The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions
Edited by R. A. Garrett, S. R. Douthwaite, A. Liljas, A. T. Matheson, P. B. Moore, and H. F. Noller

Chapter 18

Nuclear Import of Ribosomal Proteins: Evidence for a Novel Type of Nuclear Localization Signal

ROGIER STUGER, ANTONIUS C. J. TIMMERS, HENDRIK A. RAUÉ, and JAN VAN 'T RIET

The formation of eukaryotic ribosomes is a highly complex process which requires the coordinated expression of a large set of ribosomal genes, transcribed by three different RNA polymerases, to ensure production of equimolar amounts of the four rRNAs and the approximately 80 ribosomal protein (r-protein) species under all growth conditions (for a recent review, see Planta, 1997).

A further level of complexity is added to eukaryotic ribosome biogenesis by the fact that it involves different cellular compartments. Transcription of the rRNA and r-protein genes takes place in the nucleolus and nucleoplasm, respectively. The r-protein mRNAs have to be exported to the cytoplasm. After translation, the r-proteins must be imported into the nucle(ol)us, where they have to be present in equimolar amounts to be assembled with the rRNAs into ribosomal subunits. The subunits are then exported from the nucleus to take up their function in the cytosol. Thus, ribosome biogenesis in eukaryotic cells involves massive transport of macromolecules and macromolecular complexes in both directions across the nuclear envelope (reviewed in Scheer and Weisenberger, 1994). This transport not only concerns the ribosomes themselves (or their components) but also a large number of accessory factors, varying from constituents of the transcription, translation, and splicing machinery to pre-rRNA-processing and ribosome assembly factors (Tollervey, 1996).

In recent years considerable insight has been gained into the mechanisms of nucleocytoplasmic transport (see Mattaj and Englmeier, 1998; Weis, 1998; and Wen et al., 1995 for recent reviews). Interestingly, a number of strong indications were found that nuclear import of r-proteins uses a specialized import pathway different from that used by the majority of karyophilic proteins. This suggests that the nuclear localization signals or sequences (NLSs) of r-proteins may be structurally distinct from the classical NLSs present in the latter class of proteins (Chelsky et al., 1989; Dingwall and Laskey, 1991). In this chapter we review present knowledge of the mechanism and signals responsible for the nuclear import of proteins, in particular, r-proteins. A database search of the complete set of yeast r-proteins for putative NLSs by using a set of criteria derived from the comparison of all experimentally identified signals in yeast r-proteins indicates that the large majority of these proteins may indeed possess a novel type of NLS, characterized by the presence of a sequence motif consisting of three basic residues within a 4- to 7-amino acid-long sequence, of which the first, the last, or both are helix-breaking residues.

NUCLEOCYTOPLASMIC TRANSPORT

Transport from cytoplasm to nucleus and vice versa is channeled through the nuclear pore complexes (NPCs). Such NPCs are complicated, proteinaceous, three-dimensional structures spanning the nuclear envelope (for reviews, see Fahrenkrog et al., 1998; Goldberg and Allen, 1995; Panté and Aebi, 1996). Although the NPC is freely permeable to small (macro)molecules (with masses below ~40 kDa) (Roberts et al., 1987), translocation through the NPC

Rogier Stuger ■ Department of Molecular Cell Physiology, Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands. **Antonius C. J. Timmers** ■ Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, CNRS-INRA, BP 27, 31326 Castanet-Tolosan Cedex, France. **Hendrik A. Raué and Jan van 't Riet** ■ Faculty of Science, Division of Chemistry, Department of Biochemistry and Molecular Biology, Institute for Molecular Biological Sciences, BioCentrum Amsterdam, Vrije Universiteit, de Boelelaan 1083, 1081 HV Amsterdam, The Netherlands.

is energy dependent (reviewed by Dingwall, 1991) and carrier mediated, depending on the presence of an NLS within the primary structure of the nuclear protein, even if its mass is far below 40 kDa (Breeuwer and Goldfarb, 1990; Schaap et al., 1991). Similarly, nuclear export depends upon the presence of nuclear export signals (Fischer et al., 1995; Michael et al., 1995; Wen et al., 1995). Several soluble factors that are required for import and export of nuclear proteins have been identified (Görlich and Mattaj, 1996; Melchior and Gerace, 1995). In the original model for nuclear import, an NLS receptor (Görlich et al., 1995), also named karyopherin-α or importin α, binds its karyophilic transport ligand before it interacts with importin β_1, also named karyopherin-β or p97 (Adam and Adam, 1994; Moore and Blobel, 1994). With the aid of the transport factors p10 (also called pp15 or NTF2) (Nehrbass and Blobel, 1996; Paschal and Gerace, 1995) and Ran-GDP, this complex then binds to nucleoporins at the cytoplasmic face of the NPC. After translocation of the complex into the nucleus, it is dissociated by the exchange of GDP for GTP through the action of the RanGEF protein RCC1. The transport factors are recycled into the cytoplasm, using receptors that also belong to the growing family of importin β-like proteins (see Mattaj and Englmeier, 1998; Weis, 1998; and Wozniak et al., 1998, for recent reviews).

The importin α adaptor recognizes the classical mono- and bipartite NLSs, of which the ones present in the simian virus 40 large T antigen (PKKKRKV) (Kalderon et al., 1984) and in nucleoplasmin (Dingwall and Laskey, 1991) are the respective prototypes. Importin α appears to have two binding sites for the monopartite signal. The same sites are also responsible for binding the two parts of the bipartite NLS (Conti et al., 1998). Importin α itself contains a bipartite signal in its N-terminal region, the so-called IBB domain, that is recognized by the importin β receptor (Görlich and Mattaj, 1996). Recently, it was found that the IBB domain can also interact with the NLS binding domain of the same importin α molecule (Kobe, 1999). This interaction may take place in the nucleoplasm after disruption of the nuclear cargo-importin α-importin β complex by Ran-GTP, thereby preventing reassembly of the complex. In the cytoplasm this self-inhibitory effect is prevented by the high affinity of the IBB domain for the IBB binding domain of importin β, thus exposing the NLS binding domain of importin α for interaction with the nuclear cargo.

NLSs

As indicated above, two major classes of NLSs have been identified in eukaryotic proteins. The classical, monopartite NLS consists of a short stretch of 4 to 7 amino acids in which basic and hydrophobic residues predominate. Chelsky et al. (1989) have proposed the sequence KR/KXR/K (X being any amino acid) as a consensus sequence for such monopartite NLSs. However, not every sequence matching this consensus shows nuclear targeting activity (e.g., Schmidt et al., 1995). Dang and Lee (1989) observed that many monopartite NLSs contain a helix-breaking residue, either glycine or proline, which may reflect the requirement for a "random" structure of the sequence for it to be able to function as an NLS. An additional type of NLS having the sequence KIPIK, and thus similar but not identical to the monopartite type, was identified in the yeast transcriptional repressor Matα2 (Hall et al., 1984).

The bipartite NLS is characterized by the presence of a basic dipeptide separated from another sequence rich in basic amino acids by about 10 arbitrary residues (Dingwall and Laskey, 1991). It should be noted that the downstream portion of the bipartite NLS conforms to the Chelsky consensus sequence.

In the past few years a number of NLSs that cannot be grouped into either of the two classes described above have been discovered. These include NLSs of karyophilic proteins, such as the M9 domain of the hnRNP A1 protein (Aitchison et al., 1996; Pollard et al., 1996), as well as the complex signals of U snRNPs (Fischer et al., 1991). Nuclear import of these components has been found to involve adaptors and/or receptors different from the ones required for transport of components containing a classical mono- or bipartite NLS, although they clearly belong to the same family as importin β (Huber et al., 1998; Wozniak et al., 1998). In yeast the importin β family presently consists of 13 members, most of which have a vertebrate homologue (reviewed in Wozniak et al., 1998). In many cases, these proteins interact directly with the NLS-bearing protein without the intervention of an importin α-like adaptor.

In yeast, NLSs vary considerably in length and composition. Many show only weak similarity to the consensus sequences described above, although in most cases they are composed of basic and hydrophobic residues (Osborne and Silver, 1993). Nevertheless, the machinery for nuclear protein import appears to be functionally conserved between yeast and higher eukaryotes, since, regardless of their source, NLSs are functional in either type of cell (Jeeninga et al., 1996; Schaap et al., 1991; Wagner and Hall, 1993).

NUCLEAR IMPORT OF r-PROTEINS

Nuclear import of r-proteins has been most extensively investigated in yeast. So far, the NLS se-

quences of six yeast r-proteins (L3, L25, L28, S17, S22, and S25, according to the new nomenclature of Mager et al., 1997) have been identified experimentally by assessing the ability of various r-protein fragments to stimulate nuclear import of a reporter protein that normally resides in the cytoplasm. A caveat to this approach is that a region identified as having nuclear targeting activity might not be a bona fide NLS, i.e., a sequence recognized by a nuclear import receptor, but might be required for the association of the protein with another NLS-bearing protein ("piggy-back" import) (Booher et al., 1989). Although the existence of such "preassembly" complexes of r-proteins has been suggested (Fried, 1993), it is at present entirely hypothetical.

The first notable property shared by the yeast r-protein NLSs is their occurrence close to the N terminus of the protein. In the cases of both L3 (Moreland et al., 1985) and L25 (Schaap et al., 1991), the NLSs are located within a region of the r-protein that does not have a counterpart in its prokaryotic homologue (Fried, 1993; Rutgers et al., 1990). Such extensions, therefore, may have evolved to accommodate this eukaryote-specific functional signal.

A second notable feature is the presence of multiple NLSs in three of the six yeast r-proteins so far investigated. In L25 and L28 these two NLSs are located in the immediate neighborhood of each other and, since they enhance each other's activity (Underwood and Fried, 1990; Schaap et al., 1991), might together be considered a degenerate bipartite NLS (Fried, 1993). However, each of the sequences on its own acts as an NLS, as shown by mutational analysis within the context of the r-protein itself and the ability of the individual sequences to stimulate nuclear import of a reporter protein (Underwood and Fried, 1990; Schaap et al., 1991). In S25 the two NLSs are separated by about 50 residues. Moreover, the N-terminal signal is considerably more potent than the C-terminal one (Timmers et al., 1999).

Both the multiplicity and the localization close to the N terminus might increase the probability of interaction of the r-protein NLSs with the receptors and thus ensure rapid translocation of the r-protein through the nuclear pores. Nuclear import of r-proteins has to be highly efficient, since no pools of free r-proteins are present in the cytosol under any circumstances. Moreover, in yeast cells unassembled r-proteins are rapidly degraded in the cytosol (Warner et al., 1985; El Baradi et al., 1986; Maicas et al., 1988).

The yeast r-protein NLSs shown in Table 1 can be divided into three groups. Those present in L3 and the N-terminal region of S25 correspond to the classical bipartite consensus (but see below). The C-terminal, weak NLS of S25 has features in common with the import signal identified in the yeast Matα2 transcription factor (KIPIK) (Hall et al., 1984). It should be noted that the efficiency of this signal is significantly improved when adjacent sequences are present (Timmers et al., 1999). The NLSs of the yeast r-proteins S17, S22, L25, and L28, apart from a predominance of basic and hydrophobic amino acids, show only weak similarity to the classical NLSs. On the other hand, they do display mutual similarity in that they all contain a motif consisting of three basic residues clustered within a 4- to 7-amino-acid-long sequence that is either N- or C-terminally flanked by a helix-breaking Pro or Gly residue. It should be noted that the same motif is also found as part of the (supposedly) bipartite NLSs of proteins L3 and S25 (Table 1).

The existence of a specific, functionally distinct type of NLS in yeast r-proteins is supported by the recent observation that these proteins use a nuclear import pathway different from the one employed by karyophilic proteins that carry a classical mono- or bipartite NLS. The first evidence was reported by Nehrbass et al. (1993), who found that a mutation in the yeast nuclear pore protein Nsp1 did not affect nuclear import of r-protein L25 while it prevented the import of nonribosomal proteins. Many yeast r-proteins, in particular, L25, were shown to be recognized by the importin β homologues Kap123p (also called Yrb4p and yer110c), Kap121p (also called Pse1p and ymr308c), and Sxm1p (or ydr395w) but not by importin α or importin β (Rosenblum et al., 1997; Rout et al., 1997; Schlenstedt et al., 1997). A similar situation exists in mammalian cells, where the nuclear import of rat r-proteins S7, L5, and L23a could be achieved by four different vertebrate importins, namely, transportin; RanBP5, the homologue of yeast Pse1p; and Ran BP7; as well as importin β itself, but without the aid of importin α (Jäkel and Görlich, 1998).

A CONSENSUS NLS FOR YEAST r-PROTEINS

In order to obtain further support for the existence of a functionally distinct NLS (YRP-NLS) in yeast r-proteins, we performed a computer search of the complete set of yeast r-protein sequences present in the Swiss-Prot database by using the consensus sequence $(K/R)_3X_{1-4}$ either preceded or followed by a helix-breaking Gly or Pro. No acidic residues, either within or adjacent to the consensus, were allowed. The consensus is based upon the common features of

Table 1. NLSs identified experimentally in yeast r-proteins[a]

Protein		NLS	Reference
L3		1-SHRKYEAPRHGHL**GFLPRKRA**-21	Moreland et al., 1985
L25	NLS1	1-MAPSAKATAA**KKAVVKG**-17	Schaap et al., 1991
	NLS2	18-TN**GKKALKVR**TSATFRLPKTLKLAR-41	
L28	NLS1	6-**KHRKHPG**-12	Underwood and Fried, 1990
	NLS2	23-**KTRKHRG**-29	
S17		2-**GRVRTK**-7	Gritz et al., 1985
S22		20-**GKRQVLIRP**-28	Timmers et al., 1999
S25		11-AKAAAALAG**GKKSK**KKWSKKSMKDRA-36	Timmers et al., 1999
		87-GI*IKPISK*H-95	

[a] The minimal sequences identified as able to direct a reporter protein efficiently into the nucleus are shown. The numbers indicate the position of the first and last residue in the complete sequence of the protein. Underlining indicates sequences conforming to the classical bipartite consensus. The sequences in boldface match the YRP consensus discussed in the text. The sequences corresponding to the Matα2-like NLS (Hall et al., 1984) are in italics.

the r-protein NLSs discussed above, the results of the detailed mutational analysis of the L28 and S22 NLSs (Timmers et al., 1999; Underwood and Fried, 1990), and the observation that the C-terminal region of L25 contains two sequences that are very similar to its NLSs except for the absence of the flanking glycines. These sequences had no nuclear targeting activity (Schaap et al., 1991). Additional searches were carried out to score for the presence of potential bipartite and Matα2-like sequences with a strategy to identify peptide sequences containing ambiguous residues (Vodkin et al., 1996).

As shown in Table 2, 61 of the 78 individual yeast r-proteins (78%) contain one or more sequences matching the proposed YRP-NLS consensus. The emergence of these putative NLSs is not simply a consequence of the abundance of basic residues in r-proteins: of a total of 283 basic stretches, only 112 were identified by our search as containing a match to the consensus. Furthermore, for S22, S25, L25, and L28, the only matches to the YRP-NLS consensus found correspond to the experimentally identified signals. Only for S17 did we find an additional match, which is located in a region of the protein not tested for nuclear targeting activity (Gritz et al., 1985).

About half of the putative YRP-NLSs identified by our search might also be considered to be part of a putative bipartite NLS. However, in this case the abundance of basic residues may well lead to a high percentage of false positives. Nevertheless, 35 yeast r-proteins contain matches to the YRP-NLS that are not part of a bipartite consensus, and about half of these r-proteins do not have an identifiable match to the bipartite consensus at all. Only eight yeast r-proteins (10%) do not possess a sequence matching the YRP-NLS criteria but do contain a region corresponding to the bipartite consensus. Finally, a Matα2-like sequence was found in six of the yeast r-proteins, but all of these proteins also contain matches to the YRP-NLS and/or bipartite consensus.

These data clearly support the existence of a novel type of NLS in at least the majority of the yeast r-proteins that could be responsible for directing the protein to the specialized import pathway. The r-proteins that failed to give a match to the YRP-NLS consensus might use the standard importin α-importin β pathway, since they do contain a putative bipartite NLS.

No putative NLSs were found in the acidic r-proteins P1 and P2, which are known to assemble in the cytoplasm (reviewed in Ballesta et al., 1999). Acidic r-protein P0, on the other hand, which is the functional homologue of the rRNA-binding *Escherichia coli* protein L10 (Shimmin et al., 1989), does contain a YRP-like sequence. We were also unable to find matches to any of the NLS consensus sequences in a further seven nonacidic yeast r-proteins. This does not necessarily mean that these r-proteins fail to be imported into the nucleus, however. They may contain an NLS that diverges too much from our search criteria, or they could be imported by a piggyback mechanism (Booher et al., 1989) after associating with another r-protein(s) in the cytoplasm. Interestingly, a large portion of the set (proteins S27, S28, L12, and L30) was classified as "late-assembling" proteins on the basis of kinetic labeling studies carried out in our laboratory (Kruiswijk et al., 1978).

NLSs OF MAMMALIAN r-PROTEINS

Mammalian r-proteins also use a specialized nuclear import pathway not involving importin α. The situation is even more complex than in yeast, however, since four different receptors have been identified, all belonging to the importin β family:

Table 2. Putative NLSs identified by computer search of the complete set of yeast r-protein sequences[a]

YRP	Residues	Type of NLS[b]	Sequence with (putative) NLSs
S1	3–18	Y* B	**GKNKRLSKGKKGQKKR**
	150–165	B	KRHSYAQSSHIRAIRK
S2	9–27	Y* B	KRGGFG**GRNRGRPNRRG**PR
S3	63–77	B	RRINELTLLVQKRFK
S4	4–10	Y	**PKKHLKR**
	123–127	Y	**GKVKK**
S5	74–80	Y	**GRYANKR**
	100–111	Y*	**GRNNGKKLKAVR**
S6	87–99	Y B	RRDGER**KRKSVRG**
	135–142	Y	**PKRANNIR**
	181–188	Y	**PQRLQRKR**
	214–230	B	KRLSERKAEKAEIRKRR
S7	95–115	Y B	RRILPKPSRTS**RQVQKRP**RSR
	136–140	Y	**GKRVR**
S8	9–25	B	KRSATGAKRAQFRKKRK
	36–55	B M	*KIGAK*RIHSVRTRGGNKKYR
S9	38–91	B	KREIYRISFQLSKIRR
	53–70	B	RRAARDLLTRDEKDPKRL
	77–91		RRLVRVGVLSEDKKK
	169–174	Y	**GRVARR**
S10	3–20	Y* B	**GKNKRLSKGKKGQKKRVV**
	150–165	B	KRHSYAQSSHIRAIRK
S11	22–29	Y	**PKVKTSKR**
	86–104	B	RRAYLHYIPKYNRYEKRHK
S12		No matches	
S13	38–43	Y	**KYARKG**
	92–108	B	KKAVSVRKHLERNRKDK
S14	126–135	Y*	**RKKGGRRGRR**
S15	42–61	Y B	**RRRFARG**MTSKPAGFMKKLR
S16	25–29	M	*KGLIK*
	126–132	Y	**KKFGGKG**
S17	2–7	Y	**GRVRTK**
	32–49	B	KRLCDEIATIQSKRLRNK
	59–64	Y	**KRIQKG**
S18	117–124	Y	**KKIRAHRG**
S19	101–122	Y B	RKVLQALEKIGIVEIS**PKGGRR**
S20	52–59	Y	**KKGPVRLP**
	68–89	B	RKTPNGEGSKTWETYEMRIHKR
S21	3–20	Y* B	KKRASN**GRNKKGRGHVRP**
S22	20–28	Y	**GKRQVLIRP**
S23	2–8	Y	**GKGKPRG**
	13–32	B*	RKLRVHRRNNRWAENNYKKR
	69–82	B	RKCVRVQLIKNGKK
	108–115	Y	**GRKGKAKG**
S24	107–131	Y B	RKQKKNRDKKIFGT**GKRLAKK**VARR
S25	20–33	Y B	**GKKSKKK**WSKKSMK
	90–94	M	KPISK
S26	3–15	Y B	KKRASN**GRNKKGR**
S27		No matches	
S28		No matches	
S29	10–16	Y	**PRRYGKG**
S30	26–54	Y B*	**PKKPKGRAYKR**LLYTRRFVNVTLVNGKRR
S31	1–23	Y B*	**GKKRKKK**VYTT**PKKIKHK**HKKVK
P0	54–61	Y	**GKNTMVRR**
P1		No matches	
P2		No matches	
L1	91–106	B	KKLNKNKKLIKKLSKK
L2	5–10	Y	**RNQRKG**
	240–247	Y	**RRTGLLRG**
L3	3–29	Y B	RKYEAPRHGHL**GFLPRKRA**ASIRARVK
	114–127	B	KRRFYKNWYKSKKK
	235–247	B	KKLPRKTHRGLRK

Continued on next page

Table 2. —*Continued.*

YRP	Residues	Type of NLS[b]	Sequence with (putative) NLSs
L4	184–203	Y* B	KKLRA**GKGKYRNRRW**TQRRG
L5	21–35	B	RRRREGKTDYYQRKR
	258–273	B	KKFTKEQYAAESKKYR
L6A	21–26	Y	**RKAARP**
	44–50	Y	**GRFRGKR**
L6B	17–26	Y	**PKKTRKAVRP**
	44–50	Y	**GRFRGKR**
L7	93–100	Y M	***KIPPK*PRK**
	144–151	Y	**RQLVYKRG**
	153–161	Y	**GKINKQRVP**
	216–220	Y	**PRKFK**
L8	41–62	Y B	**PKRNLSR**YVKWPEYVRVQRQKK
	170–184	B	KKMGVPYAIVKGKAR
	229–250	B	KKHWGGGILGNKAQAKMDKRAK
L9	85–91	Y	**GYKYKMR**
L10	198–204	Y	**KFLSKKG**
L11	139–145	Y	**RRKRCKG**
L12		No matches	
L13	15–36	Y B	RKHWQERVKVHFDQA**GKKVSRR**
	177–199	B	KKFRGIREKRAREKAEAEAEKKK
L14	18–26	Y	**GRVVLIKKG**
	108–128	B	RRAALTDFERFQVMVLRKQKR
L15	48–77	Y* B	RRLGYKAKQGFVIY**RVRVRRGNRKRPVPKG**
	106–113	Y	**GRRAANLR**
	167–203	Y B*	**GKKSRG**INKGHKFNNTKA**GRRKTWKR**QNTLSLWRYRK
L16	113–133	B	KKKRVVVPQALRVLRLKPGRK
	158–171	B	KRKVSSAEYYAKKR
L17	122–127	Y	**PKQRRR**
	162–180	Y B	KKVVRLTSRQR**GRIAAQKR**
L18	170–183	Y B	RKFERAR**GRRRSKG**
L19	7–20	B	KRLAASVVGVGKRK
	51–63	B	KKAVTVHSKSRTR
	69–99	Y B	KREGRHSGY**GKRKGTR**EARLPSQVVWIRRLR
	161–172	B	RRLKNRAARDRR
L20	163–170	Y	**KTFSYKRP**
L21		No matches	
L22	41–48	Y	**PKRNLSRY**
	170–184	B	KKMGVPYAIVKGKAR
L23	63–88	B*	KKGKPELRKKVMPAIVVRQAKSWRRR
L24	43–77	Y B*	RKNPRRIAWTVLF**RKHHKKG**ITEEVAKKRSRKTVK
	93–127	B	RRSLKPEVRKANREEKLKANKEKKKAEKAARKAEK
L25	10–26	Y* B	**KKAVVKG**TN**GKKALKVR**
L26	50–64	Y B	RRDDEVLVV**RGSKKG**
L27	16–28	Y*	**GRYAGKKVVIVKP**
	55–73	Y B M	KKHGAKKVA**KRTKIKP***FIK*
L28	8–28	Y* B	**RKHRG**HVSAGK**GRIGKHRKHP**
	110–114	M	*KILGK*
L29	12–40	Y B*	RKAH**RNGIKKP**KTYKYPYLKGVDPKFRR
L30		No matches	
L31	17–30	B	KRLHGVYFKKRAPRAVKEIKKFAK
	60–78	B	KRGVKGVEYRLRLRIYRKR
L32	7–12	Y	**PKIVKK**
	15–48	Y B*	KKFKRHHSDRYHRVAENW**RKQKG**IDSVV**RRRFRG**
L33		No matches	
L34	36–42	Y	**KKLATRP**
	102–121	B	KKVVKEQTEAAKKSEKKSKK
L35	42–49	Y	**PKIKTVRK**
	72–107	Y B*	**GKKYQPKDLRAKK**TRALRRALTKFEASQVTEKQRKK
L36	23–30	Y	**PKISYKKG**
	54–85	B	RRLIDLIRNSGEKRARKVAKKRLGSFTRAKAK
L37	9–24	Y B	**GKRHNK**SHTLCNRCGRR
	53–74	B	KRRHTTGTGRMRYLKHVSRRFK

Continued on next page

Table 2. —*Continued.*

YRP	Residues	Type of NLS[b]	Sequence with (putative) NLSs
L38	16–46	B*	RRADVKTATVKINKKLNKAGKPFRQTKFKVR
	60–64	Y	**GKAKK**
L39	16–21	Y	**KKQNRP**
L40	21–52	Y B	RKCYARLP**PRATNCRK**RKCGHTNQLR**PKKKLK**
L41	6–25	B	RKKRTRRLKRKRRKVRARSK
L42	4–8	Y	**PKTRK**
	13–18	Y	**GKTCRK**
	85–99	B	KRCKHFELGGEKKQK
L43	5–27	B	KKVGITGKYGVRYGSSLRRQVKK
	26–49	Y B	KKLEIQQHARYDCSFC**GKKTVKRG**

[a] The complete sequences of all *S. cerevisiae* r-proteins were searched for the presence of the three types of NLSs (YRP-like, bipartite, and Matα2-like). The regions containing YRP-like sequences are shown in boldface. The underlined sequences have been shown experimentally to act as NLSs (cf. Table 1).
[b] Y, YRP-like; B, bipartite; M, Matα2-like. An asterisk indicates that the sequence in question contains multiple matches to the consensus.

transportin, which is also involved in nuclear import of hnRNP proteins; importin β itself; RanBP5, the homologue of the yeast r-protein importin Pse1p; and RanBP7. Each of these receptors was shown to be able to promote nuclear import of at least three different r-proteins (S7, L5, and L23a) directly, without the help of importin α. The r-protein S7, but not L5 or L23a, is also able to use the classical importin α-importin β pathway (Jäkel and Görlich, 1998).

Regions containing NLSs have been experimentally identified so far in five different mammalian r-proteins, including S7 and L23a (Table 3). In almost all cases the signals were found to be quite complex. They consist either of rather long stretches of amino acids (S7 and L23a) or of multiple sequences (S6 and L7a). In the latter case the individual sequences do have nuclear targeting activity by themselves, but more efficient import is observed when all of them are present (Annilo et al., 1998; Russo et al., 1997; Schmidt et al., 1995). Only L31 appears to possess a "simple" NLS (Quaye et al., 1996).

The NLS of L23a has been subjected to detailed functional characterization. This NLS encompasses residues 32 to 74, a sequence that is able to bind each of the four import receptors mentioned above but does not appear to contain separate binding sites for these receptors. Interestingly, as shown in Fig. 1, this NLS corresponds exactly to the C-terminal portion of the eukaryotic extension of its yeast homologue, L25, which contains the NLSs (Rutgers et al., 1990, 1991; Schaap et al., 1991). We have previously shown that rat L23a can be incorporated into functional yeast 60S subunits (Jeeninga et al., 1996). Presumably, therefore, L23a can use the r-protein-specific import pathway of yeast cells. While the NLS of L23a clearly is not identical to NLS2 of L25, it contains a large number of overlapping YRP motifs (Fig. 1) that could possibly allow its interaction with the yeast r-protein-specific import receptors. However, the complete region of L23a from position 32 to 74 is required for nuclear import in HeLa cells (Jäkel and Görlich, 1998). The simple

Table 3. NLSs identified experimentally in mammalian r-proteins[a]

Protein[b]	NLS	Reference
hS6	165-E**GKKPR**TKAPK-175	Schmidt et al., 1995
	182-**PRVLQHKR**RRIALKKQ-197	
	215-RMKEAKEKRQEQIAKRRRLSSLRA-239	
hS7	98-**RRILPKPTRKSR**TKN**KQKRP**R-118	Annilo et al., 1998
hL7a	23-EAKKVVNPLFEKRPKNTFGIGQDIQPKRDL-51	Russo et al., 1997
	52-TRFVKWPRYIRLQRQRAILY**KRLKVP**PAINQFTQALDRQTATQLLKLAH-100	
	101-KYRPETKQEKKQRLLARAEKKAAGKGDVPTKRPPVLRAGVNTVTTLVENKKAQLVV	
	LAHDDPIELVVFLPALCRKM-220	
hL23a	31-**GVHSHKKKKIRTSPTFRRPKTLRLRRQPKYPRKSAPR**RNKLDHYA-74	Jäkel and Görlich, 1998
rL31	81-**PYRIRVR**SLRKR-92	Quaye et al., 1996

[a] The minimal sequences identified as able to direct a reporter protein to the nucleus are underlined. The regions within these sequences matching either the mono- or bipartite consensus but not experimentally tested for NLS activity are doubly underlined. The sequences containing one or more matches to the YRP consensus are shown in boldface. The numbers indicate the positions of the first and last residues in the complete sequence of the protein.
[b] The species of origin is indicated by the prefix: h, human; r, rat.

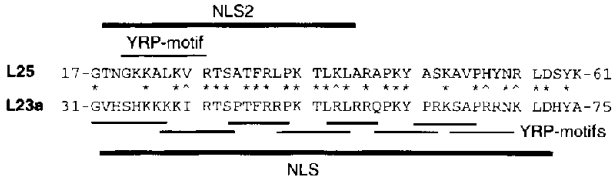

Figure 1. Comparison of the NLS of human r-protein L23a with the corresponding region in its L25 functional homologue from *S. cerevisiae*. *, identical residues; ^, similar residues. The sequences identified experimentally as having NLS activity are indicated by thick lines. YRP-like motifs are indicated by thin lines.

NLS identified in mammalian L31, on the other hand, might conform to the YRP consensus because the sequence RLSRKR, when fused to the β-galactosidase reporter, is preceded by a proline residue (Quaye et al., 1996). Similarly, YRP motifs that overlap with the NLSs are present in S6 and L7a (Table 3). Mutation of Pro^{117} in the C-terminal YRP motif significantly impaired nuclear import of S7 (Annilo et al., 1998). We have also searched the complete set of rat r-protein sequences and found that sequences matching the YRP consensus are present in 87% of these proteins. A similar analysis of the *E. coli* r-proteins, on the other hand, revealed YRP-like motifs in only 42% of the 55 species. Thus, such motifs occur considerably more frequently in eukaryotic r-proteins than in their prokaryotic counterparts.

NUCLEOLAR LOCALIZATION OF r-PROTEINS

Most of the assembly of ribosomal subunits in eukaryotic cells takes place in the nucleolus (for a review see Raué and Planta, 1991, and Mélèse and Xue, 1995). Shortly after their synthesis newly formed r-proteins accumulate rapidly in this subcompartment of the nucleus (Warner, 1979). However, the manner in which r-proteins (and other nucleolar proteins) find their way into the nucleolus is still obscure.

Sequences required for nucleolar accumulation have been identified in a number of r-proteins of mammalian origin (Schmidt et al., 1995; Quaye et al., 1996; Russo et al., 1997; Annilo et al., 1998), as well as the yeast r-protein S25. In the latter case the fragment containing the NLS of the r-protein (cf. Table 1) was also necessary and sufficient to direct the tagged r-protein to the nucleolus (Timmers et al., 1999). A similar colocalization of nuclear and nucleolar targeting activity was seen in mammalian r-protein L31, although the two activities could be separated by mutation of specific residues (Quaye et al., 1996). In mammalian r-proteins S6 (Schmidt et al., 1995), S7 (Annilo et al., 1998), and L7 (Russo et al., 1997), the nucleolar targeting activity is more complex in nature, requiring the cooperation of multiple functional domains that partly overlap with the sequences showing nuclear targeting activity.

Because nucleolar accumulation does not require passage through a membrane, the precise function of sequences that direct a protein to the nucleolus is still debated. In a number of cases, evidence has been presented that nucleolar accumulation is based upon retention of the protein by its binding to another nucleolar component(s), either protein or nucleic acid, rather than by directed transport through the action of a nucleolar targeting sequence (Schmidt-Zachmann and Nigg, 1993; Xue and Mélèse, 1994). For instance, yeast proteins like Nsr1p that have RNA binding as well as acidic (phosphorylated) domains may bring together rRNA and the (basic) r-proteins (Xue and Mélèse, 1994). In mammalian cells nucleolin may play a similar role (Bouvet et al., 1998).

CONCLUSION

In conclusion, it is now well established that most, if not all, eukaryotic r-proteins are imported into the nucleus via multiple specialized pathways, distinct from the one mediated by the importin α-importin β dimer (Rout et al., 1997; Schlenstedt et al., 1997; Jäkel and Görlich, 1998). The redundancy may be necessary to satisfy the massive requirement for import of r-proteins, particularly during rapid growth, when several hundred r-protein molecules have to be imported per NPC per minute. It may also act as a "fail-safe" mechanism to ensure the supply of r-proteins in case one of the pathways is inactivated.

The import pathways for r-proteins are simpler than the one involving the classical mono- and bipartite NLSs because they involve interaction of the r-protein only with an importin β-like receptor that delivers its cargo directly to the NPC. The domain of this receptor recognizing the r-protein differs from the domain that interacts with the bipartite signal of importin α (Jäkel and Görlich, 1998). The existence of a functionally distinct NLS implied by these observations is supported by the computer analysis of the yeast r-protein sequence database discussed in this chapter. This novel type of NLS appears to be characterized by the presence of a $[G/P](K/R)_3X_{1-4}[G/P]$ consensus motif in which there is a preference for X being hydrophobic and acidic amino acids are excluded from positions within or adjacent to the consensus motif. However, in some cases, notably in L25

(Schaap et al., 1991), the presence of adjacent sequences that themselves have no nuclear targeting activity considerably improves the activity of the region containing the YRP motif. In other yeast r-proteins, such as L28 (Underwood and Fried, 1990) and S22 (Timmers et al., 1999), the YRP motif by itself appears to have strong NLS activity (cf. Table 1).

Although matches to this consensus motif are found in many mammalian r-proteins, their NLSs appear to be more complex in nature, often distributed over multiple regions, each of which contributes part of the activity.

The work carried out in our laboratory was supported in part by the Netherlands Foundation for Chemical Science (NWO-CW), by a "STIMULANS" grant from the Netherlands Organization for Scientific Research (NWO) to A.C.J.T., and a grant to R.S. from the Royal Netherlands Chemical Society (KNCV) and the Netherlands Ministry of Economic Affairs.

REFERENCES

Adam, E. J. H., and S. A. Adam. 1994. Identification of cytosolic factors required for nuclear location: sequence-mediated binding to the nuclear envelope. *J. Cell. Biol.* 125:547–555.

Aitchison, J. D., G. Blobel, and M. P. Rout. 1996. Kap104p: a karyopherin involved in the nuclear transport of messenger RNA binding proteins. *Science* 274:624–627.

Annilo, T., A. Karis, S. Hoth, T. Rikk, J. Kruppa, and A. Metspalu. 1998. Nuclear import and nucleolar accumulation of the human ribosomal protein S7 depends on both a minimal nuclear localization sequence and an adjacent basic region. *Biochem. Biophys. Res. Commun.* 249:759–766.

Ballesta, J. P. G., M. A. Rodriguez-Gabriel, G. Bou, E. Briones, R. Zambrano, and M. Remacha. 1999. Phosphorylation of the yeast ribosomal stalk. Functional effects and enzymes involved in the process. *FEMS Microbiol. Lett.* 23:537–550.

Booher, R. N., C. E. Alfa, J. S. Hyams, and D. H. Beach. 1989. The fission yeast cdc2/cdc13/suc1 protein kinase: regulation of catalytic activity and nuclear localization. *Cell* 58:485–497.

Bouvet, P., J. J. Diaz, K. Kindbeiter, J. J. Madjar, and F. Amalric. 1998. Nucleolin interacts with several ribosomal proteins through its RGG domain. *J. Biol. Chem.* 273:19025–19029.

Breeuwer, M., and D. S. Goldfarb. 1990. Facilitated nuclear transport of Histone H1 and other small nucleophilic proteins. *Cell* 60:999–1008.

Chelsky, D., R. Ralph, and G. Jonak. 1989. Sequence requirements for synthetic peptide-mediated translocation to the nucleus. *Mol. Cell. Biol.* 9:2487–2492.

Conti, E., M. Uy, L. Leighton, G. Blobel, and J. Kuriyan. 1998. Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. *Cell* 94:193–204.

Dang, C. V., and W. M. F. Lee. 1989. Nuclear and nucleolar targeting sequences of c-*erb*-A, c-*myb*, N-*myc*, p53, Hsp70 and HIV *tat* proteins. *J. Biol. Chem.* 264:18019–18023.

Dingwall, C. 1991. Transport across the nuclear envelope: enigmas and explanations. *Bioessays* 13:213–217.

Dingwall, C. M., and R. A. Laskey. 1991. Nuclear targeting sequences—a consensus? *Trends Biochem. Sci.* 16:478–481.

El Baradi, T. T. A. L., C. A. F. M. Van der Sande, W. H. Mager, H. A. Raué, and R. J. Planta. 1986. The cellular level of yeast ribosomal protein L25 is controlled principally by rapid degradation of excess protein. *Curr. Genet.* 10:733–739.

Fahrenkrog, B., E. C. Hurt, U. Aebi, and N. Pante. 1998. Molecular architecture of the yeast nuclear pore complex—localization of Nsp1p subcomplexes. *J. Cell Biol.* 143:577–588.

Fischer, U., E. Darzynkiewicz, S. M. Tahara, N. A. Dathan, R. Lührmann, and I. W. Mattaj. 1991. Diversity in the signals required for nuclear accumulation of U snRNPs and variety in the pathways of nuclear transport. *J. Cell Biol.* 113:705–714.

Fischer, U., J. Huber, W. C. Boelens, I. W. Mattaj, and R. Lührmann. 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82:475–483.

Fried, H. M. 1993. Nucleocytoplasmic transport in ribosome biogenesis, p. 257-267. *In* A. J. P. Brown, M. F. Tuite, and J. E. G. McCarthy (ed.), *Protein Synthesis and Targeting in Yeast.* Springer-Verlag, Berlin, Germany.

Goldberg, M. W., and T. D. Allen. 1995. Structural and functional organization of the nuclear envelope. *Curr. Opin. Cell Biol.* 7: 301–309.

Görlich, D., and I. W. Mattaj. 1996. Nucleocytoplasmic transport. *Science* 271:1513–1518.

Görlich, D., F. Vogel, A. D. Mills, E. Hartmann, and R. A. Laskey. 1995. Distinct functions for the two importin subunits in nuclear protein import. *Nature* 377:246–248.

Gritz, L., N. Abovich, J. L. Teem, and M. Rosbash. 1985. Post-transcriptional regulation and assembly into ribosomes of a *Saccharomyces cerevisiae* ribosomal protein–β-galactosidase fusion. *Mol. Cell. Biol.* 5:3436–3442.

Hall, M. N., L. Hereford, and I. L. Herskowitz. 1984. Targeting of *Escherichia coli* β-galactosidase to the nucleus in yeast. *Cell* 36:1057–1065.

Huber, J., U. Cronshagen, M. Kadokura, C. Marshallsay, T. Wada, M. Sekine, and R. Lührmann. 1998. Snurportin1, an M(3)G-Cap-specific nuclear import receptor with a novel domain structure. *EMBO J.* 17:4114–4126.

Jäkel, S., and D. Görlich. 1998. Importin beta, Transportin, Ranbp5 and Ranbp7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO J.* 17:4491–4502.

Jeeninga, R. E., J. Venema, and H. A. Raué. 1996. Rat RL23a ribosomal protein efficiently competes with its *Saccharomyces cerevisiae* L25 homologue for assembly into 60 S subunits. *J. Mol. Biol.* 263:648–656.

Kalderon, D., E. L. Roberts, W. D. Richardson, and A. E. Smith. 1984. A short amino acid sequence able to specify nuclear location. *Cell* 39:499–509.

Kobe, B. 1999. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat. Struct. Biol.* 6:388–397.

Kruiswijk, T., R. J. Planta, and J. M. Krop. 1978. The course of assembly of ribosomal subunits in yeast. *Biochim. Biophys. Acta* 517:378–389.

Mager, W. H., R. J. Planta, J. P. G. Ballesta, J. C. Lee, K. Mizuta, K. Suzuki, J. R. Warner, and J. Woolford. 1997. A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 25:4872–4875.

Maicas, E., F. G. Pluthero, and J. D. Friesen. 1988. The accumulation of three yeast ribosomal proteins under conditions of excess mRNA is determined primarily by fast protein decay. *Mol. Cell. Biol.* 8:169–175.

Mattaj, I. W., and L. Englmeier. 1998. Nucleocytoplasmic transport—the soluble phase. *Annu. Rev. Biochem.* 67:265–306.

Melchior, F., and L. Gerace. 1995. Mechanisms of nuclear protein import. *Curr. Opin. Cell. Biol.* 7:310–318.

Mélèse, T., and Z. Xue. 1995. The nucleolus: an organelle formed by the act of building a ribosome. *Curr. Opin. Cell Biol.* 7:319–324.

Michael, W. M., M. Y. Choi, and G. Dreyfuss. 1995. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* 83:415–422.

Moore, M. S., and G. Blobel. 1994. Purification of a Ran-interacting protein that is required for protein import into the nucleus. *Proc. Natl. Acad. Sci. USA* 91:10212–10216.

Moreland, R. B., H. G. Nam, L. M. Hereford, and H. M. Fried. 1985. Identification of a nuclear localization signal of a yeast ribosomal protein. *Proc. Natl. Acad. Sci. USA* 82:6561–6565.

Nehrbass, U., and G. Blobel. 1996. Role of the nuclear transport factor p10 in nuclear import. *Science* 272:120–122.

Nehrbass, U., E. Fabre, S. Dihlmann, W. Herth, and E. C. Hurt. 1993. Analysis of nucleocytoplasmic transport in a thermosensitive mutant of nuclear pore protein NSP1. *Eur. J. Cell Biol.* 62:1–12.

Osborne, M. A., and P. A. Silver. 1993. Nucleocytoplasmic transport in the yeast *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* 62:219–254.

Panté, N., and U. Aebi. 1996. Molecular dissection of the nuclear pore complex. *Crit. Rev. Biochem. Mol. Biol.* 31:153–199.

Paschal, B. M., and L. Gerace. 1995. Identification of NTF2, a cytosolic factor for nuclear import that interacts with nuclear pore complex protein p62. *J. Cell Biol.* 129:925–937.

Planta, R. J. 1997. Regulation of ribosome synthesis in yeast. *Yeast* 13:1505–1518.

Pollard, V. W., W. M. Michael, S. Nakielny, M. C. Siomi, F. Wang, and G. Dreyfuss. 1996. A novel receptor-mediated nuclear protein import pathway. *Cell* 86:985–994.

Quaye, I. K. E., S. Toku, and T. Tanaka. 1996. Sequence requirement for nucleolar localisation of rat ribosomal protein L31. *Eur. J. Cell Biol.* 69:151–155.

Raué, H. A., and R. J. Planta. 1991. Ribosome biogenesis in yeast. *Prog. Nucleic Acid Res. Mol. Biol.* 41:89–129.

Roberts, B. L., W. D. Richardson, and A. E. Smith. 1987. The effect of protein context on nuclear location signal function. *Cell* 50:465–475.

Rosenblum, J. S., L. F. Pemberton, and G. Blobel. 1997. A nuclear import pathway for a protein involved in tRNA maturation. *J. Cell Biol.* 139:1655–1661.

Rout, M. P., G. Blobel, and J. D. Aitchison. 1997. A distinct nuclear import pathway used by ribosomal proteins. *Cell* 89:715–725.

Russo, G., G. Ricciardelli, and C. Pietropaolo. 1997. Different domains cooperate to target the human ribosomal L7a protein to the nucleus and to the nucleoli. *J. Biol. Chem.* 272:5229–5235.

Rutgers, C. A., P. J. Schaap, J. Van 't Riet, C. L. Woldringh, and H. A. Raué. 1990. *In vivo* and *in vitro* analysis of structure-function relationships in ribosomal protein L25 from *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* 1050:74–79.

Rutgers, C. A., J. M. J. Rientjes, J. Van 't Riet, and H. A. Raué. 1991. rRNA binding domain of yeast ribosomal protein L25. Identification of its borders and a key leucine residue. *J. Mol. Biol.* 218:375–385.

Schaap, P. J., J. Van 't Riet, C. L. Woldringh, and H. A. Raué. 1991. Identification and functional analysis of the nuclear localization signals of ribosomal protein L25 from *Saccharomyces cerevisiae*. *J. Mol. Biol.* 221:225–237.

Scheer, U., and D. Weisenberger. 1994. The nucleolus. *Curr. Opin. Cell Biol.* 6:354–359.

Schlenstedt, G., E. Smirnova, R. Deane, J. Solsbacher, U. Kutay, D. Gorlich, H. Ponstingl, and F. R. Bischoff. 1997. Yrb4p, a yeast Ran-GTP-binding protein involved in import of ribosomal protein L25 into the nucleus. *EMBO J.* 16:6237–6249.

Schmidt, C., E. Lipsius, and J. Kruppa. 1995. Nuclear and nucleolar targeting of human ribosomal protein S6. *Mol. Biol. Cell* 6: 1875–1885.

Schmidt-Zachmann, M. S., and E. A. Nigg. 1993. Protein localization to the nucleolus—a search for targeting domains in nucleolin. *J. Cell Sci.* 105:799–806.

Shimmin, L. C., G. Ramirez, A. T. Matheson, and P. P. Dennis. 1989. Sequence alignment and evolutionary comparison of the L10 equivalent and L12 equivalent ribosomal proteins from archaebacteria, eubacteria, and eucaryotes. *J. Mol. Evol.* 29:448–462.

Timmers, A. C. J., R. Stuger, P. J. Schaap, J. Van 't Riet, and H. A. Raué. 1999. Nuclear and nucleolar localization of *Saccharomyces cerevisiae* ribosomal proteins S22 and S25. *FEBS Lett.* 452:335–340.

Tollervey, D. 1996. Trans-acting factors in ribosome synthesis. *Exp. Cell Res.* 229:226–232.

Underwood, M., and H. M. Fried. 1990. Characterization of nuclear localizing sequences derived from yeast ribosomal protein L29. *EMBO J.* 9:91–99.

Vodkin, M. H., R. J. Novak, and G. L. McLaughlin. 1996. Database searches with multiple oligopeptides containing ambiguous residues. *BioTechniques* 21:1116–1117.

Wagner, P., and M. N. Hall. 1993. Nuclear protein transport is functionally conserved between yeast and higher eukaryotes. *FEBS Lett.* 321:261–266.

Warner, J. R. 1979. Distribution of newly formed ribosomal proteins in HeLa cell fractions. *J. Cell Biol.* 80:767–772.

Warner, J. R., G. Mitra, F. Schwindiger, M. Studeny, and H. M. Fried. 1985. *Saccharomyces cerevisiae* coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover. *Mol. Cell. Biol.* 5:1512–1521.

Weis, K. 1998. Importins and exportins—how to get in and out of the nucleus. *Trends Biochem. Sci.* 23:235.

Wen, W., J. L. Meinkoth, R. Y. Tsien, and S. S. Taylor. 1995. Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82:463–473.

Wozniak, R. W., M. P. Rout, and J. D. Aitchison. 1998. Karyopherins and kissing cousins. *Trends Cell Biol.* 8:184–188.

Xue, Z., and T. Mélèse. 1994. Nucleolar proteins that bind NLSs: a role in nuclear import or ribosome biogenesis? *Trends Cell Biol.* 4:414–417.