Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 327 (2005) 143–149
BBRC
www.elsevier.com/locate/ybbrc

Identification of an evolutionary conserved SURF-6 domain in a family of nucleolar proteins extending from human to yeast

Mikhail Polzikov[a,b], Olga Zatsepina[a], Charalambos Magoulas[b,*]

[a] *Laboratory of Structural Biochemistry, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia*
[b] *Institute of Cell and Molecular Science, Neuroscience Centre, Queen Mary, University of London, UK*

Received 5 November 2004
Available online 9 December 2004

Abstract

The mammalian SURF-6 protein is localized in the nucleolus, yet its function remains elusive in the recently characterized nucleolar proteome. We discovered by searching the Protein families database that a unique evolutionary conserved SURF-6 domain is present in the carboxy-terminal of a novel family of eukaryotic proteins extending from human to yeast. By using the enhanced green fluorescent protein as a fusion protein marker in mammalian cells, we show that proteins from distantly related taxonomic groups containing the SURF-6 domain are localized in the nucleolus. Deletion sequence analysis shows that multiple regions of the SURF-6 protein are capable of nucleolar targeting independently of the evolutionary conserved domain. We identified that the *Saccharomyces cerevisiae* member of the SURF-6 family, named rrp14 or ykl082c, has been categorized in yeast databases to interact with proteins involved in ribosomal biogenesis and cell polarity. These results classify SURF-6 as a new family of nucleolar proteins in the eukaryotic kingdom and point out that SURF-6 has a distinct domain within the known nucleolar proteome that may mediate complex protein–protein interactions for analogous processes between yeast and mammalian cells.

Keywords: Nucleolus; SURF-6; Nucleolar localization signals; rrp14; ykl082c; Ribosomal biogenesis; Yeast cell polarity

The nucleolus is a membrane-free nuclear organelle formed on chromosomal clusters of tandemly repeated active rRNA genes and contains the machinery for rRNA synthesis, processing, and ribosomal maturation [1,2]. In addition to ribosomal biogenesis, the nucleolus has been implicated to regulate other cellular processes such as cell cycle [3,4], RNA editing [5], maturation of telomeres [6], stem cell maintenance [7], and viral infections [2]. Although the complete functions of the nucleolus are unknown, our knowledge on the functional complexity of the human nucleolar proteome has been expanded drastically by the recent applications of mass spectrometry [8–10]. Nearly 350 proteins have been identified in the human nucleolus from which around 30% are novel proteins, one of the nucleolar proteins catalogued to have an unknown function is SURF-6 [11].

The mammalian SURF-6 protein, encoded by a member of the Surfeit gene locus, has been previously identified as a nucleolar protein by immuno-cytochemical approaches and it was shown to have a high nucleic-acid-binding capacity in vitro [12–14]. It was shown that SURF-6 has a cellular distribution similar to the major nucleolar proteins fibrillarin and B23/nucleophosmin, which are part of the nucleolar RNA processing machinery [13]. Although SURF-6 has been suggested to interact with nucleolar RNA, its sequence does not share any known functional domains. Functionally significant sequences lie inside evolutionary conserved domains that are identified in the Protein families (Pfam) database of alignments [15]. Based on such sequence alignments

* Corresponding author. Fax: +44 20 7882 7726.
E-mail address: C.B.Magoulas@qmul.ac.uk (C. Magoulas).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.11.148

the evolutionary conserved nucleolar protein families have been recently classified, and 115 known and 91 novel protein domains have been identified [16]. SURF-6, however, does not appear in this recent nucleolar protein domain repertoire. In this study, we identify a unique evolutionary conserved SURF-6 domain from human to yeast. We demonstrate, using enhanced green fluorescent protein (EGFP) as a protein tag, that the proteins containing this domain belong to a novel nucleolar protein family. Moreover, the importance of the conserved domain in the nucleolar localization capacity of SURF-6 is investigated. In addition, we identify that the *Saccharomyces cerevisiae* SURF-6 homologue is the ykl082c gene, as catalogued in the Yeast Databases (http://biodata.mshri.on.ca/yeast_grid), encoding the rrp14 protein which has been characterized based on functional proteomic studies to interact with proteins involved in cell polarity [17,18] and ribosomal biogenesis [19–21] pointing to potential functions of the identified SURF-6 domain in mammalian cells.

Materials and methods

Computer analysis. Multiple sequence alignment of SURF-6 was performed by the ClustalW alignment method in the Pfam database [15]; this database is available in the web site of the Sanger Institute (www.sanger.ac.uk/cgi-bin/Pfam). Further analysis of the multiple sequence alignment was performed by GeneDoc [22]; this software is provided in www.psc.edu/biomed/genedoc/. Secondary structure of the SURF-6 domain was predicted by the PSIPRED method [23], and by using the web site of the Bioinformatics Unit of University College London (www.bioinf.cs.ucl.ac.uk/psipred). Multiple nuclear localization sequences (NLS) were predicted by the PROSITE database [24].

Plasmids. The entire coding region of Surf-6 was isolated from different species by PCR amplification using the *pfx* DNA polymerase (Invitrogen) and cloned in-frame to the carboxy-terminal sequence of the EGFP in the pEGFP-C3 vector (CLONTECH) using the *Eco*RI–*Bam*HI cloning sites. PCR fragment was run in an agarose gel and purified by the QIAEX II gel extraction kit (Qiagen) before being digested by restriction enzymes and again before being ligated to the plasmid. Three fusion constructs containing the entire coding region from different species, namely GFP*mSURF6, GFP*drSURF6, and GFP*ySURF6, were engineered by amplifying SURF-6 from a cDNA clone of *Mus musculus* [12], from a cDNA embryonic library of *Drosophila melanogaster*, and from genomic DNA of *Schizosaccharomyces pombe* (provided by Dr. Takashi Toda from Cancer Research, UK), respectively. Similarly, a series of constructs containing various deletions of the Surf-6 gene were engineered by PCR using primers specific to the mouse sequence as shown in Table 1. For one construct, GFP*mSURF6 (1–261/288–355), an internal deletion was achieved by cloning two different PCR fragments of the coding region using an overlapping *Kpn*I restriction site. The correctness of the amplified PCR sequence was confirmed by ABI automated DNA sequence analysis. For all transfections DNA constructs were purified by the plasmid purification kit (Qiagen) and were expressed transiently in P19 embryonic carcinoma cells.

Cell culture and transfection. The embryonic carcinoma P19 cell line was purchased from ATCC. Cells were grown in minimum essential medium (α-MEM) with ribonucleosides and deoxyribonucleosides (Gibco), containing 7.5% bovine calf serum and 2.5% fetal bovine serum (Sigma) at 37 °C and 5% CO_2 incubator. Cells were transfected with the above-described plasmids by using the Lipofectamine2000 reagent (Invitrogen) according to the manufacturer's instruction. Transiently transfected cells were cultured under standard conditions for another 24 h before proceeding to fluorescence microscopy.

Fluorescence imaging analysis. For the analysis in the sub-cellular distribution of EGFP-tagged proteins, transiently transfected cells were fixed with 3% paraformaldehyde in phosphate-buffered saline (PBS) for 10 min at room temperature (RT). After this step cells were analyzed by confocal fluorescence microscopy. For immunocytochemistry, fixed cells were treated with 0.2% Triton X-100 in PBS for 10 min and incubated with a mouse specific SURF-6 antibody as we previously described [12]. After washing with PBS, cells were incubated for 1 h at RT with a secondary antibody, Texas red-conjugated goat-anti-rabbit immunoglobulins at 1:250 dilution (Jackson ImmunoResearch Laboratories). After several rinses by PBS, immunoreactivity was visualized with a confocal laser scanning microscope LSM 510

Table 1
Oligonucleotides used for SURF-6 constructs

Constructs	5′-primers	3′-primers
Species specific constructs	*Eco*RI	*Bam*HI
GFP*mSURF6	5′-cggaattctgatggcttctctcctggccaaggat	5′-gcgggatcctcaggagaggcctgcacgctccag
GFP*drSURF6	5′-cggaattctgatgacccaggcggagatcgtgaaa	5′-gcgggatcctaatagccgggtatgatgcgacc
GFP*ySURF6	5′-cggaattctgatggaaacaactgaaggaa	5′-gcgggatccctatttttcttggatggc
Deletion constructs		
GFP*mSURF6 (1–195)	5′-cggaattctgatggcttctctcctggccaaggat	5′-gcgggatcctcagttaaagatcagccctgattcctg
GFP*mSURF6 (1–99)	5′-cggaattctgatggcttctctcctggccaaggat	5′-gcgggatcctcatttaggggacccccagggagctcaa
GFP*mSURF6 (100–195)	5′-cggaattctggatagccaaggcacagcacgggag	5′-gcgggatcctcagttaaagatcagccctgattcctg
GFP*mSURF6 (198–355)	5′-cggaattctggaagtgactgaagaggagccagcc	5′-gcgggatcctcaggagaggcctgcacgctccag
GFP*mSURF6 (288–355)	5′-cggaattctgaagcgcaaagagaagcgccgggca	5′-gcgggatcctcaggagaggcctgcacgctccag
GFP*mSURF6 (318–355)	5′-cggaattctgaagcggcgccagaacctgcgcaag	5′-gcgggatcctcaggagaggcctgcacgctccag
GFP*mSURF6 (198–318)	5′-cggaattctggaagtgactgaagaggagccagcc	5′-gcgggatcctcaccgcttgtcctgtcgctgctgcat
GFP*mSURF6 (198–287)	5′-cggaattctggaagtgactgaagaggagccagcc	5′-gcgggatcctcacagggcttcctgcagcaggcgttc
GFP*mSURF6 (1–261/288–355)	5′-cggaattctgatggcttctctcctggccaaggat	5′-gcgggatcctcaggagaggcctgcacgctccag
	*Kpn*I	*Kpn*I
	5′-cggggtaccaagcgcaaagagaagcgccgggca	5′-cggggtacccatcttggcctccagctcctgggc

Oligonucleotides used for the amplification of various SURF-6 sequences by PCR to engineer EGFP fusion constructs as described in Materials and methods. Restriction sites used for cloning are underlined.

(Carl Zeiss, Germany) using excitation wavelengths of 450 and 543 nm.

Results and discussion

The evolutionary conserved SURF-6 domain

We identified the SURF-6 domain in a sequence alignment of distantly related species extending from human to yeast, see Fig. 1, by searching the Protein families database. The domain is localized in the carboxy-terminal of all eukaryotic proteins and has an average length of 191 amino acids with an average residue identity of 36% between different species. The amino-terminal of the SURF-6 proteins is variable in length and does not share any significant sequence homology between distantly related species. The conserved domain is very basic, being very rich in arginine/lysine residues, and its predicted secondary structure consists predominantly of α-helices in all species as shown in Fig. 1. Similar secondary structures have been found to be widespread in nucleic acid-binding proteins [25,26]. Although we have previously shown that the mammalian SURF-6 is a nucleolar-matrix protein and has nucleic-acid-binding properties [13], the SURF-6 domain has not been identified to be similar to any known nucleic-acid-binding proteins based on exhaustive bioinformatic analyses of the nucleolar proteome [9,16]. Moreover, the SURF-6 domain is not similar to any known nucleolar domains necessary for protein–protein interaction characteristic to cell-cycle regulators sequestered in the nucleolus [16]. Within the SURF-6 domain there is a highly conservative core of about 60 amino acids containing nine identical residues, between a highly conserved tryptophan and asparagine, in all species (Fig. 1). The core residues can play an important role both in a proper domain folding and in potential molecular interactions of SURF-6 with protein partners and/or nucleic acids.

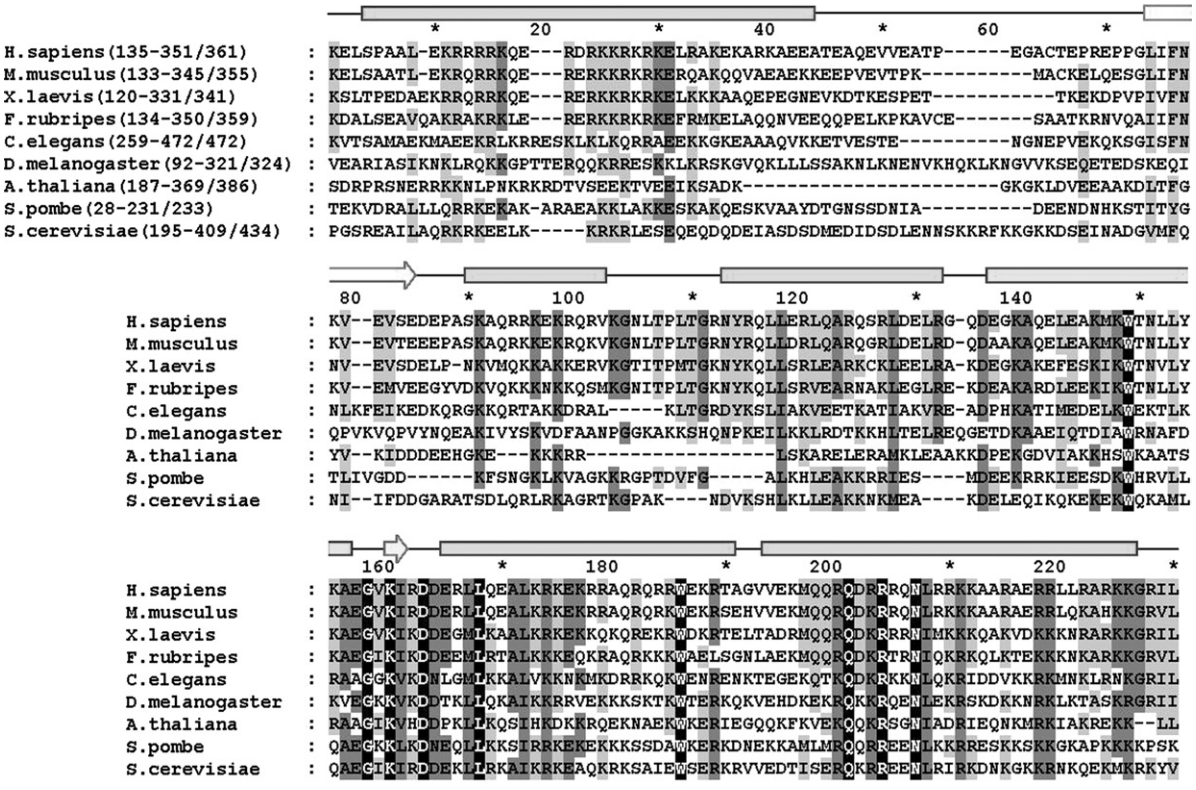

Fig. 1. The evolutionary conserved domain of SURF-6. A significant amino acid identity was found in the carboxy-terminal of the SURF-6 protein between distantly related species as shown in this comparative protein sequence alignment, adapted from the Pfam database (Accession No. PF0495). The region containing the conserved domain and the total length of the protein are indicated in parentheses for each species. Shading mode for the degree of amino acid conservation was performed by GeneDoc [22]: 100% identity of residues within all species—black; less than 100% but more than 70% conserved residues—gray; and less than 70% but more than 50% conserved residues—grayish. The secondary structure for the SURF-6 domain is the same in all species, as predicted by the PSIPRED analysis program [23]; box = α-helix, arrow = β-strand, and line = coil. The GenBank accession numbers of the sequences are: AF186772 (*Homo sapiens*), X92842 (*Mus musculus*), AJ276843 (*Xenopus laevis*), Y15171 (*Fugu rubripes*), U29380 (*Caenorhabditis elegans*), AE003728 (*Drosophila melanogaster*), AY039882 (*Arabidopsis thaliana*), AL021837 (*Schizosaccharomyces pombe*), and Z28082 (*Saccharomyces cerevisiae*). The *S. cerevisiae* SURF-6 homologue is the ykl082c open reading frame encoding the rrp14 protein (www.yeastgenome.org).

Nucleolar localization of SURF-6 proteins is conserved in the eukaryotic kingdom

Then, we examined the cellular localization of proteins containing the SURF-6 domain from distantly related species, covering the entire eukaryotic kingdom, in order to determine whether the SURF-6 domain is related to nucleolar functions in all eukaryotes. For this purpose, we isolated the full coding sequence from genes encoding the SURF-6 domain, as identified in the sequence alignment of Fig. 1, in mouse (*M. musculus*), fruit fly (*D. melanogaster*), and yeast (*S. pombe*).

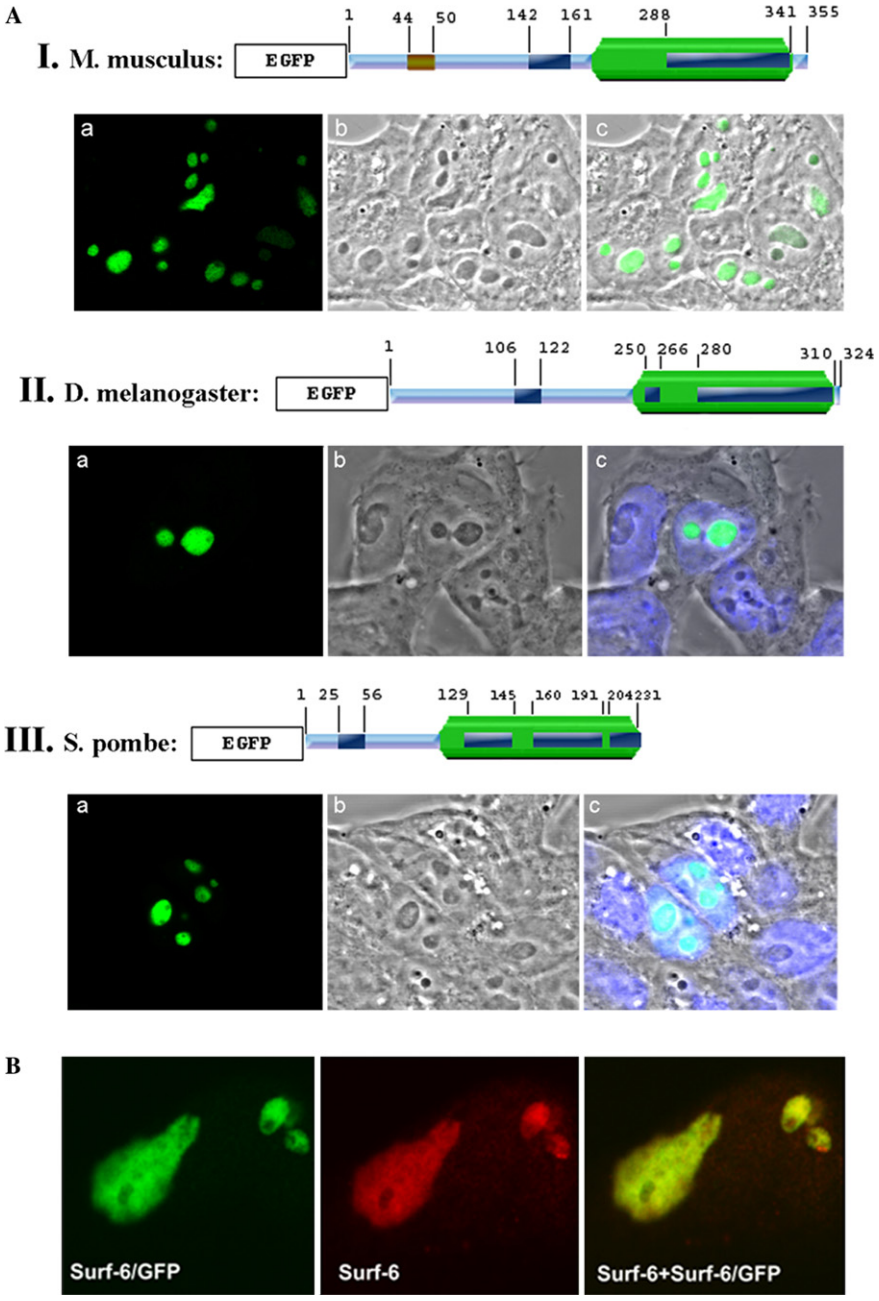

Fig. 2. Nucleolar localization of SURF-6 in eukaryotes. (A) Constructs having EGFP fused with SURF-6 proteins from distantly related eukaryotic taxa, *M. musculus* (panel I), *D. melanogaster* (panel II), and *S. pombe* (panel III), are shown schematically. The SURF-6 domain located at the carboxy terminal of the proteins is presented as a large gray box. Also, small shaded boxes show the location of identified bipartite and monopartite nuclear localization sequences described in Table 2. Constructs were transiently expressed in mouse P19 embryonic carcinoma cells and EGFP localization was observed by fluorescence microscopy (a) and was co-localized in the nucleolus (c) as compared to the phase contrast images (b). (B) Nucleolar co-localization of the EGFP*mouse SURF-6 fusion protein, Surf-6/GFP (green), with endogenous mouse SURF-6, Surf-6 (red), in cells as shown in enlarged images of nucleoli. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

Isolated sequences from each species were fused with the enhanced variant of the green fluorescent protein (EGFP), as described in Materials and methods, in order to assay their cellular localization by fluorescent microscopy (see Fig. 2A). We had previously shown by immuno-cytochemistry that the mouse SURF-6 is localized in the nucleolus and particularly in the granular component, a structure which is involved in ribosome maturation [13,12]. Therefore, we first examined whether our EGFP fusion experiments reflect a characteristic cellular localization by co-localizing the green fluorescence from the mouse SURF-6*EGFP fusion with the endogenous SURF-6 antigen, recognized by a mouse antibody, in mouse P19 embryonic carcinoma cells (Fig. 2B). It is clearly shown in magnified nucleoli that GFP is targeted to the nucleolus and it is co-localized with the endogenous SURF-6 protein. Similarly, it is found that SURF-6 sequences from the other distantly related species, i.e., fruit fly and yeast, target GFP into the nucleolus (Fig. 2A, panels II and III). These results point out that all eukaryotic SURF-6 proteins share nucleolar functions and belong to a nucleolar protein family. Moreover, they suggest that the SURF-6 proteins may contain evolutionary conserved signals for nucleolar targeting.

Identification of nucleolar localization signals in SURF-6

We engineered a series of plasmid constructs having EGFP tagged with various deletions of the mouse SURF-6 sequence in order to determine the regions responsible for nucleolar targeting (see schematic presentation of constructs in Fig. 3). We tested the cellular localization of ten different EGFP-tagged constructs and found that strikingly all constructs but one yielded fusion proteins with a nucleolar localization capacity. The strength of the nucleolar localization, however, was different for some constructs. Specifically two constructs, plus the control construct containing the entire SURF-6 sequence, gave a strong nucleolar localization, i.e., localized exclusively in the nucleolus (Figs. 3B and J). One of these constructs contains only the N-amino terminal of the SURF-6 protein, which does not have any sequences of the evolutionary conserved domain, and the other construct has part of the core of the conserved domain deleted (i.e., the region with the identical residues found in all species), indicating that the highly conserved residues in the domain are not necessary for nucleolar localization. These results were confirmed reversely by showing that the deleted core region of the domain alone does not have any nucleolar localization properties, as it is localized

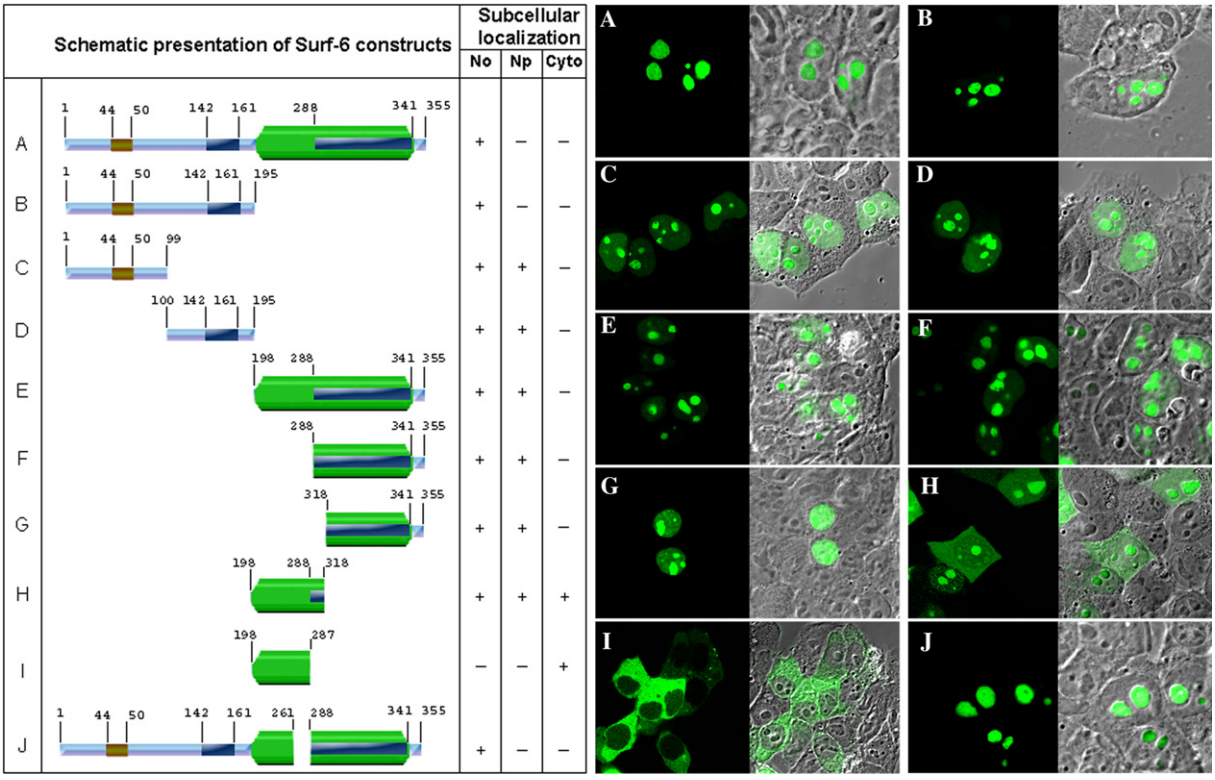

Fig. 3. Examining for nucleolar localization signals in SURF-6. Constructs having deleted sequences of the mouse SURF-6 protein fused with EGFP are shown schematically. The engineering of these constructs is described in Materials and methods. The cellular localization of the construct having the entire sequence of the mouse SURF-6 protein is described in Fig. 1. The scoring of subcellular localization of the EGFP-tagged proteins (+ or −) for nucleolus (No), nucleus (Np), and cytoplasm (Cyto) is based on the shown images.

Table 2
Identification of multiple nuclear localization sequences in the SURF-6 protein

M. musculus	*D. melanogaster*	*S. pombe*
		25–41: Rkatekvdralllqrrk
44–50: kkkrk	106–122: KKgptterqqkrreskk	39–55: RRkekakaraeakklak
142–158: KRqrrkqererkkrkrk	250–266: KKvkddtkllqkaikkr	40–56: RKekakaraeakklakk
145–161: RRkqererkkrkrkerq	280–296: RKqkvehdkekrqkkrq	129–145: KKrriesmdeekrrkie
288–304: KRkekrraqrqrkwekr	290–306: KRqkkrqenlekrskdk	160–176: KKlkdneqllkksirrk
303–319: KRsehvvekmqqrqdkr	293–309: KKrqenlekrskdkknr	175–191: RKekekkkssdawkerk
318–334: KRrqnlrkkkaaraer	294–310: KRqenlekrskdkknrk	204–220: RReenlkkrreskkskk
325–341: KKkaaraerrlqkahkk		210–226: KKrreskkskkgkapkk
		211–227: KRreskkskkgkapkkk
		212–228: RReskkskkgkapkkkk
		216–232: KKskkgkapkkkkpskk

Bipartite and monopartite nuclear localization sequences were predicted by the PROSITE motif analysis program [24].

exclusively in the cytoplasm (Fig. 3I). Two different subregions of the N-amino terminal gave a less strong nucleolar localization, i.e., localized in both the nucleolus and nucleus (Figs. 3C and D). Similarly, we found that also the carboxy-terminal of the protein, which contains the evolutionary conserved domain, has different regions containing weaker signals for a nucleolar/nuclear localization (Figs. 3E, F, G, and H). These results indicate that multiple regions of the SURF-6 protein have the capacity for nucleolar localization. We identified, by using the PROSITE motif analysis program [24], that SURF-6 contains multiple nuclear localization signals (NLS) within the regions localized in the nucleolus (see Table 2). Nucleolar localization signals (NoLS), also known as nucleolar targeting signal (NOS), are regarded to be clustered basic amino acids resembling NLS, usually forming bipartite NLS (bNLS) [27]. In some proteins, however, the correct nucleolar localization is completed by a functional domain [28,29]. Such a requirement for nucleolar localization is characteristic to the major nucleolar proteins nucleolin [30], B23 [31], and fibrillarin [32]. In SURF-6, nucleolar localization appears to be achieved by sequences containing bNLS and without the putative functional conserved domain. The strongest nucleolar localization is determined by the N-amino terminal region that contains bNLS being apart from each other, indicating that the degree of nucleolar localization is determined by the spacing of bNLS rather than by their accumulative effect. The same distribution of multiple bNLS is found in SURF-6 proteins from all species, examples from three evolutionary distant species are shown in Table 2, indicating that the signaling mechanisms for nucleolar localization are evolutionary conserved between mammalian and yeast cells. It is interesting that some sequences within the SURF-6 domain are also capable for nucleolar targeting. This is in agreement with the observation that domains rich in arginine and lysine contain bipartite nuclear localization signals which can overlap with DNA- and RNA-binding properties [33].

The identified SURF-6 domain is a candidate for nucleic-acid binding, possibly for RNA binding as the mammalian SURF-6 protein has stronger binding affinity for RNA than DNA in vitro and is localized in the granular component of the nucleolus where rRNA processing occurs [12]. However, the structure of the SURF-6 domain is unique as it does not belong to any of the known families of RNA-binding proteins [16,34,35]. Alternatively, the SURF-6 domain may present a novel domain for protein–protein interactions since the yeast homologue of SURF-6, ykl082c/rrp14, has been found to interact with proteins necessary for ribosomal biogenesis [20] and cell polarity [17,18]. Further functional genomic and proteomic studies on yeast have illuminated the potential functions of the mammalian SURF-6. It was found from knock-out yeast mutants that the ykl082c gene is necessary for viability [36], and in a similar genetic screening ykl082c has been excluded as being a candidate for pre-rRNA splicing [37]. Rather, the ykl082 protein is found to interact with proteins from the proteome of the pre-ribosomal 60 S subunit implicating that it is a pre-ribosomal assembly factor [19–21]. From these yeast studies, it emerges that rrp14 acts as a chaperone between ribosomal biosynthesis and cell polarity. We can speculate that the mammalian SURF-6 has analogous functions implicating that it may link ribosome biogenesis and cell-cycle processes in metazoa. Such interactions in the mammalian cell remain to be elucidated, however, this study can help us to direct our future research on particular molecular processes in order to define the importance of the SURF-6 protein structure in cellular function.

Acknowledgments

We are grateful to thank M.C. Liu and K.S. Mukhar'iamova for help with confocal microscopy imaging. M. Polzikov has been supported by INTAS and the European Science Foundation. This work was partially supported by the Russian Foundation for Basic Research (Grant No. 04-04-48391).

References

[1] T. Melese, Z. Xue, The nucleolus: an organelle formed by the act of building a ribosome, Curr. Opin. Cell Biol. 7 (1995) 319–324.

[2] M.O. Olson, K. Hingorani, A. Szebeni, Conventional and nonconventional roles of the nucleolus, Int. Rev. Cytol. 219 (2002) 199–266.

[3] J.B. Bachant, S.J. Elledge, Mitotic treasures in the nucleolus, Nature 398 (1999) 757–758.

[4] M. Carmo-Fonseca, L. Mendes-Soares, I. Campos, To be or not to be in the nucleolus, Nat. Cell Biol. 2 (2000) E107–E112.

[5] C.L. Sansam, K.S. Wells, R.B. Emeson, Modulation of RNA editing by functional nucleolar sequestration of ADAR2, Proc. Natl. Acad. Sci. USA 100 (2003) 14018–14023.

[6] M.T. Teixeira, K. Forstemann, S.M. Gasser, J. Lingner, Intracellular trafficking of yeast telomerase components, EMBO Rep. 3 (2002) 652–659.

[7] R. Bernardi, P.P. Pandolfi, The nucleolus: at the stem of immortality, Nat. Med. 9 (2003) 24–25.

[8] J.S. Andersen, C.E. Lyon, A.H. Fox, A.K. Leung, Y.W. Lam, H. Steen, M. Mann, A.I. Lamond, Directed proteomic analysis of the human nucleolus, Curr. Biol. 12 (2002) 1–11.

[9] A.K. Leung, J.S. Andersen, M. Mann, A.I. Lamond, Bioinformatic analysis of the nucleolus, Biochem. J. 376 (2003) 553–569.

[10] A. Scherl, Y. Coute, C. Deon, A. Calle, K. Kindbeiter, J.C. Sanchez, A. Greco, D. Hochstrasser, J.J. Diaz, Functional proteomic analysis of human nucleolus, Mol. Biol. Cell. 13 (2002) 4100–4109.

[11] A.K. Leung, A.I. Lamond, The dynamics of the nucleolus, Crit. Rev. Eukaryot. Gene Expr. 13 (2003) 39–54.

[12] C. Magoulas, M. Fried, The Surf-6 gene of the mouse surfeit locus encodes a novel nucleolar protein, DNA Cell Biol. 15 (1996) 305–316.

[13] C. Magoulas, O.V. Zatsepina, P.W. Jordan, E.G. Jordan, M. Fried, The SURF-6 protein is a component of the nucleolar matrix and has a high binding capacity for nucleic acids in vitro, Eur. J. Cell Biol. 75 (1998) 174–183.

[14] C. Magoulas, M. Fried, Isolation and genomic analysis of the human surf-6 gene: a member of the Surfeit locus, Gene 243 (2000) 115–123.

[15] A. Bateman, L. Coin, R. Durbin, R.D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E.L. Sonnhammer, D.J. Studholme, C. Yeats, S.R. Eddy, The Pfam protein families database, Nucleic Acids Res. 32 (Database issue) (2004) D138–D141.

[16] E. Staub, P. Fiziev, A. Rosenthal, B. Hinzmann, Insights into the evolution of the nucleolus by an analysis of its protein domain repertoire, Bioessays 26 (2004) 567–581.

[17] B.L. Drees, B. Sundin, E. Brazeau, J.P. Caviston, G.C. Chen, W. Guo, K.G. Kozminski, M.W. Lau, J.J. Moskow, A. Tong, L.R. Schenkman, A. McKenzie III, P. Brennwald, M. Longtine, E. Bi, C. Chan, P. Novick, C. Boone, J.R. Pringle, T.N. Davis, S. Fields, D.G. Drubin, A protein interaction map for cell polarity development, J. Cell Biol. 154 (2001) 549–571.

[18] P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J.M. Rothberg, A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*, Nature 403 (2000) 623–627.

[19] A. Fatica, A.D. Cronshaw, M. Dlakic, D. Tollervey, Ssf1p prevents premature processing of an early pre-60S ribosomal particle, Mol. Cell 9 (2002) 341–351.

[20] E.W. Horsey, J. Jakovljevic, T.D. Miles, P. Harnpicharnchai, J.L. Woolford Jr., Role of the yeast Rrp1 protein in the dynamics of pre-ribosome maturation, RNA 10 (2004) 813–827.

[21] C. Shirai, T. Takai, M. Nariai, C. Horigome, K. Mizuta, Ebp2p, the yeast homolog of Epstein–Barr virus nuclear antigen 1-binding protein 2, interacts with factors of both the 60 S and the 40 s ribosomal subunit assembly, J. Biol. Chem. 279 (2004) 25353–25358.

[22] K.B. Nicholas, H.B. Nicholas Jr., D.W. Deerfield, GeneDoc: analysis and visualization of genetic variation, EMBNEW. NEWS 4 (1997) 14–15.

[23] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, J. Mol. Biol. 292 (1999) 195–202.

[24] C.J. Sigrist, L. Cerutti, N. Hulo, A. Gattiker, L. Falquet, M. Pagni, A. Bairoch, P. Bucher, PROSITE: a documented database using patterns and profiles as motif descriptors, Brief. Bioinform. 3 (2002) 265–274.

[25] N.M. Luscombe, J.M. Thornton, Protein–DNA interactions: amino acid conservation and the effects of mutations on binding specificity, J. Mol. Biol. 320 (2002) 991–1009.

[26] A.G. Murzin, S.E. Brenner, T. Hubbard, C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, J. Mol. Biol. 247 (1995) 536–540.

[27] M. Hatanaka, Discovery of the nucleolar targeting signal, Bioessays 12 (1990) 143–148.

[28] J.D. Weber, M.L. Kuo, B. Bothner, E.L. DiGiammarino, R.W. Kriwacki, M.F. Roussel, C.J. Sherr, Cooperative signals governing ARF-mdm2 interaction and nucleolar localization of the complex, Mol. Cell. Biol. 20 (2000) 2517–2528.

[29] K.T. Etheridge, S.S. Banik, B.N. Armbruster, Y. Zhu, R.M. Terns, M.P. Terns, C.M. Counter, The nucleolar localization domain of the catalytic subunit of human telomerase, J. Biol. Chem. 277 (2002) 24764–24770.

[30] M.S. Schmidt-Zachmann, E.A. Nigg, Protein localization to the nucleolus: a search for targeting domains in nucleolin, J. Cell Sci. 105 (Pt 3) (1993) 799–806.

[31] Y. Nishimura, T. Ohkubo, Y. Furuichi, H. Umekawa, Tryptophans 286 and 288 in the C-terminal region of protein B23.1 are important for its nucleolar localization, Biosci. Biotechnol. Biochem. 66 (2002) 2239–2242.

[32] S. Snaar, K. Wiesmeijer, A.G. Jochemsen, H.J. Tanke, R.W. Dirks, Mutational analysis of fibrillarin and its mobility in living human cells, J. Cell Biol. 151 (2000) 653–662.

[33] E.C. LaCasse, Y.A. Lefebvre, Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins, Nucleic Acids Res. 23 (1995) 1647–1656.

[34] A.C. Messias, M. Sattler, Structural basis of single-stranded RNA recognition, Acc. Chem. Res. 37 (2004) 279–287.

[35] J.M. Perez-Canadillas, G. Varani, Recent advances in RNA-protein recognition, Curr. Opin. Struct. Biol. 11 (2001) 53–58.

[36] E.A. Winzeler, D.D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J.D. Boeke, H. Bussey, A.M. Chu, C. Connelly, K. Davis, F. Dietrich, S.W. Dow, M. El Bakkoury, F. Foury, S.H. Friend, E. Gentalen, G. Giaever, J.H. Hegemann, T. Jones, M. Laub, H. Liao, R.W. Davis, Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis, Science 285 (1999) 901–906.

[37] W.T. Peng, M.D. Robinson, S. Mnaimneh, N.J. Krogan, G. Cagney, Q. Morris, A.P. Davierwala, J. Grigull, X. Yang, W. Zhang, N. Mitsakakis, O.W. Ryan, N. Datta, V. Jojic, C. Pal, V. Canadien, D. Richards, B. Beattie, L.F. Wu, S.J. Altschuler, S. Roweis, B.J. Frey, A. Emili, J.F. Greenblatt, T.R. Hughes, A panoramic view of yeast noncoding RNA processing, Cell 113 (2003) 919–933.