AAO51188) proteins are predicted from respective genome sequencing projects and have unknown function.

Comparative sequence analysis of nine ARRS proteins (Figure 1A and B) highlighted two conserved amino acid motifs, $P(X)_3N(X)_7HF(X)_2FG(X)_3N$ and $A(X)_2A(X)_2S(X)_5$ $NNRFI(X)_3W$ (boxed in Figure 1A, panel iii), that are unique to the RRMs of these ARRS proteins. Other conserved sequences corresponded to Psc1 amino acids 843–892 (the C domain), and a terminal RSWR sequence in all nine proteins except NP_498234 and AAO51188, located at the C-terminus adjacent to the acidic rich region (Figure 1D). The proline-rich region could not be identified in NP_609976, XP_318628 or AAO51188 but was present in all other proteins (Figure 1D) and the RG repeat sequence was confined to the vertebrate members of the family (Figure 1B). The PG-rich region was unique to the Psc1 and KIAA1311 ARRS clade (Figure 1C).

Psc1 exhibits novel nuclear and cytoplasmic localization

Subcellular localization of Psc1 was analysed by over expression of an epitope-tagged protein in fixed and viable COS-1 cells (Figure 2). Psc1-HA (Figure 2A) and GFP-Psc1 (data not shown), both colocalized with nuclear speckles (identified as anti-SC35 localization) and were excluded from the nucleoli (right arrow in Figure 2A, panels iv and v). No instance of Psc1 exclusion from nuclear speckles was observed, however, additional punctate regions of Psc1 localization in the nucleus were observed that contained Psc1, but were not stained with the anti-SC35 antibody. These were often smaller than the nuclear speckles, did not share the same irregular morphology and were frequently located adjacent to the nuclear membrane (left arrow in Figure 2A, panels iv and v, and arrow in Figure 2D).

Punctate foci containing Psc1 were also detected within the cytoplasm (Figure 2B). Three distinct subcellular localization profiles were observed in GFP-Psc1 expressing cells (Figure 2C): nuclear only in ~50% of cells (Figure 2A), cytoplasmic only in <1% of cells (Figure 2B, panel i) or nuclear and cytoplasmic in ~49% of cells (Figure 2B, panel ii). Cytospeckles were observed in up to 50% of transfected cells, varied in size from <0.1 μm to ~1 μm in diameter and numbered from 50 to 1000. No correlation was observed between the number of cytoplasmic and nuclear speckles. When

Figure 1. Arrangement of conserved elements in Psc1 and ARRS proteins and homologues. (**A**) Diagrammatic representation of conserved protein motifs and domains within the 1005 amino acid sequence of Psc1. N domain; RS domain, arginine/serine dipeptide repeat; Zn finger, $C(X)_8C(X)_5C(X)_3H$ zinc finger motif; P, proline-rich region; PG, proline/glycine repeats; RRM, RNA binding motif; C domain, shared region of homology between ARRS proteins and homologues; RG, arginine/glycine repeats; acidic rich, C-terminal aspartate/glutamate-rich region; (i–iii). Homologies between the Psc1 N domain (i), Zn finger (ii) and RRM (iii) and proteins of known function or consensus derived from the NCBI conserved domain database of known Zn finger or RNA binding domains. Identical residues shown as bold. The RRM motifs $P(X)_3N(X)_7HF(X)_2FG(X)_3N$ and $A(X)_2A(X)_2S(X)_5NNRFI(X)_3W$ that are unique to ARRS proteins and homologues are boxed (iii). (**B**) Alignment of representative ARRS proteins and homologues with Psc1. Representative proteins from human, mouse, amphibian, fruit fly, mosquito, nematode worm and slime-mould. Conservation extends to all ARRS proteins including fish, chicken and rat (data not shown). Elements are drawn to scale and the positions of motifs described in (A) are indicated. Percentages indicate degree of amino acid identity to the equivalent domain in Psc1. (**C**) Unrooted distance neighbour-joining tree showing a phylogeny of ARRS proteins. Sequences for predicted ARRS proteins from fish (accession codes SINFRUP00000133230; SINFRUP00000133187), chicken (accession codes ENSGALG00000007516; ENSGALG00000016910,) and rat (accession code ENSRNOG00000009836) were taken from the ENSEMBL database (http://www.w3.org/1999/xlink). Sequences for mouse (accession code BAC34721), human (accession codes KIAA1311 and AAH41655), amphibian (*Xenopus laevis*; AAH43744), fruit fly (*D.melanogaster*, NP_609976), mosquito *(A.gambiae*, XP_318628), nematode worm (*C.elegans*, NP_498234) and slime-mould (*D.discoideum*, AAO51188), were taken from the NCBI database (http://www.w3.org/1999/xlink). Sequences were aligned using 'CLUSTAL W' (29) with the BLOSUM 62 scoring matrix, with gap opening and gap extension penalties of 10.0 and 0.1, respectively, followed by some minor manual corrections to conform to known structural features. The tree was constructed with PAUP* (76) using standard distances and mean character differences. High confidence was confirmed via congruent tree topology using Parsimony treatment (PAUP*, data not shown) and high resampling statistics indicated at nodes (1000 bootstrap replications represented as percentage values; distance above and Parsimony below). Ellipses (shaded) define two proteins, Psc1 and BAC34721 clades, the result of a putative gene duplication (*) that occurs in the vertebrate lineage. Shown are clades composed of orthologous proteins from vertebrate and invertebrate organisms. Scale bar indicates a distance of 0.1 amino acid substitutions per position in the sequence. (**D**) Sequence comparison of RS domain, proline-rich and acidic rich elements in ARRS proteins and homologues. The asterisk in the sequence of AA051188 represents an intervening 11 amino acids not shown. All other sequences are contiguous with periods used to align areas of similarity. DNE; does not exist.

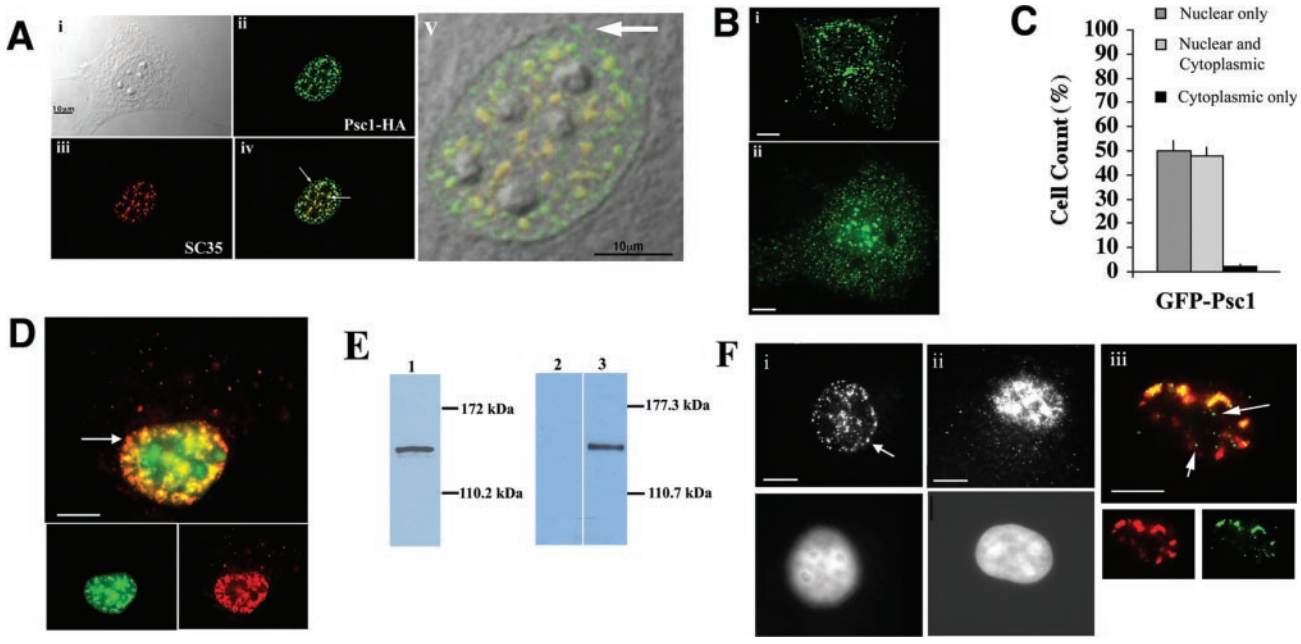


Figure 2. Subcellular localization of Psc1. (**A**) Nuclear localization: Psc1-HA transfected COS-1 cell (i) visualized with anti-HA FITC antibody (ii) and anti-SC35 antibody (iii). (iv), Merged image of (ii) and (iii). Left arrow indicates Psc1 nuclear speckles that do not contain SC35 and tend to be associated with the nuclear periphery, right arrow indicates Psc1 and SC35 colocalization in yellow. (v), merged image of panels (i) and (iv). The arrow shows the proximity to the nuclear membrane of Psc1-containing nuclear speckles that do not contain SC35. (**B**) Cytoplasmic localization: (i), Psc1-HA transfected COS-1 cell visualized with anti-HA FITC antibody showing cytoplasmic localization in the absence of nuclear localization. (ii), GFP–Psc1 transfected COS-1 cell showing nuclear and cytoplasmic localization. (**C**) Subcellular distribution of GFP–Psc1 within a population of GFP–Psc1 transfected COS-1 cells. Error bars indicate standard deviation. (**D**) COS-1 cells were co-transfected with Psc1-HA and GFP–SF2/ASF and visualized with anti-HA TRITC (lower right) or by direct fluorescence (lower left). The top panel shows merged images of the lower panels. (**E**) Specificity of purified polyclonal anti-Psc1 antibody raised against Psc1 amino acids 635–713. Western blot analysis of: lane 1, untransfected COS-1 cells (10^7) were lysed and the pellet fraction probed with purified anti-Psc1 antibody; lane 2, 5 μl of 50 μl unprimed rabbit reticulocyte lysate transcription/translation reaction probed with anti-FLAG monoclonal antibody; lane 3, 5 μl of 50 μl His-*Psc1*-Flag primed rabbit reticulocyte lysate transcription/translation reaction probed with anti-FLAG monoclonal antibody. (**F**) Endogenous Psc1 in COS-1 cells. Untransfected COS-1 cells were visualized with anti-Psc1 antibody. Panels show endogenous Psc1 in nuclear speckles in the absence (i) or presence (ii) of cytoplasmic speckles. (iii) Merged image of COS-1 cell visualized with anti-SC35 (red, lower left panel), and anti-Psc1 (green, lower right panel). Arrows indicate endogenous Psc1-containing nuclear speckles that do not contain SC35. Nuclei were stained with 5 μg/ml Hoechst 33258 (lower panels i and ii). Size bars represent 10 μm. Fluorescence images in (A) were obtained using confocal microscopy. All other images were obtained by conventional microscopy.

assayed in the same manner GFP–SF2/ASF, which is known to shuttle between the nucleus and the cytoplasm (22), localized to nuclear speckles as described previously (34), with no evidence of punctate cytoplasmic localization (Figure 2D). Even in cells co-transfected with SC35 + Psc1-HA (data not shown) or SF2/ASF + Psc1-HA (Figure 2D), the former proteins remained confined to nuclear speckles while Psc1-HA localized to nuclear speckles, additional punctate regions in the nucleus (arrow, Figure 2D) and cytospeckles within the same transfected cell.

The existence of Psc1-containing cytospeckles was validated by immunofluorescent detection of endogenous Psc1 in untransfected COS-1 cells using an affinity-purified polyclonal antibody (Figure 2E) directed against amino acids 635–713 of Psc1. This region of Psc1 shares no significant similarity with BAC34721 and would, therefore, not be expected to detect this protein or its homologues. However, the Psc1 human homologue, KIAA1311 shares 83% identity across this region, suggesting that the mammalian homologues of Psc1 may be recognized by the anti-Psc1 antibody. Endogenous protein was detected in speckles either in the nucleus only (Figure 2F, panel i), or in the nucleus and cytoplasm of ~8% of cells (Figure 2F, panel ii). Within the nucleus, both $SC35^+$ and $SC35^-$ Psc1-containing speckles could be identified (Figure 2F, panel iii), and speckles adjacent to the nuclear membrane were clearly evident (arrow, Figure 2F, panel i). Consistent with the results obtained for subcellular localization of other SR proteins using these assay conditions (10,21), the distribution of endogenous Psc1 protein was, therefore, reminiscent of over expressed Psc1-HA and GFP-Psc1 in transfected COS-1 cells in both the nuclear and the cytoplasmic compartments. The low signal for endogenous protein compared with over expressed Psc1 may reflect protein levels within COS-1 cells or reduced antibody affinity for the monkey protein, consistent with the requirement for large numbers of cells for detection of endogenous protein by western blot (Figure 2E). A similar distribution of endogenous Psc1 was observed in mouse EPL cells (31) (data not shown).

The subcellular distribution of Psc1, therefore, differed from that of the other RS domain proteins such as the splicing factors SF2/ASF and SC35 in two respects; first, it was assembled into additional speckles in the nucleus that were often peripheral to the nuclear membrane and did not contain the SF2/ASF or SC35, and second, it localized throughout the cytoplasm in cytospeckles.

Psc1-containing cytospeckles are motile

The size of nuclear speckles and cytospeckles allows for real time observation of the subcellular motility of Psc1 within

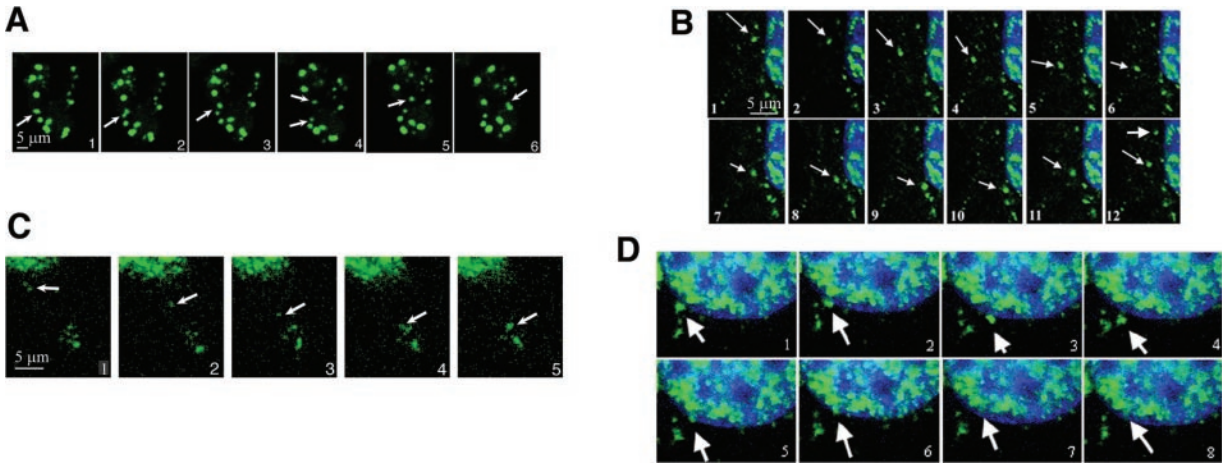


Figure 3. Real time analysis of GFP–Psc1 motility in COS-1 cells. GFP–Psc1 transfected COS-1 cells were analysed 10 h post transfection by confocal microscopy. Images were captured at 15–30s intervals. Cells were stained with Hoechst 33342 to identify the nucleus and maintained in fresh growth medium at 37°C for the timecourse of the experiment. (**A**) Motility of Psc1 nuclear speckles. The arrow indicates speckle motility and budding across the nucleus. The lower arrow in panel 4 indicates a speckle originated from a budding event. Panels 1–6 were captured at 0, 30, 90, 270, 420 and 540s, respectively. (B–D) Motility of Psc1 cytospeckles. (**B**) Random and stationary cytospeckle motility. The cytospeckle indicated by the arrow in panels 1–12 shows a change of direction in panel 10. Upper arrow in panel 12 shows a stationary cytospeckle. Panels 1–12 were captured at 0, 15, 30, 45, 75, 90, 105, 120, 150, 165, 210 and 225, respectively. (**C**) Directed motility and fusion of cytospeckles. The cytospeckle indicated by the arrow moves directionally away from the nucleus and fuses with distant cytoplasmic speckles. Panels 1–5 were captured at 0, 630, 780, 900 and 930s, respectively. (**D**) Motility of Psc1 cytospeckles, (arrow, panels 1–8). Panel 4 shows the speckle move slightly away from the nucleus before apparent nuclear translocation in panels 5–8. Panels 1–8 were captured at 0, 15, 45, 90, 105, 120, 135 and 150s, respectively. GFP was visualized by direct fluorescence under excitation at 480 nm using confocal microscopy. Size bars represent 5 μm.

both the nucleus and the cytoplasm (Figure 3). GFP-Psc1 transfected COS-1 cells were analysed by confocal microscopy from 10 h post transfection for up to 4 h by capture of images at either 15 s (Figure 3B and D) or 30 s (Figure 3A and C) intervals. Nuclear speckles were largely stationary throughout the analysis although infrequent large-scale movements were observed, with speckles traversing the diameter of the nucleus, fusing and budding (Figure 3A).

By contrast, cytospeckles displayed considerable motility and could be classified into four classes: static, random, directional and tethered. Although continual shape changes were observed, static cytospeckles (10% of cytospeckles) did not move from their position in the cytoplasm throughout the time course (e.g. top arrow in Figure 3B, panel 12). Random movement (5% of cytospeckles) was characterized by short (<5 μm), rapid movement (<15 s) and apparent random directional changes with pauses (15 s–1 min) between movements (Figure 3B). Directional movement (5% of cytospeckles) resulted in straight line movements at ranges of 5–8 μm over a period of ∼15.5 min (Figure 3C). The most abundant class, tethered cytospeckles (80% of cytospeckles), showed no net migration through the cytoplasm with mobility restricted to an estimated 1 μm radius. Cytospeckle size correlated with the patterns of movement. The majority of cytospeckles were <0.5 μm in diameter, evenly distributed throughout the cytosol and demonstrated tethered motility. Larger speckles, in the order of 1 μm, were more likely to demonstrate directional movement. Within the cytoplasm, numerous speckle–speckle interactions were observed, resulting in cycles of budding and fusion (Figure 3C, panel 5) of cytospeckles throughout the time course. Cytospeckle trafficking was consistent in both the presence and absence of Hoechst 33342 stain (data not shown).

A subpopulation of larger cytospeckles (<1%) was observed in close proximity to the nuclear membrane (Figure 3D) and demonstrated an apparent translocation into the nucleus associated with distortion to a crescent shape during translocation (Figure 3D, panel 5). Throughout the course of the analysis no nuclear export of speckles were observed, suggesting that Psc1 aggregation occurs in the cytoplasm, either by recruitment to cytospeckles or *de novo* formation.

The RS domain of Psc1 facilitates nuclear import and assembly into nuclear speckles but is not required for cytospeckle formation

The significance of the RS domain for Psc1 subcellular localization was investigated by analysing the cytoplasmic and nuclear distribution of a Psc1 RS deletion mutant lacking residues 141–240 (GFP–Psc1ΔRS), and an RS domain fusion protein (GFP–RS) containing Psc1 amino acids 141–240. The percentage of cells showing cytoplasmic localization of GFP–Psc1ΔRS (cytoplasmic alone or nuclear + cytoplasmic) increased by 78% compared with full-length Psc1, while GFP–RS was always localized in the nucleus and demonstrated a 40% decrease in cytoplasmic localization compared with GFP–Psc1 (Figure 4A). Cells in which GFP–Psc1ΔRS was excluded from the nucleus (Figure 4B) increased 750% compared with full-length Psc1. These observations are indicative of a role for the Psc1 RS domain in nuclear import.

While GFP–Psc1ΔRS integrated into punctate nuclear compartments (Figure 4C and D), these were often observed to partially overlap or localize to confined regions within SC35-containing nuclear speckles (arrows, Figure 4C) and were also associated with a varying degree of diffuse background staining, indicating a requirement for the RS domain for faithful nuclear targeting of Psc1. While GFP–RS demonstrated a diffuse background nuclear distribution, it also assembled into nuclear speckles, which colocalized with

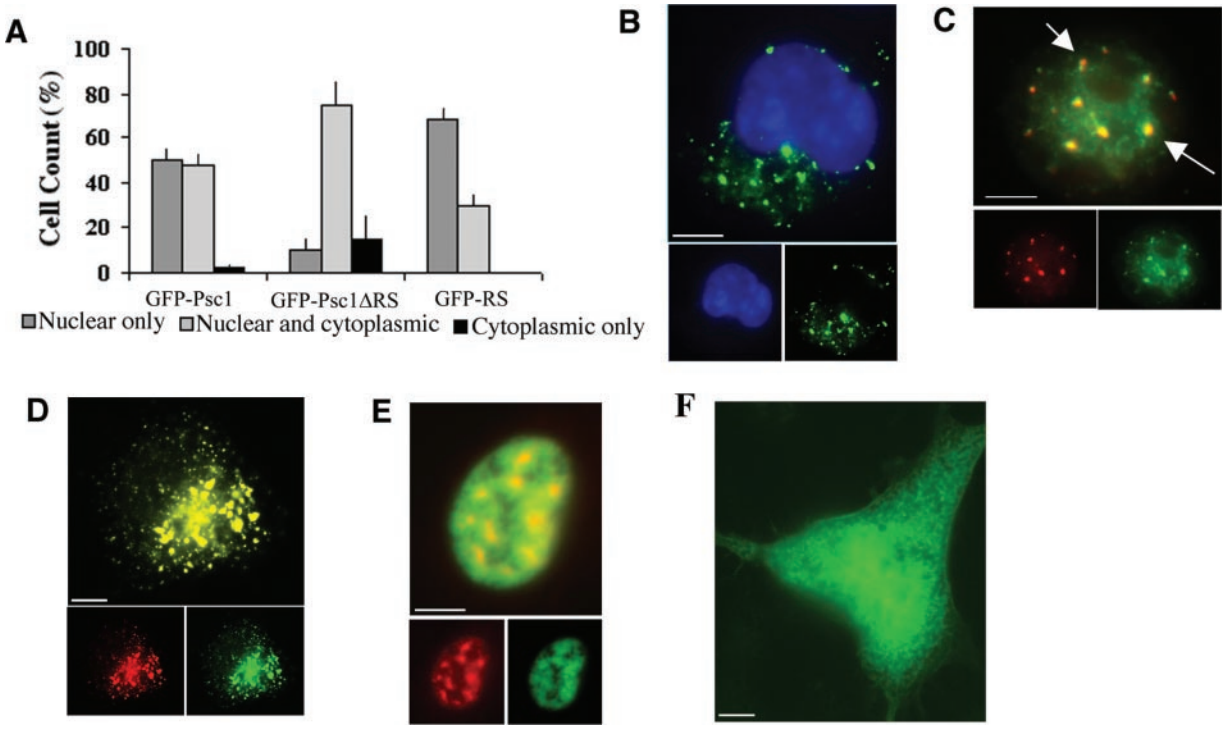


Figure 4. The role of the Psc1 RS domain in Psc1 subcellular localization. (**A**) Subcellular distribution of Psc1 protein in Psc1, GFP–Psc1ΔRS and GFP–RS transfected COS-1 cells. Error bars indicate standard deviation. (**B**) COS-1 cell transfected with GFP–Psc1ΔRS and visualized by direct fluorescence (lower right). Nuclei stained with Hoechst 333258 (lower left). The top panel shows merged images. (**C**) COS-1 cell transfected with GFP–Psc1ΔRS and visualized by direct fluorescence (lower right) or anti-SC35 antibody (lower left). Top panel shows merged images. (**D**) COS-1 cell co-transfected with GFP–Psc1ΔRS and Psc1-HA, and visualized by direct fluorescence (lower right) or anti-HA TRITC (lower left). Top panel shows merged images. (**E**) COS-1 cell transfected with GFP–RS and visualized by direct fluorescence (lower right) or anti SC35 antibody (lower left). Top panel shows merged images. (**F**) COS-1 cell transfected with GFP–RS and visualized by direct fluorescence showing diffuse distribution in the nucleus and the cytoplasm as observed in 30% of cells. All images were captured using conventional microscopy. Size bars represent 10 μm.

SC35 (Figure 4E). Nuclear speckle localization was not always apparent, however, with ~30% of GFP–RS transfected cells showing diffuse staining (Figure 4F). The RS domain is therefore necessary but not sufficient for assembly of Psc1 into nuclear speckles.

GFP–Psc1ΔRS localized to speckles in the cytoplasm (Figure 4B) and colocalized with cytospeckles in cells cotransfected with Psc1-HA (Figure 4D). Where GFP–RS was localized in the cytoplasm, the staining was diffuse with no cytoplasmic speckle formation in any of the cells analysed (Figure 4F). This confirms that the Psc1 RS domain contains no information relevant to cytoplasmic localization.

The RRM of Psc1 is a functional RNA binding motif

An *in vitro* RNA binding assay was used to determine the ability of the Psc1 RRM to interact with RNA (Figure 5A). As attempts to purify full-length Psc1 were unsuccessful, a GST–RRM fusion protein, encompassing residues 475–685 which include the two strictly conserved amino acid motifs of the ARRS protein RRM, was incubated with *in vitro* transcribed adenovirus major late transcript (42), *Psc1* RNA or *CRTR-1* RNA (43) and analysed by gel electrophoresis. As reported by others (44), GST did not bind RNA (Figure 5A). Presence of a band at 49 kDa, the predicted size of GST–RRM, indicated that GST–RRM interacted with all the three transcripts (data not shown) (Figure 5A). While the specificity of interaction was not addressed by this analysis, binding was abolished by the addition of total RNA from ES cells. These results confirm that the Psc1 RRM can bind to RNA and indicate that there exist transcript(s) within pluripotent cells that can compete for binding with the assayed transcripts.

The RRM is necessary for nuclear localization and is both necessary and sufficient for the integration of Psc1 into cytospeckles

The contribution of the RRM to subcellular localization was analysed using a deletion mutant for the binding domain GFP–Psc1ΔRRM (deleted amino acids 527–623), and a GFP–RRM fusion protein containing Psc1 amino acids 475–685. Plasmids were transfected into COS-1 cells for 10 h prior to scoring the transfected populations for subcellular localization of protein in the nuclear and cytoplasmic compartments. Both GFP–RRM and GFP–Psc1ΔRRM were localized in the nucleus of almost 100% of cells (Figure 5B). In the case of GFP–Psc1ΔRRM, the distribution was exclusively nuclear in over 99% of cells, in contrast to the localization of GFP–Psc1 to the cytoplasm of 50% of transfected cells. The exclusion of GFP–Psc1ΔRRM from the cytoplasm points to a critical role for this motif in cytoplasmic localization of Psc1 or integration of Psc1 into cytospeckles.

Within the nucleus, the localization of both GFP–Psc1ΔRRM (Figure 5C) and GFP–RRM (Figure 5D), was

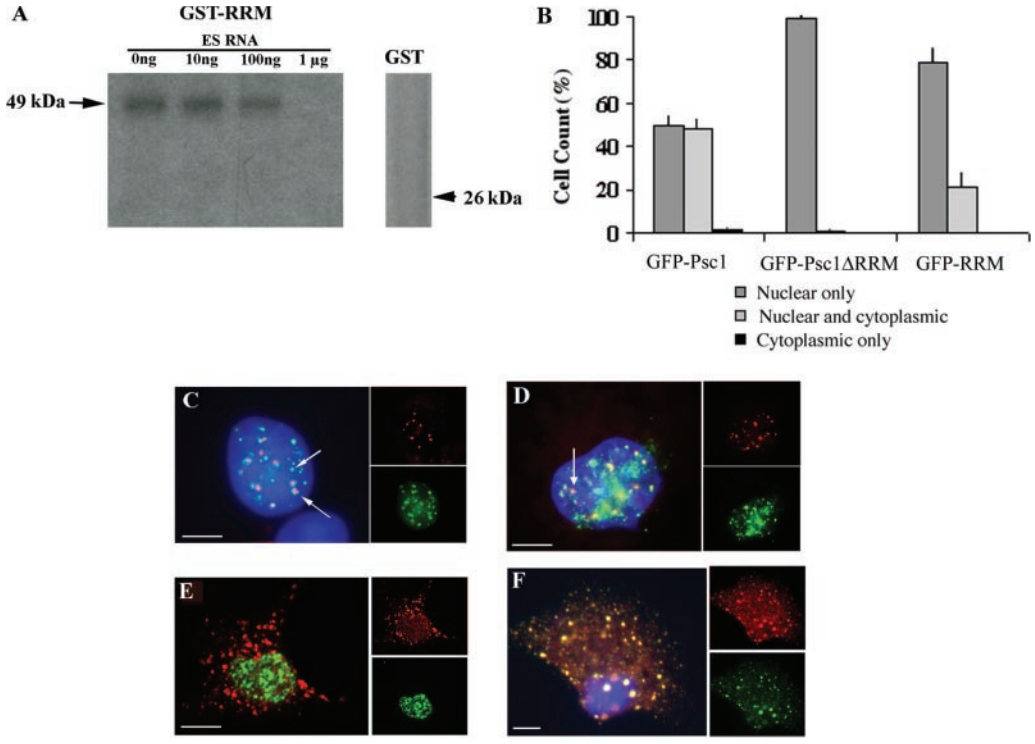


Figure 5. The role of the Psc1 RRM in Psc1 subcellular localization. (**A**) RNA binding assay. GST–RRM and GST were expressed in bacteria and purified. An aliquot of 1 µg of the protein was cross linked for 10 min at 254 nm with [^{32}P]UTP adenovirus RNA, treated with RNase A and analysed by 12.5% SDS–PAGE followed by autoradiography. Unlabelled ES cell total RNA was used as a competitor. The positions of the GST–RRM and GST proteins are indicated. (**B**) Subcellular distribution of Psc1 protein in Psc1, GFP–Psc1ΔRRM and GFP–RRM transfected COS-1 cells. Error bars indicate standard deviation. (**C**) COS-1 cell transfected with GFP–Psc1ΔRRM visualized by direct fluorescence (lower right) or anti-SC35 antibody (top right). Left panel shows merged images. (**D**) COS-1 cell transfected with GFP–RRM visualized by direct fluorescence (lower right) or anti-SC35 antibody (top right). Left panel shows merged images (**E**) COS-1 cell cotransfected with GFP–Psc1ΔRRM and Psc1–HA visualized by direct fluorescence (lower right) or anti-HA TRITC (top right). Left panels show merged images. (**F**) COS-1 cell cotransfected with GFP–RRM and Psc1–HA and visualized by direct fluorescence (lower right) or anti-HA TRITC (top right). Left panel shows merged images. Size bars represent 10 µm. Panel E is a confocal image, all other images were obtained by conventional microscopy.

diffuse with defined punctate regions which partially overlapped (arrow, Figure 5D), or localized to confined regions within nuclear speckles, similar to the mislocalization observed for GFP–Psc1ΔRS. Both GFP–Psc1ΔRRM^{-}/SC35^{+} (bottom arrow Figure 5C) and GFP–Psc1ΔRRM^{+}/SC35^{-} (top arrow, Figure 5C) containing speckles were observed. These results suggest that the RRM and perhaps, RNA binding is necessary but not sufficient for proper localization of Psc1 to nuclear speckles.

In cells cotransfected with GFP–Psc1ΔRRM and Psc1–HA, GFP–Psc1ΔRRM remained nuclear, while full-length protein was found in speckles in both the nucleus and the cytoplasm (Figure 5E). By contrast, GFP–RRM protein in the cytoplasm was localized to punctate structures reminiscent of cytospeckles, and in cells cotransfected with GFP–RRM and Psc1–HA colocalization of the proteins was observed in cytospeckles (Figure 5F). These results indicate that the RRM, and perhaps RNA binding, is obligatory and sufficient for localization of Psc1 to cytospeckles.

The Psc1 C-terminal domain may be required for trafficking between the nucleus and cytoplasm via microtubules

The C domain was identified solely on the basis of homology between ARRS proteins and homologues, and data base analysis failed to identify any putative role for this domain. The contribution of the C domain to subcellular localization was analysed using a C domain deletion mutant GFP–Psc1ΔCD (deleted amino acids 738–893), and a GFP–CD fusion protein inclusive of the C domain and RG repeats (Psc1 amino acids 731–961). Plasmids were transfected into COS-1 cells for 10 h prior to scoring the transfected populations for subcellular localization of protein in the nuclear and cytoplasmic compartments. There was a 76% decrease in cells containing GFP–Psc1ΔCD in the nucleus compared with GFP–Psc1 (Figure 6A), suggesting a role for the C domain in nuclear entry and/or retention. Exclusion of GFP–CD from the nucleus (Figure 6A) suggests that the former is a more probable explanation. Within the nucleus, while Psc1–HA and all nuclear speckles colocalized with GFP–Psc1ΔCD (arrowheads, Figure 6B and arrows, Figure 6C), the distribution of GFP–Psc1ΔCD often extended beyond the nuclear speckle as defined by staining for Psc1–HA (Figure 6B) or SC35 (Figure 6C).

Within the cytoplasm, GFP–Psc1ΔCD formed punctate structures but these did not localize reliably with Psc1–HA containing cytospeckles (upper arrows, Figure 6B). GFP–CD (Figure 6D) was restricted to the cytoplasm (Figure 6A), where it did not form cytospeckles but colocalized with α-tubulin (Figure 6A) in the presence of varying degrees of diffuse cytoplasmic staining. Analysis of GFP–Psc1ΔCD and GFP–CD therefore, suggests an association between the Psc1 C

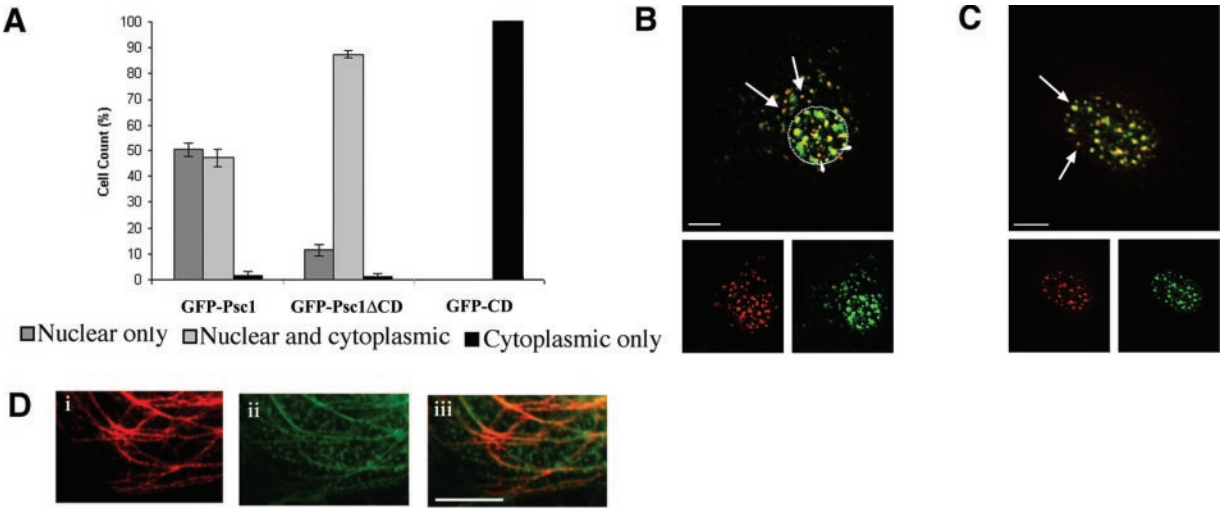


Figure 6. The role of C-terminal elements in Psc1 subcellular localization. (**A**) Subcellular distribution of Psc1 protein in Psc1, GFP–Psc1ΔCD and GFP–CD transfected COS-1 cells. Error bars indicate standard deviation. (**B**) COS-1 cell cotransfected with GFP–Psc1ΔCD and Psc1–HA, visualized by direct fluorescence (lower right) or anti-HA TRITC (lower left). Top panel shows merged images. The nucleus is outlined by a dotted line. (**C**) COS-1 cell transfected with GFP–Psc1ΔCD, visualized by direct fluorescence (lower right) or anti-SC35 antibody (lower left). GFP–Psc1ΔCD is restricted to the nucleus in this example. Top panel shows merged images. (**D**) Representative cytoplasmic section of COS-1 cells transfected with GFP–CD and visualized by anti α-tubulin antibody (i) or by direct fluorescence (ii). Merged images shown in (iii). Size bars represent 10 μm. All images were obtained by conventional microscopy.

domain and the microtubule component of the cytoskeleton that affects the subcellular distribution of GFP–Psc1 between the nuclear and cytoplasmic compartments.

DISCUSSION

ARRS proteins are conserved in evolution

Psc1 and BAC34721 were identified as related proteins in mouse with a domain structure which defines ARRS domain containing proteins. ARRS proteins are typically large, in the order of 800–1100 amino acids, and are defined by the sequential arrangement of an N-terminal domain with homology to Hprp3p, RS domain, RRM with unique conserved motifs, C-domain homology, an acidic rich region adjacent to the C-terminus and with the exception of the *C.elegans* and *D.discoideum* homologues, a C-terminal RSWR/K motif. Phylogenetic analyses (Figure 1C) indicate that ARRS proteins share a common evolutionary origin. ARRS proteins remain monophyletic to a single putative gene ancestor. Our analyses show the slime-mould protein (AAO51188) is close to the centre (root) of a hypothetical evolutionary tree highlighting the deep biological origin of this protein family. Orthologues of a single gene were easily identified in the mosquito, fruit fly and the nematode worm. However, a putative gene duplication event specific to the vertebrate lineage obscures the order of descent of the two conserved vertebrate homologues, represented by Psc1 and BAC34721 in the mouse (Figure 1C). Slightly deeper evolutionary nodes and longer branch lengths suggest the Psc1 clade may be parental to the BAC34721 clade raising the possibility that Psc1 is the orthologue of the invertebrate ancestor. Arguably, a high degree of structural conservation between ARRS proteins reflects conserved functional roles for these proteins in eukaryotes. This diversification of function appears to have arisen within the vertebrate lineage at least 450 million years ago, the estimated time of divergence between human and puffer fish. Interestingly, the duplicated vertebrate proteins have structural differences in that human KIAA1311 and mouse Psc1 proteins contain a PG repeat domain not found in the BAC34721 clade, raising the possibility of diverged functional roles between the two proteins.

Members of the SR protein family are a well characterized family of RS domain proteins, and have been shown to mediate protein–protein and protein–RNA interactions in the spliceosome. However, it is clear that this description understates SR protein function. Tissue expression variability (45), apoptotic regulation (46), developmental requirement (47), differential RNA binding specificities (2) and roles in cancer/disease states (48,49), demonstrate the extent to which SR proteins are involved in the regulation of cellular events. The presence of features consistent with SR proteins such as an RS domain, a functional RRM and localization to nuclear speckles, suggests that at least one aspect of ARRS protein function is likely to be involved with RNA processing.

Determinants of Psc1 nuclear localization

GFP–Psc1 was localized to punctate areas of the nucleus and colocalized to all nuclear speckles stained with anti SC35, a distribution Lamond and Spector define as diagnostic for proteins involved in pre-mRNA splicing (8). The additional diffuse background nuclear staining observed for over expressed GFP–Psc1 has been reported for over expressed splicing factors including SF2/ASF in Hela cells (21).

GFP–Psc1 and Psc1–HA also localized to additional speckles within the nucleus that did not contain either SC35 or SF2/ASF, and were often located at or near the nuclear membrane (Figure 2A, panel v). A similar subcellular localization pattern was observed for endogenous Psc1 protein in COS-1 cells. The apparent ingression of cytospeckles observed in real time is a possible explanation for these additional sites of GFP–Psc1

localization in the nucleus and consistent with the absence of other RS domain proteins from both these sites and cytospeckles.

Nuclear localization was observed in GFP fusion proteins containing either the RRM or the RS domain, suggesting that both of these domains contribute to nuclear targeting. In addition, the inefficient nuclear localization observed for the GFP–Psc1ΔRS protein suggests a central role for the Psc1 RS domain in nuclear import, consistent with that reported for many SR proteins (50). Partial colocalization of GFP–RS, GFP–Psc1ΔRS, GFP–RRM and GFP–Psc1ΔRRM with SC35 suggests that both the RS domain and the RRM of Psc1 contain some of the information required for nuclear speckle localization. These results support a model for cooperativity of the RRM and RS domains of Psc1 in the regulation of protein trafficking and subcellular nuclear localization. A cooperative relationship between these domains has also been reported for SF2/ASF (21).

Regions rich in arginine and glycine are capable of mediating protein–protein interactions (51), subcellular localization and RNA binding (52). All identified vertebrate ARRS proteins contain an RG-rich region, however, each lacks the RGG Box typically observed in RNA binding proteins (53) and reported to contribute to a number of nuclear functions such as nucleolar/nuclear targeting (54) and, protein and RNA interactions (55). The Psc1 RG-rich region consists of interspersed and consecutive RG dipeptide repeats, similar to the RG Box found in p80 coilin which localizes SMN to cajal bodies (56). While cajal bodies are not sites of active splicing, they are biogenic sites for snRNPs, which subsequently traffic to nuclear speckles and are involved in pre-messenger RNA processing.

Nuclear speckles containing SR proteins are heterogeneous

Subnuclear localization patterns of Psc1 and Psc1 mutant proteins point to the existence of heterogeneity amongst nuclear speckles in two respects. First, full-length Psc1 and GFP–Psc1ΔRRM localized to nuclear speckles that did not contain SR proteins such as SC35, suggesting a diversity of molecular composition amongst SR protein containing structures in the nucleus. Variability amongst nuclear speckles has been described by Zhang *et al*. (57) and others, who suggested a relationship between shape and function, with irregular speckles active in the recruitment/trafficking of splicing factors, and regular, rounded speckles forming in the absence of active transcription. The presence of splicing factors such as SC35 in interchromatin granules, sites implicated in spliceosome assembly (58) and perichromatin fibrils, associated with active pre-mRNA transcription and processing (59), also indicates a relationship between speckle localization/composition and function. In this case, we identify sub nuclear localization as an indicator of speckle heterogeneity, with speckles containing Psc1 but not marked by anti-SC35 often associated with the nuclear periphery. Spatial heterogeneity within individual speckles was evident from the fact that both GFP–Psc1ΔRRM and GFP–Psc1ΔRS localized to discrete regions within nuclear speckles that overlapped partially but not completely with the anti-SC35 nuclear speckle marker. This is consistent with the localization of cyclin T1, Cdk9 (60) and β-actin mRNA (61), each of which demonstrate partial or limited overlap with nuclear speckles. Partial overlap may be associated with the formation of 'subdomains' which have been described by Mintz and Spector (62) as 5–50 spherical structures within nuclear speckles, heterogenous in size and composed of SR proteins and snRNPs, implying a function separate to factors uniformly found within nuclear speckles.

Psc1 is located within discrete cytoplasmic structures called cytospeckles

Given the precedents for RS domain protein localization in the nucleus, the presence of Psc1-containing speckles in the cytoplasm of interphase cells (called cytospeckles) was unexpected. Observation of these structures in both monkey kidney cells (COS-1) and mouse pluripotent cells (EPL) indicates that they are not cell type- or species-specific. Novel aspects of cytoplasmic speckling are likely to be common to all ARRS proteins given the observation of *Drosophila* NP_609976 and human se70-2 speckling in the cytoplasm of SL3 cells and HeLa cells, respectively (data not shown). Psc1-containing cytospeckles did not colocalize with endoplasmic reticulum, mitochondria, lysosomes, actin, γ or α-tubulin, and their distribution or morphology were not affected by treatment of transfected cells with the microtubule depolymerizing agents, nocodazole and colchicine, alteration of COS-1 cell seeding densities, or variation in transfection time from 8 to 72 h (data not shown). Psc1-containing cytospeckles are therefore, not identical to or associated with these structures. RS domain proteins have previously been identified in cytoplasmic structures. SR proteins have been observed in the two-cell stage of the nematode *Ascaris lumbricoides* (23). However, unlike Psc1 containing cytospeckles, the nematode cytosolic speckles are only observed prior to zygotic gene activation and contain SC35. The yeast actin binding protein, Sla1p contains an RS domain and in addition to suggested nuclear roles (1), is localized to cortical actin in the cytoplasm and regulates actin assembly with a role in endocytosis. Although the function of the Sla1p RS domain is unknown, deletion mutants inclusive of this domain did not perturb cytoplasmic localization (63). A cytoplasmic localization profile can be conferred upon SF2/ASF by amino acid substitution of the RS residues for RG residues (64). This localization was largely diffuse, and although cytoplasmic punctate structures were apparent, these did not resemble Psc1-containing cytospeckles. Cytospeckles are not a prerequisite for nuclear speckle formation as the GFP–Psc1ΔRRM mutant was capable of forming nuclear speckles in the absence of cytospeckle formation (Figure 2D). Formation of Psc1-containing cytospeckles did not appear to result from the export of nuclear speckles and did not require integration of Psc1 into nuclear speckles as Psc1 and GFP–RRM containing cytospeckles were observed in the absence of nuclear speckles. It is therefore, assumed that these structures form *de novo* within the cytoplasm.

Cytospeckles do not appear to be associated with sites of RNA degradation as they do not colocalize with GW182, a marker for GW/DCP bodies (65,66) (data not shown), nor do they resemble stress granules which form under conditions of oxidative stress (66,67), although this observation has not been verified experimentally. The appearance of cytospeckles

is reminiscent of RNA containing granules (68), large cytoplasmic complexes which contain multiple proteins and RNA (69–72). Statistical analysis suggests that at least in the case of A2RE/hnRNP RNA, each RNA granule is heterogeneous with respect to RNA content and contains ~30 RNA molecules (73). Consistent with the molecular composition of RNA granules, cytospeckles are deduced to contain multiple protein molecules since GFP–Psc1 cytospeckles could be visualized easily using light microscopy, suggesting that multiple Psc1 molecules are integrated into these structures. Further, the demonstrated ability of the Psc1 RRM to bind at least two transcripts expressed within pluripotent cells and the intron containing adenovirus transcript (Figure 5A), together with the obligatory requirement for the RRM to direct GFP–Psc1 to cytospeckles, suggests that cytospeckles are likely to contain a heterogeneous RNA population. It is possible that RNA binding specificity is directed by domains outside the Psc1 RRM, in which case cytospeckle RNA content could be restricted to a limited repertoire of cellular transcripts.

Trafficking of RNA granules occurs via continuous cycles of anchoring and active transport associated with the cytoskeletal network. Fusco *et al*. (72) report distinct patterns of RNA motility including completely immobile, corralled and non-restricted diffusion, similar to these observed during real time analysis of Psc1 cytospeckles. The failure of cytoplasmic Psc1 to colocalize with F-actin and α-tubulin suggests that, if Psc1 cytospeckles traffic via microtubule/actin networks, their associations with these components must be transient. Further investigation is required to determine the rate and pattern of movement given the possible involvement of bi-directional motorized transport (74) and/or treadmilling in association with the ends of dynamic microtubules (75).

Association with cytoskeletal filaments may result from interaction with the C domain, shown to colocalize with microtubles (Figure 6D). For those cytospeckles whose fate is proposed to be nuclear entry (Figure 3D), cytoskeletal association via the C domain may be a significant contributor to the nuclear import pathway given the increased cytoplasmic compartmentalization observed for GFP–Psc1ΔCD (Figure 6A).

RNA granules are proposed to contain all machinery components required for translation and play a role in the regulation of site-specific and temporal translational regulation (69). While cytospeckles may, by association, be involved in translational regulation, a further possibility is a role for Psc1-containing complexes in the cytoplasm in the storage of Psc1 or Psc1-associated proteins/RNA. In this context, Psc1 may have no function within the cytospeckle, but await signals that mediate transport to sites of functional relevance. The complex subcellular localization and trafficking is consistent with a novel role for Psc1 in the coordination of cytoplasmic events and nuclear RNA metabolism.

ACKNOWLEDGEMENTS

We thank Dr Joy Rathjen and Dr Rebecca Keough for helpful discussions and critical reading of the manuscript. We are also grateful to Dr Peter Kolesik for assistance with confocal microscopy and Mr Brian Denton for preparation of materials used in the real time analysis. This work was supported by grants from the Australian Research Council and the ARC SRC for the Molecular Genetics of Development. Funding to pay the Open Access publication charges for this article was provided by the Australian Research Council Special Research Centre in Molecular Genetics, University of Adelaide, Adelaide, 5005, South Australia.

REFERENCES

1. Boucher,L., Ouzounis,C.A., Enright,A.J. and Blencowe,B.J. (2001) A genome-wide survey of RS domain proteins. *RNA*, **7**, 1693–1701.
2. Graveley,B.R. (2000) Sorting out the complexity of SR protein functions. *RNA*, **6**, 1197–1211.
3. Tacke,R. and Manley,J.L. (1999) Determinants of SR protein specificity. *Curr. Opin. Cell Biol*., **11**, 358–362.
4. Valcarcel,J. and Green,M.R. (1996) The SR protein family: pleiotropic functions in pre-mRNA splicing. *Trends Biochem. Sci*., **21**, 296–301.
5. Fu,X.-D. (1995) The superfamily of arginine/serine rich splicing factors. *RNA*, **1**, 663–680.
6. Hedley,M.L., Amrein,H. and Maniatis,T. (1995) An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proc. Natl Acad. Sci. USA*, **92**, 11524–11528.
7. Fu,X.-D. and Maniatis,T. (1990) Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature*, **343**, 437–441.
8. Lamond,A.I. and Spector,D.L. (2003) Nuclear speckles: a model for nuclear organelles. *Nature Rev. Mol. Cell Biol*., **4**, 605–612.
9. Jimenez-Garcia,L.F. and Spector,D.L. (1993) *In vivo* evidence that transcription and splicing are coordinated by a recruiting mechanism. *Cell*, **73**, 47–59.
10. Misteli,T., Caceres,J.F. and Spector,D.L. (1997) The dynamics of a pre-mRNA splicing factor in living cells. *Nature*, **387**, 523–527.
11. Bregman,D.B., Du,L., van der Zee,S. and Warren,S.L. (1995) Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell Biol*., **129**, 287–298.
12. Dostie,J., Lejbkowicz,F. and Sonenberg,N. (2000) Nuclear eukaryotic initiation factor 4E (eIF4E) colocalizes with splicing factors in speckles. *J. Cell Biol*., **148**, 239–247.
13. Zhao,K., Wang,W., Rando,O.J., Xue,Y., Swiderek,K., Kuo,A. and Crabtree,G.R. (1998) Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. *Cell*, **95**, 625–636.
14. Manley,J.L. and Tacke,R. (1996) SR proteins and splicing control. *Genes Dev*., **10**, 1569–1579.
15. Graveley,B.R. (2004) A protein interaction domain contacts RNA in the prespliceosome. *Mol. Cell*, **13**, 302–30416.
16. Allemand,E., Dokudovskaya,S., Bordonne,R. and Tazi,J. (2002) A conserved *Drosophila* transportin-serine/arginine-rich (SR) protein permits nuclear import of *Drosophila* SR protein splicing factors and their antagonist repressor splicing factor 1. *Mol. Biol. Cell*, **13**, 2436–2447.
17. Kataoka,N., Bachorik,J.L. and Dreyfuss,G.J. (1999) Transportin-SR, a nuclear import receptor for SR proteins. *Cell Biol*., **145**, 1145–1152.
18. Lai,M.C., Lin,R.I. and Tarn,W.Y. (2001) Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc. Natl Acad. Sci. USA*, **98**, 10154–10159.
19. Li,H. and Bingham,P.M. (1991) Arginine/serine-rich domains of the su(wa) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell*, **18**, 335–342.
20. Philipps,D., Celotto,A.M., Wang,Q.Q., Tarng,R.S. and Graveley,B.R. (2003) Arginine/serine repeats are sufficient to constitute a splicing activation domain. *Nucleic Acids Res*., **15**, 6502–6508.
21. Caceres,J.F., Misteli,T., Screaton,G.R., Spector,D.L. and Krainer,A.R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol*., **138**, 225–238.
22. Caceres,J.F., Screaton,G.R. and Krainer,A.R. (1998) A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev*., **12**, 55–66.
23. Sanford,J.R. and Bruzik,J.P. (2001) Regulation of SR protein localization during development. *Proc. Natl Acad. Sci. USA*, **98**, 10184–10189.

24. Huang,Y., Yario,T.A. and Steitz,J.A. (2004) A molecular link between SR protein dephosphorylation and mRNA export. *Proc. Natl Acad. Sci. USA*, **101**, 9666–9670.
25. Gilbert,W. and Guthrie,C. (2004) The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. *Mol. Cell*, **13**, 201–212.
26. Tacke,R., Chen,Y. and Manley,J.L. (1997) Sequence specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40 specific splicing enhancer. *Proc. Natl Acad. Sci. USA*, **94**, 1148–1153.
27. Pelton,T.A., Sharma,S., Schulz,T.C., Rathjen,J. and Rathjen,P.D. (2002) Transient pluripotent cell populations during primitive ectoderm formation: correlation of *in vivo* and *in vitro* pluripotent cell development. *J. Cell Sci.*, **115**, 329–339.
28. Chapman,G. and Rathjen,P.D. (1995) Sequence and evolutionary conservation of the murine Gbx-2 homeobox gene. *FEBS Lett.*, **364**, 289–292.
29. Stockham,C., Wang,L.S. and Warnow,T. (2002) Statistically based postprocessing of phylogenetic analysis by clustering. *Bioinformatics*, **18**(Suppl. 1), 285–293.
30. Rathjen,P.D., Toth,S., Willis,A., Heath,J.K. and Smith,A.G. (1990) Differentiation inhibiting activity is produced in matrix-associated and diffusible forms that are generated by alternate promoter usage. *Cell*, **62**, 1105–1114.
31. Rathjen,J., Lake,J.A., Bettess,M.D., Washington,J.M., Chapman,G. and Rathjen,P.D. (1999) Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors. *J. Cell Sci.*, **112**, 601–612.
32. Durand,R.E. and Olive,P.L. (1982) Cytotoxicity, mutagenicity and DNA damage by Hoechst 33342. *J. Histochem. Cytochem.*, **30**, 111–116.
33. Zuccotti,M., Ponce,R.H., Boiani,M., Guizzardi,S., Govoni,P., Scandroglio,R., Garagna,S. and Redi,C.A. (2002) The analysis of chromatin organisation allows selection of mouse antral oocytes competent for development to blastocyst. *Zygote*, **10**, 73–78.
34. Krainer,A.R., Conway,G.C. and Kozak,D. (1990) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev.*, **4**, 1158–1171.
35. Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosawa,M. and Ohara,O. (2000) Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins *in vitro*. *DNA Res.*, **7**, 65–73.
36. Gonzalez-Santos,J.M., Wang,A., Jones,J., Ushida,C., Liu,J. and Hu,J. (2002) Central region of the human splicing factor Hprp3p interacts with Hprp4p. *J. Biol. Chem.*, **277**, 23764–23772.
37. Screaton,G.R., Caceres,J.F., Mayeda,A., Bell,M.V., Plebanski,M., Jackson,D.G., Bell,J.I. and Krainer,A.R. (1995) Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J.*, **14**, 4336–4349.
38. Marchler-Bauer,A., Anderson,J.B., DeWeese-Scott,C., Fedorova,N.D., Geer,L.Y., He,S., Hurwitz,D.I., Jackson,J.D., Jacobs,A.R., Lanczycki,C.J. *et al.* (2003) CDD: a curated Entrez database of conserved domain alignments. *Nucleic Acids Res.*, **31**, 383–387.
39. Misteli,T. and Spector,D.L. (1998) The cellular organization of gene expression. *Curr. Opin. Cell Biol.*, **10**, 323–331.
40. Carninci,P. and Hayashizaki,Y. (1999) High-efficiency full-length cDNA cloning. *Meth. Enzymol.*, **303**, 19–44.
41. Eichmüller,S., Usener,D., Dummer,R., Stein,A., Thiel,D. and Schadendorf,D. (2001) Serological detection of cutaneous T-cell lymphoma-associated antigens. *Proc. Natl Acad. Sci. USA*, **98**, 629–634.
42. Graham,F.L. and Prevec,L. (1991) Manipulation of adenovirus vectors. *Methods Mol. Biol.*, **7**, 109–128.
43. Rodda,S., Sharma,S., Scherer,M., Chapman,G. and Rathjen,P. (2001) CRTR-1, a developmentally regulated transcriptional repressor related to the CP2 family of transcription factors. *J. Biol. Chem.*, **276**, 3324–3332.
44. Igel,H., Wells,S., Perriman,R. and Ares,M.,Jr (1998) Conservation of structure and subunit interactions in yeast homologues of splicing factor 3b (SF3b) subunits. *RNA*, **4**, 1–10.
45. Zahler,A.M., Neugebauer,K.M., Lane,W.S. and Roth,M.B. (1993) Distinct functions of SR proteins in alternative pre-mRNA splicing. *Science*, **260**, 219–222.
46. Jiang,Z.-H., Zhang,W.-J., Rao,Y. and Wu,J.Y. (1998) Regulation of Ich-1 pre mRNA alternative splicing and apoptosis by mammalian splicing factors. *Proc. Natl Acad. Sci. USA*, **95**, 9155–9160.
47. Jumaa,H., Wei,G. and Nielson,P.J. (1999) Blastocyst formation is blocked in mouse embryos lacking the splicing factor SRp20. *Curr. Biol.*, **9**, 899–902.
48. Yang,L., Embree,L.J. and Hickstein,D.D. (2000) TLS–ERG leukemia fusion protein inhibits RNA splicing mediated by serine-arginine proteins. *Mol. Cell. Biol.*, **20**, 3345–3354.
49. Young,P.J., DiDonato,C.J., Hu,D., Kothary,R., Androphy,E.J. and Lorson,C.L. (2002) SRp30c-dependent stimulation of survival motor neuron (SMN) exon 7 inclusion is facilitated by a direct interaction with hTra2 beta 1. *Hum. Mol. Genet.*, **11**, 577–587.
50. Kataoka,N., Bachorik,J.L. and Dreyfuss,G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. *J. Cell Biol.*, **145**, 1145–11452.
51. Bouvet,P., Diaz,J.J., Kindbeiter,K., Madjar,J.J. and Amalric,F. (1998) Nucleolin interacts with several ribosomal proteins through its RGG domain. *J. Biol. Chem.*, **273**, 19025–19029.
52. Miller,M.M. and Read,L.K. (2003) *Trypanosoma brucei*: functions of RBP16 cold shock and RGG domains in macromolecular interactions. *Exp. Parasitol.*, **105**, 140–148.
53. Burd,C.G. and Dreyfuss,G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
54. Gendra,E., Moreno,A., Alba,M.M. and Pages,M. (2004) Interaction of the plant glycine-rich RNA-binding protein MA16 with a novel nucleolar DEAD box RNA helicase protein from *Zea mays*. *Plant J.*, **38**, 875–886.
55. Heine,M.A., Rankin,M.L. and DiMario,P.J. (1993) The Gly/Arg-rich (GAR) domain of *Xenopus* nucleolin facilitates *in vitro* nucleic acid binding and *in vivo* nucleolar localization. *Mol. Biol. Cell*, **4**, 1189–1204.
56. Gall,J.G. (2001) A role for Cajal bodies in assembly of the nuclear transcription machinery. *FEBS Lett.*, **498**, 164–167.
57. Zhang,G., Taneja,K.L., Singer,R.H. and Green,M.R. (1994) Localization of pre-mRNA splicing in mammalian nuclei. *Nature*, **372**, 809–812.
58. Puvion-Dutilleul,F., Bachellerie,J.P., Visa,N. and Puvion,E. (1994) Rearrangements of intranuclear structures involved in RNA processing in response to adenovirus infection. *J. Cell Sci.*, **107**, 1457–1468.
59. Cmarko,D., Verschure,P.J., Martin,T.E., Dahmus,M.E., Krause,S., Fu,X.-D., van Driel,R. and Fakan,S. (1999) Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol. Biol. Cell*, **10**, 211–223.
60. Herrmann,C.H. and Mancini,M.A. (2001) The Cdk9 and cyclin T subunits of TAK/P-TEFb localize to splicing factor-rich nuclear speckle regions. *J. Cell Sci.*, **114**, 1491–1503.
61. Xing,Y., Johnson,C.V., Moen,P.T., McNeil,J.A. and Lawrence,J. (1995) Nonrandom gene organization: structural arrangements of specific pre-mRNA transcription and splicing with SC-35 domains. *J. Cell Biol.*, **131**, 1635–1647.
62. Mintz,P.J. and Spector,D.L. (2000) Compartmentalization of RNA processing factors within nuclear speckles. *J. Struct. Biol.*, **129**, 241–251.
63. Gourlay,C.W., Dewar,H., Warren,D.T., Costa,R., Satish,N. and Ayscough,K.R. (2003) An interaction between Sla1p and Sla2p plays a role in regulating actin dynamics and endocytosis in budding yeast. *J. Cell Sci.*, **116**, 2551–2564.
64. Cazalla,D., Zhu,J., Manche,L., Huber,E., Krainer,A.R. and Caceres,J.F. (2002) Nuclear export and retention signals in the RS domain of SR proteins. *Mol. Cell. Biol.*, **22**, 6871–6882.
65. Eystathioy,T., Chan,E.K., Takeuchi,K., Mahler,M., Luft,L.M., Zochodne,D.W. and Fritzler,M.J. (2003) Clinical and serological associations of autoantibodies to GW bodies and a novel cytoplasmic autoantigen GW182. *J. Mol. Med.*, **81**, 811–818.
66. Cougot,N., Babajko,S. and Seraphin,B. (2004) Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.*, **165**, 31–40.
67. Kedersha,N. and Anderson,P. (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem. Soc. Trans.*, **30**, 963–969.
68. Carson,J.H., Cui,H. and Barbarese,E. (2001) The balance of power in RNA trafficking. *Curr. Opin. Neurobiol.*, **11**, 558–563.
69. Barbarese,E., Koppel,D.E., Deutscher,M.P., Smith,C.L., Ainger,K., Morgan,F. and Carson,J.H. (1995) Protein translation components are colocalized in granules in oligodendrocytes. *J. Cell Sci.*, **108**, 2781–2790.

70. Carson,J.H., Worboys,K., Ainger,K. and Barbarese,E. (1997) Translocation of myelin basic protein mRNA in oligodendrocytes requires microtubules and kinesin. *Cell Motil. Cytoskeleton*, **38**, 318–328.
71. Schuman,E.M. (1999) mRNA trafficking and local protein synthesis at the synapse. *Neuron*, **23**, 645–648.
72. Fusco,D., Accornero,N., Lavoie,B., Shenoy,S.M., Blanchard,J.M., Singer,R.H. and Bertrand,E. (2003) Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.*, **13**, 161–167.
73. Mouland,A.J., Xu,H., Cui,H., Krueger,W., Munro,T.P., Prasol,M., Mercier,J., Rekosh,D., Smith,R. and Barbarese,E. (2001) RNA trafficking signals in human immunodeficiency virus type 1. *Mol. Cell. Biol.*, **21**, 2133–2143.
74. Ligon,L.A., Tokito,M., Finklestein,J.M., Grossman,F.E. and Holzbaur,E.L. (2004) A direct interaction between cytoplasmic dynein and kinesin I may coordinate motor activity. *J. Biol. Chem.*, **279**, 19201–19208.
75. Shaw,S.L., Kamyar,R. and Ehrhardt,D.W. (2003) Sustained microtubule treadmilling in *Arabidopsis* cortical arrays. *Science*, **300**, 1715–1718.
76. Swofford,D.L. (2000) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.