Nucleolar Binding Sequences of the Ribosomal Protein S6e Family Reside in Evolutionary Highly Conserved Peptide Clusters

Swarupa Kundu-Michalik, Marc-Angelo Bisotti, Edgar Lipsius, Andreas Bauche, Antonina Kruppa, Thomas Klokow, Gertrud Kammler,1 and Joachim Kruppa

Center of Experimental Medicine, Institute of Molecular Cell Biology, Hamburg University, D-20246 Hamburg, Germany

Proteomic analyses of the nucleolus have revealed almost 700 functionally diverse proteins implicated in ribosome biogenesis, nucleolar assembly, and regulation of vital cellular processes. However, this nucleolar inventory has not unveiled a specific consensus motif necessary for nucleolar binding. The ribosomal protein family characterized by their basic nature should exhibit distinct binding sequences that enable interactions with the rRNA precursor molecules facilitating subunit assembly. We succeeded in delineating 2 minimal nucleolar binding sequences of human ribosomal protein S6 by fusing S6 cDNA fragments to the 5′ end of the LacZ gene and subsequently detecting the intracellular localization of the β-galactosidase fusion proteins. Nobis1 (nucleolar binding sequence 1), comprising of 4 highly conserved amino acid clusters separated by glycine or proline, functions independently of the 3 authentic nuclear localization signals (NLSs). Nobis2 consists of 2 conserved peptide clusters and requires the authentic NLS2 in its native context. Similarly, we deduced from previous publications that the single Nobis of ribosomal protein S25 is also highly conserved. The functional protein domain organization of the ribosomal protein S6e family consists of 3 modules: NLS, Nobis, and the C-terminal serine cluster of the phosphorylation sites. This modular structure is evolutionarily conserved in vertebrates, invertebrates, and fungi. Remarkably, nucleolar binding sequences of small and large ribosomal proteins reside in peptide clusters conserved over millions of years.

Introduction

The nucleolus is a highly dynamic nuclear structure (Andersen et al. 2005) formed around tandemly repeated genes of rDNA coding for preribosomal RNA (Hadjiloy 1985). The major function of the nucleolus is the assembly of ribosomal subunits (Raska et al. 2006). During ribosome biogenesis, the rRNA precursor is transcribed at the border between the fibrillar center and dense fibrillar component (DFC), and in the DFC (Koberna et al. 2002). Ribosomal RNA processing and base modifications as well as the association of rRNA with ribosomal proteins give rise to intermediates of preribosomal particles (Grandi et al. 2002; Tschochner and Hurt 2003). These progress vectorially through functional subdomains of the DFC into the granular component (GC) for further maturation and completion of ribosomal subunit assembly (Scheer and Hock 1999; Gerbi et al. 2003). A myriad of small nucleolar ribonucleoprotein particles (RNPs) assist the processing and modification of the primary rRNA transcript (Fromont-Racine et al. 2003; Granneman et al. 2004) so that ultimately mature 5.8S, 18S, and 25S rRNA molecules are produced in Saccharomyces cerevisiae (Venema and Tollervey 1999).

The advent of proteomic analyses of the nucleolus (Andersen et al. 2002; Scherl et al. 2002) as well as of preribosomal particles (Takahashi et al. 2003) have revealed countless trans-acting factors participating in the complex pattern of preribosome maturation (Milkereit et al. 2003), which is accompanied by a dynamic change in the composition and subnuclear distribution of ribosomal synthesis factors (Leung et al. 2003; Hinsby et al. 2006). Fluorescence recovery after photobleaching analyses have demonstrated that some nucleolar components involved in ribosome biogenesis rapidly shuttle between the nucleolus and the nucleoplasm (Chen and Huang 2001).

Ribosomes are complexes essential for cell proliferation and cell growth present in all organisms (Mayer and Grummt 2006; Bernstein et al. 2007). In bacteria, the catalytically active rRNA forms the framework upon which 55 ribosomal proteins are assembled, thereby stabilizing the 3-dimensional structure of the ribosomal subunits (Schuwirth et al. 2005). Analogous to all other nuclear proteins, eukaryotic ribosomal proteins are synthesized in the cytoplasm, subsequently imported through the nuclear pore complex into the nucleoplasm, and ultimately accumulate in the nucleolus (Fromont-Racine et al. 2003). Their association with the newly transcribed pre-rRNA generates a 90S preribosomal particle that is processed in a highly complex pattern into 40S and 60S ribosomal subunits, which are exported through the nuclear pore complex into the cytoplasm (Grandi et al. 2002; Fromont-Racine et al. 2003; Tschochner and Hurt 2003). In mammalian cells, nuclear import of ribosomal proteins (Schmidt et al. 1995) is mediated by classical nuclear localization signals (NLSs) of either monopartite or bipartite nature (Dingwall and Laskey 1991; Lipsius et al. 2005) and assisted by a variety of import factors (Jäkel and Görlich 1998).

A specific mechanism for targeting proteins into the nucleolus is not needed because the nucleolus is not surrounded by a lipid bilayer (Carro-Fonseca et al. 2000). Furthermore, analysis of the nucleolar proteome has demonstrated that a simple universal targeting motif shared by all nucleolar proteins does not exist (Andersen et al. 2002; Scherl et al. 2002) suggesting that they are instead sequenced and retained by binding to other nucleolar components (Carro-Fonseca et al. 2000). Within proteins of retroviruses (Hatanaka 1990), short nucleolar targeting signals of 12–20 amino acids can target a reporter protein from the cytoplasm into the nucleolus.

It is assumed that the ribosome arose on a single occasion as a particle composed completely of RNA capable
of catalyzing peptide bond formation (Noller 1991; Wool et al. 1995). Numerous established rRNA sequences of different phyla and species support this assumption by providing compelling evidence for the homology of rRNAs of all species, which is more apparent in the comparison of secondary structure than of nucleotide sequences. In the 1970’s, Wool and Stöffler (1974) were already convinced that ribosomal protein sequences from distant species are also evolutionary related. Nowadays, genomic sequence comparisons have confirmed the existence of homologous proteins in ribosomes of Archaea, Bacteria, and Eukarya (Lecompte et al. 2002; Mears et al. 2002). Presumably, evolutionary highly conserved clusters in ribosomal protein sequences fulfill a crucial role in ribosome assembly and function.

Research on the biogenesis of ribosomal subunits concentrating on rRNA transcription, processing, and maturation has neglected the influence of ribosomal proteins in the assembly process. However, in 2005, Milkereit’s group (Ferreira-Cerca et al. 2005) demonstrated that eukaryotic ribosomal proteins play vital roles in the maturation and transport of the pre-18S rRNA. Implyed by this result is the existence of specific amino acid sequences within ribosomal proteins, which optimally interact with their corresponding nucleotide sequences so that distinct pre-rRNA processing steps can occur (Ferreira-Cerca et al. 2005). Unfortunately, the exact sequences and structural features of such binding domains in eukaryotic ribosomal proteins have not yet been systematically investigated. Therefore, we characterized the topogenic sequences of human ribosomal protein S6 in detail by delineating the regions required for nuclear targeting and nucleolar binding in order to understand the organization and interrelationship of these functional domains (Fukami-Kobayashi et al. 2007). To this end, we demonstrate that S6 has 2 nucleolar binding sequences — Nobis1 (nucleolar binding sequence 1) in the center functioning independently of the 3 authentic NLSs and Nobis2 in the C-terminal region depending on the presence of NLS2—both of which reside in regions of highly conserved peptide clusters. Furthermore, we show that S6 seems to be a bipartite import signal.

### Materials and Methods

#### Constructs for the Delineation of Nobis1 in the Central Region of S6

A set of deletion mutants was constructed to identify a Nobis1 in the central region of S6. To this end, 2 polymerase chain reaction (PCR) fragments were designed using the primers S6 (A1) and S6 (Nhe170) as well as S6 (Nhe210) and S6 (H249) (table 1). The PCR fragments were digested with AflII/NheI and NheI/HindIII, respectively, and cloned together into the pASH vector (Lipsius et al. 2005), which had been restricted with AflII/HindIII. Since the PCR primer S6 (Nhe170) also hybridized to the S6 DNA sequence at the nucleotides 414–420, an additional S6 fragment corresponding to amino acids 1–139 was synthesized so that a second deletion mutant S6(1–249,Δ140–210) was obtained. In order to narrow down Nobis1, additional constructs were designed beginning at amino acid residues 53, 64, 72, 75, 100, and 115, respectively, by employing the indicated primers (table 1) and the S6(1–249,Δ140–210) construct as a template.

Additional constructs with consecutively shorter DNA segments were prepared containing the exogenous SV40 NLS. The DNA fragments were generated by PCR using the plasmid pBluescript KS(+) as a template.

### Site-Directed Mutagenesis of NLS3 in the Fragment S6(215–240) and in the Tetrapeptide

Arginine 231 of NLS3 was mutated into valine, glycine, and glutamic acid. Hybrids were prepared from synthetic oligonucleotides as summarized in table 2 to obtain the tetrapeptides for fusion to β-galactosidase. These hybrids were ligated into the vector pKHlacZ (Annino et al. 1998), which had been restricted with NotI and HindIII. The correct constructs of the clones obtained were verified by sequencing. Plasmid DNA was used for transfections as previously described (Schmidt et al. 1995).

The fragment S6(215–240) was subcloned into the vector M13mp19 and an oligonucleotide-dependent in vitro mutagenesis was carried out as described by the manufacturer (Amersham-Buchler, Braunschweig, Germany). A set of oligonucleotides was synthesized by using a mixture of A, C, G, and T at position N in the sequence: 5'-ATTGGC-GAAGGNCAGCGACTT-3'. Competent Escherichia coli XL1-Blue cells were transformed with the mutated

### Table 1

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence in 5′ → 3′</th>
</tr>
</thead>
<tbody>
<tr>
<td>S6 (A1)</td>
<td>GAAAGGATCCCCTTTAGATGAAAGCTGCAAGCTCACATTC</td>
</tr>
<tr>
<td>S6 (A55)</td>
<td>TTGGATCTTACCAGGACTGTCGTTGGGAAAGCAGCAAAACAG</td>
</tr>
<tr>
<td>S6 (A64)</td>
<td>TTGGATCTTACCAAGAGCCGGGTCTGGATGCAAC</td>
</tr>
<tr>
<td>S6 (A72)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A75)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A115)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A170)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A173)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A180)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (A192)</td>
<td>TTGGATCTTACCAAGGTCGGTGCTGACAGTCAGT</td>
</tr>
<tr>
<td>S6 (H94)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H96)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H102)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H120)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H203)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H210)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H229)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (H249)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (Nhe170)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
<tr>
<td>S6 (Nhe210)</td>
<td>GGAACGGGACCTCGTTCTTCTTGCAAGT</td>
</tr>
</tbody>
</table>

Note.—The orientation in which the PCR primers have been used is indicated, and the restriction sites are underlined.
Table 2
Oligonucleotides Used for Site-Directed Mutagenesis and Hybrid Formation

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLS3 wt</td>
<td>5’-GGCCGCATAAGAGAAGAATA-3’</td>
</tr>
<tr>
<td>NLS3 wt</td>
<td>3’-CGTTTCTCTCTTTATTGCAGA-5’</td>
</tr>
<tr>
<td>R230G</td>
<td>5’-GGGGCGCAGGAAGAAGAATA-3’</td>
</tr>
<tr>
<td>R231G</td>
<td>3’-CGTTTCTCTCTTTATTGCAGA-5’</td>
</tr>
<tr>
<td>R231V</td>
<td>5’-GGGGCGCAGGAAGAAGAATA-3’</td>
</tr>
<tr>
<td>R231V</td>
<td>3’-CGTTTCTCTCTTTATTGCAGA-5’</td>
</tr>
<tr>
<td>R231E</td>
<td>5’-GGGGCGCAGGAAGAAGAATA-3’</td>
</tr>
<tr>
<td>R231E</td>
<td>3’-CGTTTCTCTCTTTATTGCAGA-5’</td>
</tr>
<tr>
<td>R215/216</td>
<td>5’-CCAGTGAATTCGCGCATGAGGGAGGTAAAG-3’</td>
</tr>
<tr>
<td>Δ223/224</td>
<td>5’-AAGGAGGCTAAGGAGCAAGAACAATTCGG-3’</td>
</tr>
<tr>
<td>T-Ag1</td>
<td>5’-GGCCGGCGGAAATATACCGAAAGATGGA-3’</td>
</tr>
<tr>
<td>T-Ag2</td>
<td>3’-CGGTCTTTTTTTTGGGTTCACCCTAA-5’</td>
</tr>
</tbody>
</table>

Note.—Oligonucleotides were purchased from MWG Biotech AG (Ebersberg, Germany) and used either for site-directed mutagenesis or hybrid formation. The nucleotides in bold represent codons of site-directed mutagenesis.

M13mp19 RF DNA that contained mutated S6 fragments. The clones were isolated and sequenced. Only 3 of the 4 theoretically possible mutations were obtained, that is, valine, glycine, and glutamic acid but not alanine. The individual S6 cDNA fragments were isolated from the corresponding M13 constructs, sequenced, and subcloned into pBluescript II/XL1-Blue cells, the plasmid was isolated and purified. The fragment containing the SV40 NLS tag and the S6 sequence was excised by NotI and HindIII and ligated into the corresponding sites of pKHLacZ in order to fuse the S6 construct to the 5’ end of the β-galactosidase gene. The hybrid gene was in the correct reading frame of the ATG codon of the pKHLacZ vector as was verified by sequencing.

To produce the SV40 NLS–tagged S6 constructs, the oligonucleotides T-Ag1 and T-Ag2 (table 2) were hybridized giving rise to a double-stranded DNA fragment with a NotI site at the 5’ end and an EcoRI site at the 3’ end. The SV40 NLS tag was inserted into NotI and EcoRI sites of the plasmid KS-S6ANLS1-3, which contained the S6(3–240) fragment with the triple NLS deletion in its EcoRI and HindIII sites. The correct frame of the tag insertion was verified by sequencing. After transformation of E. coli XL1-Blue cells, the plasmid was isolated and purified. The sequence was verified. After transformation of E. coli XL1-Blue cells, the plasmid was isolated and purified.

The deletion mutants were constructed as follows: the fragment S6(215–240) containing the valine mutation was excised from the corresponding pBluescript II/XL1-Blue plasmid and ligated into pUC19. To construct the deletion mutants S6(215–240, Δ215/216) and S6(215–240, Δ223/224), the primers in table 2 were used in the site-directed mutagenesis. The mutants were selected after a Scal restriction in the E. coli strain BMH71-18 mutS. After a second Scal restriction, the isolated plasmids were transformed into E. coli JM105 cells, and their correct sequence was verified.

A NotI linker had to be added to the excised S6 fragment after filling up the cohesive ends with the Klenow fragment of DNA polymerase I. The modified constructs were subcloned into the NotI/HindIII sites of the vector pKHLacZ where they were in frame with β-galactosidase. After ligation, the correct frame was verified by sequencing.

Construction of S6 Fragments for the Nucleolar Binding Domain (Nobis1) Encompassing NLS2

Starting from the S6 cDNA, several constructs containing a single NLS, either NLS2 or NLS3, were designed and synthesized by PCR using the primers indicated in table 1. All the fragments were purified by the QIAEX II gel extraction kit (Qiagen, Hilden, Germany) before restriction digestion with AflII and HindIII and then subcloned into the corresponding sites of the expression vector pASH. The final expression constructs were verified by sequencing.

Results

Delineation of Nobis1 in the Central Region of S6

In prokaryotes, early ribosomal binding proteins typically fold into globular structures that bind to multistem junctions, thereby folding and fixing the conformation of tRNA helices in their vicinity (Brodersen et al. 2002). Human ribosomal protein S6 is categorized as a primary RNA-binding protein because it associates with early
precursors to the 18S rRNA in the nucleolus (Bernstein et al. 2004; Ferreira-Cerca et al. 2005). Comparison of eu-
karyotic S6 amino acid sequences revealed a high sequence
conservation between amino acids 54 and 154 implying that
this region plays a vital role in the 40S subunit assembly
and/or ribosomal function.

The β-galactosidase fusion construct containing the
fragment S6(3–183) was inefficiently targeted to the nu-
ucleus, which was attributed to the weak nature of NLS1
(Schmidt et al. 1995). Appending the C-terminal segment,
S6(210–249) containing the stronger NLS3 to the con-
structs greatly increased the efficiency of nuclear import
of the corresponding fusion proteins and made it possible
to examine the amino-terminal region of the S6 protein
(1–170) for its nucleolar binding capability (fig. 1; supple-
mental fig. S1, Supplementary Material online for a color
version). The first construct, S6(1–249, Δ171–209),
directed the β-galactosidase reporter into the nucleolus
implying that the amino-terminal half of S6 contains a nu-
cleolar binding domain (fig. 1).

A rough estimate of the location of this Nobis was ob-
tained using the construct S6(1–249, Δ140–209) fused to
β-galactosidase, which accumulated exclusively in the nu-
cleolus (fig. 1). A series of amino-terminal deletions of S6
was studied to narrow down the region capable of nucleolar
binding: 115–139, 100–139, 75–139, 64–139, and 53–139.
Only the β-galactosidase fusion proteins containing
S6(53–249, Δ140–209) (data not shown) and S6(64–249,
Δ140–209) (fig. 1) accumulated in the nucleolus, whereas
all shorter fragments starting with S6(75–249, Δ140–209)
were merely imported into the nucleoplasm (fig. 1). The
amino-terminal border of Nobis1 was clearly defined by
the construct S6(72–249, Δ140–209) that included the
highly conserved tripeptide RVR as an N-terminal exten-
sion of the S6(75–249, Δ140–209) fragment (fig. 1 and sup-
plementary fig. S3, Supplementary Material online).

To further narrow down Nobis1 by defining its C-
terminal border, additional constructs containing the
exogenous SV40 NLS were generated by consecutively
shortening the C-terminus of the S6(72–139) fragment
(fig. 1). The fragments S6(72–120) and S6(72–102) gave
rise to a succinct nucleolar staining, whereas the
S6(72–98) fragment showed, in addition to a distinct nucle-
olar accumulation, a fainter staining of the nucleoplasm.
Shortening by 4 amino acids at the C-terminus produced
fragment S6(72–94), which was completely excluded from
the nucleolus recognizable by the dark spots appearing in
the stained nucleoplasm (fig. 1).

<table>
<thead>
<tr>
<th>Construct</th>
<th>Nobis Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>S6(1–249, Δ171–209)</td>
<td>No</td>
</tr>
<tr>
<td>S6(1–249, Δ140–209)</td>
<td>No</td>
</tr>
<tr>
<td>S6(64–249, Δ140–209)</td>
<td>No</td>
</tr>
<tr>
<td>S6(75–249, Δ140–209)</td>
<td>N</td>
</tr>
<tr>
<td>S6(72–249, Δ140–209)</td>
<td>No</td>
</tr>
<tr>
<td>SV40-S6(72–120)</td>
<td>No</td>
</tr>
<tr>
<td>SV40-S6(72–102)</td>
<td>No</td>
</tr>
<tr>
<td>SV40-S6(72–98)</td>
<td>No</td>
</tr>
<tr>
<td>SV40-S6(72–94)</td>
<td>N</td>
</tr>
</tbody>
</table>

**Fig. 1.** Nobis1 in the central region of S6. A set of deletion mutants starting from S6(1–170) was constructed by employing the endogenous
bipartite NLS3 to identify the minimal nucleolar binding sequence, Nobis1. Decreasingly shorter N-termini of the fragment S6(1–139) were tested for
their nucleolar binding ability. A second set of constructs with an exogenous SV40 NLS attached to the N-terminus was generated by PCR, in which the
C-terminus was consecutively shortened. All fusion proteins were expressed in COS cells. N, nuclear; No, nucleolar. Scale bar, 5 μm.
In summary, the 27 amino acids of S6(72–98) representing Nobis1 have the capability to lead to an accumulation of the fusion protein in the nucleolus. This S6 peptide sequence most probably interacts with the ribosomal 18S rRNA, thereby sequestering the fusion protein in the nucleolus. These experiments demonstrate that a Nobis lacking any endogenous NLS resides in the central and evolutionarily highly conserved region of S6 (supplementary fig. S3, Supplementary Material online).

NLS3 of Human Ribosomal Protein S6 Is of Bipartite Nature

Incorporation of the stronger NLS3 not only improved the efficiency of nuclear import but also surprisingly allowed for nucleolar accumulation, which prompted us to carry out a detailed characterization of NLS3. The following substitutions in the tetrapeptide sequence of the arginine at position 231, R231V, R231G, and R231E, gave rise to nonfunctional NLSs unable to direct the fused β-galactosidase into the nucleoplasm (fig. 2; supplementary fig. S2, Supplementary Material online for a color version). This experimental result was predicted from the NLS consensus sequence K-R/K-X-R/K postulated by Chelsky et al. (1989). Remarkably, this bipartite signal has been conserved over almost 1 billion years because it is also present in ribosomal protein S6A of S. cerevisiae (Lipsius et al. 2005).

Nobis2: A Minimal Nucleolar Binding Domain Encompassing NLS2

The sequence of human ribosomal protein S6 contains the peptide 188KRRIALKKQRTKKNK203, which is highly reminiscent of the nucleolar localization signals of the retroviral proteins rex, rev, and tat (Hatanaka 1990). Interestingly, this peptide sequence is the core element of the rather large fragments, S6(138–198) and S6(183–240), which targeted β-galactosidase into the nucleolus (Schmidt et al. 1995). Based on these observations, the existence of a shorter nucleolar binding domain in the region of NLS2 was postulated. To test this hypothesis, the S6(180–210) fragment fused to the N-terminus of β-galactosidase was constructed (fig. 3). The corresponding fusion protein was imported into the nucleus due to the presence of NLS2, however, it was excluded from the nucleolus (fig. 3). Extension of this peptide sequence by 10 amino acids at the N-terminus gave rise to the S6 fragment S6(170–210) which, when fused to β-galactosidase, was able to target the reporter protein exclusively into the nucleolus (fig. 3). In an attempt to define the C-terminal boundary of Nobis2, 12 amino acids were removed from S6(170–210) giving rise to S6(170–198), which abolished nucleolar accumulation (data not shown).

A sequence comparison of the eukaryotic S6e family members called our attention to an evolutionary highly conserved S6 peptide cluster between 2 proline residues: 174PKIQRLVTP182 (supplementary fig. S3, Supplementary Material online). The shortest construct, S6(173–203), containing this peptide imported the fusion protein into the nucleolus (fig. 3). Presumably, the evolutionary conserved cluster between amino acids 174 and 182 functions as an essential domain vital for nucleolar binding of this S6 fusion protein (fig. 3). The fusion constructs, S6(180–229) and S6(170–229), support the observation that the highly conserved peptide leads to an accumulation of the fusion proteins in the nucleolus (fig. 3). The fragment S6(192–249), which is an N-terminal extension of S6(210–249) used for the characterization of Nobis1, is also imported into the nucleus due to NLS3. A distinct nucleolar staining is absent because NLS2 and the highly conserved peptide cluster were removed (fig. 3).

Taken together, the minimal Nobis2 reminiscent of retroviral nucleolar targeting signals consists of a cluster of evolutionary highly conserved amino acids encompassing NLS2 as the core element (supplementary fig. S3, Supplementary Material online).

Nobis1 Suffices for Accumulation of S6 in the Nucleolus

The 13 basic amino acids of the 3 NLSs contribute 13 positive charges, which may prove essential for ionic...

**FIG. 2.—NLS3 of S6 is a bipartite import signal. The tetrapeptide 230KRRR233 and the fragment S6(215–240) were mutated by site-directed mutagenesis, fusion proteins with β-galactosidase were expressed in COS cells, and the intracellular distribution of the constructs was visualized by X-gal staining. C, cytoplasmic; N, nuclear. Scale bar, 5 μm.**
interactions with the 18S rRNA or other nucleolar factors. Deletion of the 3 NLSs gives rise to an S6 molecule that fails to promote the nuclear import of β-galactosidase (Schmidt et al. 1995). Fusing the SV40 NLS to the amino-terminal end of this triple deletion mutant resulted in nuclear import and recovery of the nucleolar targeting ability of S6 (fig. 4), most probably due to the presence of Nobis1, which functions independently of the 3 authentic NLSs, as demonstrated by the constructs SV40-S6(3–240, A1–3) and SV40-S6(72–120). This interpretation was supported by an SV40 construct comprising amino acids 173–249 but lacking all authentic NLSs, which was imported into the nucleoplasm and excluded from the nucleoli (fig. 4). Hence, Nobis2 in the C-terminal half of S6 is dependent on the presence of NLS2 in its native, evolutionary highly conserved context because the SV40 NLS could not restore nucleolar accumulation of the S6(173–249, A1–3) fragment (fig. 4).

Although the presence of the endogenous NLS2 is vital for the functioning of Nobis2, the deletion of all 3 NLSs still leads to nucleolar accumulation of S6 because Nobis1 remains intact in the construct SV40-S6(3–240, A1–3). Thus, Nobis1, residing in the highly conserved central region of S6, suffices for nucleolar binding of S6.

Both Nucleolar Binding Sequences of the S6e Family Are Evolutionary Conserved

The delineated minimal sequences Nobis1 and Nobis2 from human ribosomal protein S6 were aligned with the corresponding sequences of 9 selected species from vertebrates, invertebrates, and fungi (supplementary table S1, Supplementary Material online) to create a consensus logo. The combined logos of all 27 selected species are depicted in figure 5, and the logos of the subgroups are shown in supplementary figure S3 (Supplementary Material online for a color version).

The 4 boxes of Nobis1 designated A–D are almost completely conserved in the 9 vertebrate species used for comparison (supplementary fig. S3, Supplementary Material online). The Nobis1 of rainbow trout differs only by 3 amino acids: 1 in box A and 2 in box D. A variation in the C-terminal amino acids of box A and C is present in Nobis1 from invertebrates, whereas the clusters of box B and D are almost completely invariant. When comparing the 9 S6 of fungi, a similar pattern of sequence conservation can be deduced. Combining the vertebrates, invertebrates, and fungi produces the following sequence pattern of Nobis1: the N-termini of box A and C as well as the entirety of box B and D are highly conserved, whereas the C-termini of box A and C are variable (fig. 5). A similar pattern of sequence conservation with minor deviations can be observed in the Nobis1 of plants with established sequences: Nicotiana tabacum, Asparagus officinalis, Zea mays, Arabidopsis thaliana, and Oryza sativa. The basic structure and pattern of sequence conservation characteristic of Nobis1 are also displayed in the S6e family member Pyrococcus horikoshii of the Archaea.

Nobis2 is almost fully conserved in the vertebrate group (fig. 5 and supplementary fig. S3, Supplementary Material online for a color version). Invertebrates and fungi display a peptide cluster in box A that is similar to the vertebrates, whereas an invariant distribution of positive charges is exhibited in box B, albeit the amino acid sequences vary considerably (supplementary fig. S3, Supplementary Material online). Characteristic features of Nobis2 are the complete invariance of box A—when conserved amino acid exchanges are taken into account—and the high
density of positive charges distributed in 3 unequally sized clusters that are separated by 2 doublets of uncharged, mostly aliphatic amino acids in box B (fig. 5). This pattern is also present in the Nobis2 of plants (N. tabacum, A. officinalis, Z. mays, A. thaliana, and O. sativa).

Interestingly, the S6e family members of Archaea, which have shorter protein sequences that are discontinued soon after Nobis1, do not contain an rRNA-binding sequence comparable to Nobis2.

Discussion
Differential Import Efficiencies of the 3 S6 NLSs

The efficiency of the S6 delivery to the nucleus is exceptionally high because no free S6 protein, that is, not assembled into 40S subunits, can be detected in the cytoplasm during interphase (Wool and Stöffler 1976). Apparently, the import machinery flawlessly and rapidly translocates nearly 3,000 S6 molecules per min per HeLa cell from the cytoplasm into the nucleus (Kruppa A, Kruppa J, unpublished data).

Human ribosomal protein S6 has 3 NLSs each able to direct the reporter protein β-galactosidase into the nucleus, thereby leading to the accumulation of the corresponding fusion protein in the nucleolus (Schmidt et al. 1995). Detailed analysis revealed the bipartite nature of NLS3 in contrast to the monopartite NLS1 and NLS2 (fig. 2).

Human S6 and yeast S6A protein are endowed with 3 distinct import signals differing in their transport efficiencies (Schmidt et al. 1995; Lipsius et al. 2005), albeit a single NLS is sufficient to direct a reporter protein into the nucleus in vivo. Phosphorylation of the serine cluster at the C-terminus may affect the adjacent NLS3 by decreasing the affinity of the target sequence for the import factors. Hence, nuclear import would primarily have to rely on NLS2 because NLS1 is a weak import signal (Schmidt et al. 1995). Eukaryotic cells possess different importins with distinct NLS-binding specificities; therefore, the affinity of the importin–NLS interaction (Fontes et al. 2003) is the critical parameter in determining transport efficiency. The limiting cellular concentration of these importins as well as of the guanosine triphosphate–binding factor Ran influence the extent of nucleocytoplasmic transport (Jans et al. 2000).

Differences in import efficiencies within the 3 functionally distinct NLSs of ribosomal protein L5 have also been observed (Claussen et al. 1999; Rosorius et al. 2000). Only L5-NLS1 binds to a number of importins in vitro. However, in vivo L5-NLS1 and L5-NLS3 mediate the nuclear import of the 5S RNP. On the contrary, L5-NLS2 is incapable of translocating a heterologous 5S RNP into the nucleoplasm (Claussen et al. 1999).

Remarkably, the number, nature, and position of the targeting signals in S6 and L5 are evolutionary conserved. Specifically, the number of NLSs—2 monopartite and 1 bipartite—as well as their positions at the C-terminus of
ribosomal protein S6 have been conserved in vertebrates, invertebrates, and fungi (Ruvinsky and Meyuhas 2006).

Characteristic Features of the 2 S6 Nucleolar Binding Sequences Conserved throughout Evolution

After entering the nucleoplasm, the import complexes dissociate releasing the S6 protein. Given that there is no structural or functional evidence for a membrane barrier separating the nucleolus from the surrounding nucleoplasm, the imported soluble S6 protein should migrate into the nucleolar compartment where it associates with the pre-rRNA as an early binding protein required for efficient early cleavages (Ferreira-Cerca et al. 2005). S6 has been localized to the small head region of the 40S subunit near the mRNA/tRNA-binding site (Volarevic and Thomas 2001). Cross-linking studies have revealed that a region of the S6 molecule extends to the subunit interphase where it cross-links with L24 (Uchiumi et al. 1986). S6 interacts with components of the ternary initiation complex, thus contributing to the formation of the P-site on the small subunit (Bommer et al. 1980).

For the prokaryotic 3OS and 50S ribosomal subunits, the specific binding interactions between ribosomal proteins and the corresponding rRNA molecules have been precisely defined by X-ray crystallography (Brodersen et al. 2002; Klein et al. 2004). The atomic resolution structures of these subunits have demonstrated that most ribosomal proteins are located at the surface of the respective particle. The 3-dimensional structures of these proteins display globular domains with either internal ex-}

respective particle. The 3-dimensional structures of these subunits have demonstrated that most ribo-

molecular proteins are located at the surface of the 3-

dimensional architecture of the subunits, eukaryotic ribo-

somal proteins should also possess 1 or several regions that serve as rRNA– and/or protein–protein binding domains (Ramakrishnan and Moore 2001) giving rise to nucleolar localization.

A single Nobis, either Nobis1 or Nobis2, is sufficient to lead to an exclusive accumulation of S6 in the nucleolus (figs. 1 and 3). The 27-amino-acid-long Nobis1—

(G)RVRLLLSKGHSCYRPRPTGERKKRSVR(G)—contains 12 basic residues amounting to 41% in this oligopeptide compared with only 27% in the entire S6 protein. The N- and C-termini of Nobis1 are flanked by glycines that provide a certain conformational flexibility to Nobis1. Inherent to the amino acid sequence are several constraints that affect the 3-dimensional folding of Nobis1 preventing the formation of an extended α-helical structure: 1) electrostatic repulsion between successive basic residues (bold); 2) bulkiness of adjacent R-groups (underlined) in Ser, Cys, and Thr; and 3) occurrence of Pro and Gly residues (bold italics). Thus, it is assumed that the specific features of this Nobis are a perfect match for the 18S rRNA; this S6 protein–rRNA interaction occurs early during particle assembly in the GC of the nucleolus (Krüger et al. 2007) giving rise to the small subunit processome (Bernstein et al. 2004). Specific binding of ribosomal protein S6 to distinct nucleolar proteins was not detected during protein interaction mapping in S. cerevisiae (Gavin et al. 2006; Reguly et al. 2006).

Taken together, the sequence comparisons (fig. 5 and supplementary fig. S3, Supplementary Material online) demonstrate that Nobis1 plays an important role in the phylogenetic domain of Eukarya and was already established in the Archaea, which lack a nuclear compartment, implying that the highly basic Nobis1 is interacting with a specific, presumably evolutionary conserved rRNA sequence. It has not escaped our notice that nature employed existing basic amino acid sequences as nucleolar binding sequences after the appearance of the nucleus.

As mentioned in the results, Nobis2 is dependent on the functional NLS2 (fig. 4). Nobis2—APKIQQRLVTPRVLQHKRRRALKQTKKK—comprises of 31 amino acids, 12 of which are basic in nature. The basic residues correspond to 39% in this oligopeptide compared with only 27% in the entire S6 protein. The structural constraints mentioned for Nobis1 also hold true for Nobis2, thereby creating a positively charged, flexible binding sequence capable of optimally interacting with the 18S rRNA backbone.

Nobis1 and Nobis2 differ in 1 important aspect: Nobis1 resides in a region of the S6 protein sequence that lacks any NLS, whereas the functioning of Nobis2 is crucially dependent on NLS2 because deletion of NLS2 in SV40-S6(173–249,Δ1–3) concomitantly results in a loss of nucleolar accumulation (fig. 4). Thus, integration of an NLS into the second Nobis of S6 ascents the efficient recognition of the newly synthesized ribosomal protein by the nuclear import machinery, thereby leading to a rapid import into the nucleolus. Clearing the cytoplasm of the highly basic ribosomal proteins seems vital to avoid the deleterious binding of ribosomal proteins to mRNAs that would otherwise interfere with the translation process (Wool and Stöffler 1976; Jäkel and Görlich 1998). The nucleolar binding sequences are essential for driving the subunit assembly in the nucleolus, thus positively influencing the growth rate of cells. The presence of an NLS in a Nobis increases the efficiency of the ribosomal protein delivery into the nucleus, for example, for Nobis2 of S6e and for the Nobis of S25e. An NLS as an integral building block is also found in the nucleolar binding sequences of human S7 and L7α (table 3).

In essence, although nuclear targeting signals increase import efficiency, NLSs are superfluous because ribosomal proteins can diffuse through the nuclear pore into the nucleoplasm due to their low molecular weight, yet the presence of a Nobis is indispensable for nucleolar binding, thereby facilitating subunit assembly.

Does the C-terminal Serine Cluster of S6 Contribute to the 18S rRNA Binding?

The C-terminus of ribosomal protein S6 contains the serine cluster of phosphorylation sites, which play an important role as determinants of cell size, cell proliferation,
and glucose homeostasis (Ruvinsky et al. 2005; Ruvinsky and Meyuhas 2006). In vertebrates, the sequence of the serine cluster with its 5 phosphorylation sites is completely conserved, whereas the amino acid sequence varies in invertebrates (Ruvinsky and Meyuhas 2006). In contrast, fungi have to rely on a shorter C-terminus with just 2 serine residues (Lipsius et al. 2005).

By proton nuclear magnetic resonance spectroscopy, Katahira et al. (1996) obtained a solution structure for the C-terminal end, S6(217–249), containing the serine phosphorylation sites. This peptide folds into an α-helix between E222 and R238, a distorted helical structure for the following 3 residues, and a flexible tail lacking any secondary structure starting from S242 that must remain freely accessible to the p70 S6 kinase, which uses the tetrapeptide KRRR of NLS3 as its recognition site. By proton nuclear magnetic resonance spectroscopy, Katahira et al. (1996) obtained a solution structure for the C-terminal end, S6(217–249), containing the serine phosphorylation sites. This peptide folds into an α-helix between E222 and R238, a distorted helical structure for the following 3 residues, and a flexible tail lacking any secondary structure starting from S242 that must remain freely accessible to the p70 S6 kinase, which uses the tetrapeptide KRRR of NLS3 as its recognition site.

Evolutionary Highly Conserved Nucleolar Binding Sequences Occur in Numerous Ribosomal Proteins

Our detailed analyses clearly show that the nucleolar binding sequences of S6 reside in evolutionary highly conserved peptide clusters (fig. 5 and supplementary fig. S3, Supplementary Material online). Does this observation also hold true for the nucleolar binding sequences of other ribosomal proteins that have been published? Our conclusion is corroborated by the 2 publications (Kubota et al. 1999; Timmers et al. 1999) that delineated a single Nobis of yeast and human S25, respectively, which consist of a highly conserved cluster that has been maintained over more than 1 billion years of evolution (table 3 and supplementary fig. S4, Supplementary Material online). Previously, Annino et al. (1998) have identified a single Nobis in human S7 containing a high percentage of basic amino acids that is built from 2 clusters, which are evolutionary highly conserved, analogous to S6. In human S19, 2 short nucleolar binding sequences have been characterized: 1 located at the N-terminus and the other at the C-terminus (Da Costa et al. 2003). A single amino acid exchange in either Nobis fails to localize S19 in the nucleolus and leads to Diamond–Blackfan anemia (DBA). The C-terminal Nobis of S19 is inserted between 2 glycine residues similar to Nobis1 of S6. The mutations in S19 of DBA patients clearly demonstrate the importance of proper localization of ribosomal proteins for the development of a healthy organism.

Ribosomal proteins of the large subunit also contain evolutionary highly conserved nucleolar binding sequences as exemplified by human L22 (Shu-Nu et al. 2000) and rat L31 (Quaye et al. 1996), which are 6 and 5 residues long, respectively, whereas human L5 (Rosorius et al. 2000) displays 2 nucleolar binding sequences, both of which have an integrated NLS (table 3).

Supplementary Material

Supplementary table S1 and figures S1–S4 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Antonie Kruppa for her critical reading of the manuscript and her valuable suggestions.

Literature Cited


Annino T, Karis A, Hoth S, Rikk T, Kruppa J, Metspalu A. 1998. Nuclear import and nucleolar accumulation of the human ribosomal protein S7 depends on both a minimal nuclear

Table 3  
Nuclear Localization Signals and Nucleolar Binding Sequences in Ribosomal Proteins  

<table>
<thead>
<tr>
<th>Species</th>
<th>RP</th>
<th>Number</th>
<th>Length</th>
<th>Type</th>
<th>Nobis</th>
<th>Number</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens</td>
<td>S6</td>
<td>3</td>
<td>4/17</td>
<td>2 monopartite/1 bipartite</td>
<td>2</td>
<td>27/31</td>
<td>This paper</td>
</tr>
<tr>
<td></td>
<td>S7</td>
<td>1</td>
<td>21</td>
<td>Bipartite-like</td>
<td>1</td>
<td>21</td>
<td>Annino et al. 1998</td>
</tr>
<tr>
<td></td>
<td>S19</td>
<td>1</td>
<td>21</td>
<td>Bipartite</td>
<td>2</td>
<td>17/23</td>
<td>Da Costa et al. 2003</td>
</tr>
<tr>
<td></td>
<td>S25</td>
<td>1</td>
<td>4</td>
<td>Monopartite</td>
<td>1</td>
<td>17</td>
<td>Kubota et al. 1999</td>
</tr>
<tr>
<td></td>
<td>L5</td>
<td>3</td>
<td>17/11</td>
<td>Monopartite</td>
<td>2</td>
<td>17/11</td>
<td>Rosorius et al. 2000</td>
</tr>
<tr>
<td></td>
<td>L7a</td>
<td>3</td>
<td>29/49/120</td>
<td>ND</td>
<td>1</td>
<td>49</td>
<td>Russo et al. 1997</td>
</tr>
<tr>
<td></td>
<td>L22</td>
<td>1</td>
<td>4</td>
<td>Monopartite</td>
<td>1</td>
<td>6</td>
<td>Shu-Nu et al. 2000</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>L31</td>
<td>1</td>
<td>6</td>
<td>Monopartite</td>
<td>1</td>
<td>6</td>
<td>Quaye et al. 1996</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>S6A</td>
<td>3</td>
<td>14/18/17</td>
<td>2 monopartite/1 bipartite</td>
<td>2</td>
<td>62/42</td>
<td>Lipsius et al. 2005</td>
</tr>
<tr>
<td></td>
<td>S25</td>
<td>2</td>
<td>26/9</td>
<td>Bipartite/Mats2-like</td>
<td>1</td>
<td>26</td>
<td>Timmers et al. 1999</td>
</tr>
</tbody>
</table>

Note.—RP, ribosomal protein; ND, not determined.
localization sequence and an adjacent basic region. Biochem Biophys Res Commun. 249:759–766.
protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. Genes Dev. 19:2199–2211.

Manolo Gouy, Associate Editor

Accepted December 28, 2007