The EMBO Journal Vol.16 No.4 pp.793–806, 1997

In vivo selection of RNAs that localize in the nucleus

Christian Grimm[1], Elsebet Lund and James E.Dahlberg

Department of Biomolecular Chemistry, 1300 University Avenue, University of Wisconsin, Madison, WI 53706, USA

Present address: Department of Ophthalmology, University Hospital-Zürich, 8091 Zürich, Switzerland

[1]Corresponding author

Nuclear localization of an RNA is affected by *cis*-acting elements (NLEs) that lead to nuclear import or retention or to blockage of export from the nucleus. To identify such elements, we selected and analyzed transcripts that localized in the nuclei of *Xenopus laevis* oocytes. The RNAs were isolated from a collection of m^7G-capped RNAs in which a combinatorial library ($n = 20$) of sequences had been inserted. One class of selected RNAs (Sm^+) had a consensus Sm binding site (AAUUUUUGG) and bound Sm proteins in the cytoplasm; these RNAs resembled small nuclear RNAs like U1 and U5 RNAs in their bi-directional nucleo–cytoplasmic transport and their 5′-cap hypermethylation. Another class, Sm^- RNAs, contained sequences that masked the m^7G-caps of the RNAs and promoted interaction with La protein. These RNAs were retained within nuclei after nuclear injection and were imported when injected into the cytoplasm. Their nuclear import and retention were independent of a 5′-cap, required an imperfect double-stranded stem near the 5′ end, and depended on interaction with La protein. Import of the Sm^- RNAs, while using the import pathway of proteins, was distinct from that of U6 RNA.

Keywords: La protein/nuclear localization elements (NLEs)/RNA transport/selection of RNAs *in vivo*/Sm consensus site

Introduction

The distribution of RNAs within various sub-cellular compartments results from a balance of export, import and retention. In several instances, *cis*-acting sequences in the RNA and *trans*-acting cellular factors have been shown to contribute to these processes. However, relatively few of the actual proteins and RNA signals are known. Intracellular RNA transport is an active process that often involves translocation of the RNAs (or RNPs) across the nuclear envelope, through the nuclear pore complexes (NPCs; Davis, 1995). Whereas most RNAs (mRNA, tRNA and scRNA) are transported unidirectionally from the nucleus to the cytoplasm, the precursors of many small nuclear RNAs (pre-snRNAs) and 5S ribosomal RNA are translocated through the pores in both directions (Zapp, 1992; Izaurralde and Mattaj, 1992, 1995). Other nuclear RNAs, like the small nucleolar RNAs (snoRNAs) and U6 snRNA are not exported to the cytoplasm, but are retained in the nucleus (Vankan *et al.*, 1990; Terns and Dahlberg, 1994; Boelens *et al.*, 1995; Terns *et al.*, 1995).

Several pathways exist for nuclear import of RNA (Fischer *et al.*, 1991; Michaud and Goldfarb, 1992). Efficient import of most spliceosomal snRNAs into nuclei, as snRNPs, requires two *cis*-acting elements: a binding site for the Sm proteins and a trimethylguanosine cap structure (reviewed in Nigg *et al.*, 1991; Izaurralde and Mattaj, 1992). In contrast, U6 snRNA (Hamm and Mattaj, 1989) and 5S rRNA (Allison *et al.*, 1993) lack both of these signals, but nevertheless can be imported into nuclei of *Xenopus* oocytes, bound to proteins that contain nuclear localization signals (NLSs) (Fischer *et al.*, 1991).

Likewise, several distinct pathways are likely to exist also for RNA export, as demonstrated by the lack of competition in this process between different classes of RNA molecules (Jarmolowski *et al.*, 1994; Pokrywka and Goldfarb, 1995) and by the differential inhibition of export by various inhibitors of NPC function (E.Lund and J.E.Dahlberg, in preparation; Powers *et al.*, 1997). However, several RNAs also share aspects of common export pathways. For example, saturation of the Rev-mediated export of RRE-containing RNAs (Fischer *et al.*, 1994) affects export of pre-snRNAs and 5S RNA, but not that of mRNAs and tRNAs (Fischer *et al.*, 1995).

The monomethylated cap structure of pre-snRNAs and mRNAs facilitates export of these RNAs to the cytoplasm (Hamm and Mattaj, 1990; Terns *et al.*, 1993; Jarmolowski *et al.*, 1994). This m^7G-cap is recognized by the proteins of the cap binding complex (CBC) which mediates export, at least of snRNAs (Izaurralde *et al.*, 1995; reviewed in Görlich and Mattaj, 1996). Other structural elements of these RNAs are also important for export (Hamm and Mattaj, 1990; Eckner *et al.*, 1991; Terns *et al.*, 1993; Jarmolowski *et al.*, 1994). Several RNA binding proteins have been implicated in the export of different classes of RNAs, but only CBP20 and CBP80 (the proteins of CBC; Izaurralde *et al.*, 1995), and the Rev protein (Fischer *et al.*, 1995), have been shown to promote export of bound RNA (reviewed in Izaurralde and Mattaj, 1995; Görlich and Mattaj, 1996).

The lectin wheat germ agglutinin (WGA), which binds to *N*-acetylglucosamine residues of NPC proteins, is an effective inhibitor of many types of nucleo–cytoplasmic transport (reviewed by Forbes, 1992). Import of most proteins carrying an NLS is inhibited by WGA treatment (Finlay *et al.*, 1987; Dabauvalle *et al.*, 1988) as is the import of U6 RNA, which apparently occurs by the same pathway (Fischer *et al.*, 1991; this study). In contrast, import of snRNPs containing Sm proteins is relatively insensitive to the lectin. Export of most RNAs is inhibited

by WGA (E.Lund and J.E.Dahlberg, manuscript in preparation), but some of these effects may be secondary to inhibition of import of proteins that are required as carriers.

Several mechanisms could account for the localization of RNAs in nuclei. RNAs may contain sequences that lead to their retention within the nucleus or to their import from the cytoplasm, as has been demonstrated for various classes of snRNAs. Examples include a sequence element common to most snoRNAs (box D), that is essential for nuclear retention of U8 snoRNA (Terns *et al.*, 1995), and the binding site for Sm proteins of the spliceosomal RNAs U1, U2, U4 and U5, which promotes import of these molecules into the nucleus. Alternatively, structures or sequences that ordinarily would direct export to the cytoplasm may either be absent or masked in other RNAs that localize in nuclei. Here we define the term nuclear localization element (NLE) as any *cis*-acting RNA sequence or structural feature that promotes localization of the RNA in the nucleus.

In this study we developed and used an iterative selection to identify signals and mechanisms that contribute to the localization of RNAs within nuclei. At least two classes of molecules were isolated that had the capacity to be localized within nuclei. One class of RNAs contains a consensus sequence for the Sm binding site (AAUUUUUGG) that promotes nuclear localization of snRNAs; isolation of this expected class of molecules validated the method. Another class contains a structural motif that participates in nuclear localization of Sm$^-$ RNAs. Interaction of this structure with La protein appears to promote retention of the RNA in the nucleus by masking the 5′-cap, which otherwise acts as an export signal; furthermore, by binding to La protein in the cytoplasm the RNA can be imported in an NLS-dependent fashion.

Results

In vivo selection of RNA elements that promote nuclear localization

To select for NLEs (both nuclear import and nuclear retention signals), we used a short derivative of U1 snRNA as the carrier molecule for a 20 nucleotide long, randomized sequence (N20; Figure 1A; see also Grimm *et al.*, 1995). Since the carrier moiety of this *in vitro* transcribed RNA contains a strong nuclear export signal in the form of the m^{7}G-cap (Hamm and Mattaj, 1990; Izaurralde *et al.*, 1995) but no nuclear import signal, most of the stable molecules of the starting pool localized in the cytoplasm within 20 h after they had been injected into nuclei (Figure 1B; see also Figure 5A, top panel). However, some RNAs received a sequence through the N20 insert that caused nuclear localization of the RNA, and therefore could be isolated from nuclei and amplified through multiple rounds of selection. The percentage of RNAs in the experimental RNA pool that localized in the nucleus increased with each round, indicating a gradual enrichment of the pool with RNAs containing NLEs. In contrast, the distribution of several non-selected control snRNAs remained constant throughout the selection procedure (Figure 1B).

One known NLE is the Sm site of snRNAs to which Sm proteins bind, promoting snRNP import (Fischer *et al.*, 1993). Therefore, if this selection process works as

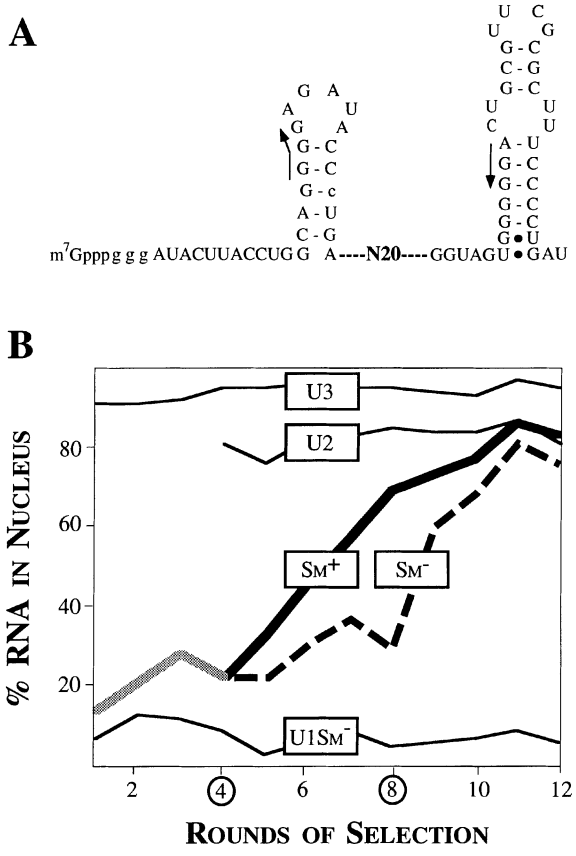

Fig. 1. Selection of RNAs with nuclear localization signals. (**A**) Structure of the RNA used for the selection of NLEs. Fixed nucleotide sequences from the 5′ and 3′ ends of U1 snRNA are in upper case letters, and nucleotides in lower case letters indicate changes in the U1 sequences that were made to ensure efficient *in vitro* transcription by T7 RNA polymerase (addition of Gs at the 5′ end) and increased *in vivo* stability of the RNA (A to C change in the 5′ stem). N20 represents the 20 nucleotide randomized sequence. Arrows indicate the endpoints of the primers used for reverse transcription and PCR amplification of the RNAs. (**B**) Enrichment of the pools for RNAs that localize in the nucleus. The percentages of RNAs that localized in the nucleus at 20–24 h after nuclear injection were determined by PhosphorImager analysis of gels run after each round of selection and were expressed as [N/(N+C)]×100. Thick lines, experimental RNA pools; thin lines, control U1Sm$^-$, U2 and U3 RNAs that were exported to the cytoplasm, exported and imported back into the nucleus and retained in the nucleus, respectively. Sm$^-$ and Sm$^+$ pools of RNAs were generated at rounds 4 and 8 (circles) by immunoprecipitation of nuclear extracts with anti-Sm antibodies; the RNAs in the precipitate (Sm$^+$) and supernatant (Sm$^-$) fractions were injected separately after round 4.

expected, one class of the molecules we isolate should contain Sm protein binding sites. To distinguish between molecules that contain such Sm binding sites and molecules with other NLEs, we used anti-Sm antibodies to precipitate RNAs bound to Sm proteins (at rounds 4 and 8, as indicated by circles, Figure 1B). The resulting Sm$^-$ and Sm$^+$ RNA pools were treated separately after round 4. After 12 rounds of injection and selection, RNAs of both pools localized in the nucleus as efficiently as did U2 snRNA.

Similarities in transport and maturation of Sm$^+$ nuclear localized RNAs (NL-RNAs)

Individual cDNAs, made from nuclear localized Sm$^+$ NL-RNAs were cloned and sequenced (Figure 2A). All of the

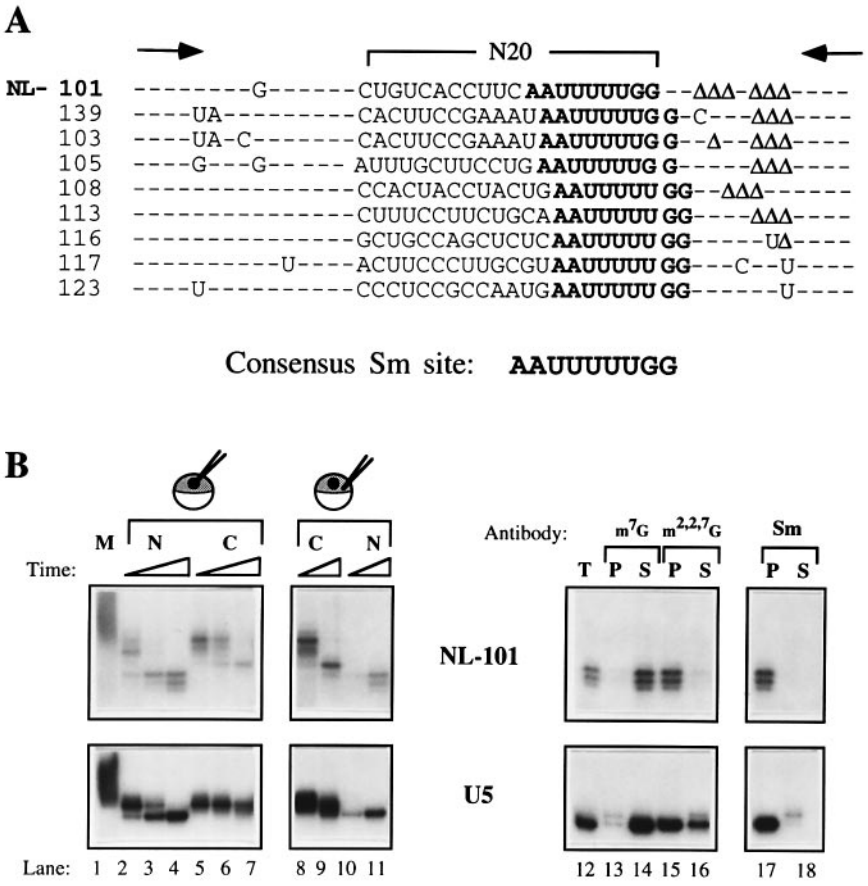

Fig. 2. Selected Sm^+ NL-RNAs. (**A**) Sequences of Sm^+ NL-RNAs after 12 rounds of selection–amplification. The randomized region (N20), nucleotide changes and deletions (Δ) within the fixed sequence of the carrier RNA (dashed lines) are indicated; the 3′ ends of the primers used for reverse transcription and PCR amplification are shown by arrows. Sequences of Sm protein binding sites are shown in bold letters. NL-101 RNA (top line) was used for testing the functionality of the selected Sm sites (see text). (**B**) Transport and maturation of NL-101 RNA. 1–2 fmol each of ^{32}P-labeled, m^7G-capped NL-101 and U5 RNAs were co-injected into nuclei (lanes 2–7) or cytoplasms (lanes 8–11) of oocytes. Oocytes were fractionated into nuclei (N) and cytoplasms (C) at 2 h (lanes 2 and 5), 3 h (lanes 8 and 10), 5 h (lanes 3 and 6) or 24 h (lanes 4, 7, 9 and 11) after injection. Total RNAs of each fraction were analyzed by denaturing polyacrylamide gel electrophoresis (PAGE) and autoradiography of the gel. M, RNAs prior to injection. Cap hypermethylation and RNP formation was assayed 24 h after nuclear injection (cf. lane 4) by immunoprecipitation of total nuclear RNAs (T, lane 12) or nuclear extracts using antibodies specific for the mono- (m^7G; lanes 13 and 14) or hypermethylated ($m^{2,2,7}$G; lanes 15 and 16) cap structures, or for Sm proteins (lanes 17 and 18), respectively. RNAs in the precipitate (P) and supernatant (S) fractions were analyzed by denaturing PAGE.

Sm^+ RNAs contained a consensus sequence (AAUUUUUGG) which resembled a typical Sm site; ~10% of the cloned RNAs were Sm^- contaminants that did not localize in nuclei, when tested individually (data not shown). Interestingly, this consensus sequence is always located close to the U1 3′ stem–loop of the carrier RNA. This preference for position is in agreement with the finding that the function of an Sm site in nucleo–cytoplasmic trafficking is dependent on the presence and nature of adjacent stem–loop structures (Jarmolowski and Mattaj, 1993). The occasional deletions and nucleotide changes outside the randomized region are probably caused by the amplification method, since they occur at positions close to the endpoints of the primers; whether they are important parts of the selected sequences of Sm^+ NL-RNAs is unclear.

The individual Sm^+ NL-RNAs follow the transport and maturation pathway of spliceosomal snRNAs like U1 and U5. When injected into nuclei of oocytes, m^7G-capped NL-101 RNA was exported to the cytoplasm where its 3′ end was shortened (Figure 2B, lanes 2–7). After import back into the nucleus, the 3′ end was further trimmed by one or two nucleotides (compare lanes 7 and 4, or lanes 9 and 11), as was observed also with U1 RNA (Yang *et al.*, 1992) and U5 RNA. When injected into the cytoplasm (lanes 8–11), similar 3′ end shortenings occurred prior to (lane 9) and after (lane 11) nuclear import. Immunoprecipitations (lanes 13–18) of RNAs or RNPs present in the nucleus at 24 h after nuclear injection (lane 4 and lane 12) demonstrated that all of the NL-101 RNA was associated with Sm proteins (lanes 17 and 18) and had acquired a hypermethylated $m^{2,2,7}$G-cap structure (lanes 13–16). Therefore, transport, 3′ end trimming, Sm protein binding and cap hypermethylation of NL-101 RNA closely resembled comparable steps in the maturation of snRNAs and their precursors. These results show that the method used here is capable of selecting RNA molecules on the basis of their abilities to be transported within cells.

Three groups of secondary structures common to Sm^- NL-RNAs

In contrast to Sm^+ NL-RNAs, no strongly conserved sequence motif was evident in the N20 region of the selected Sm^- NL-RNAs, other than a bias against

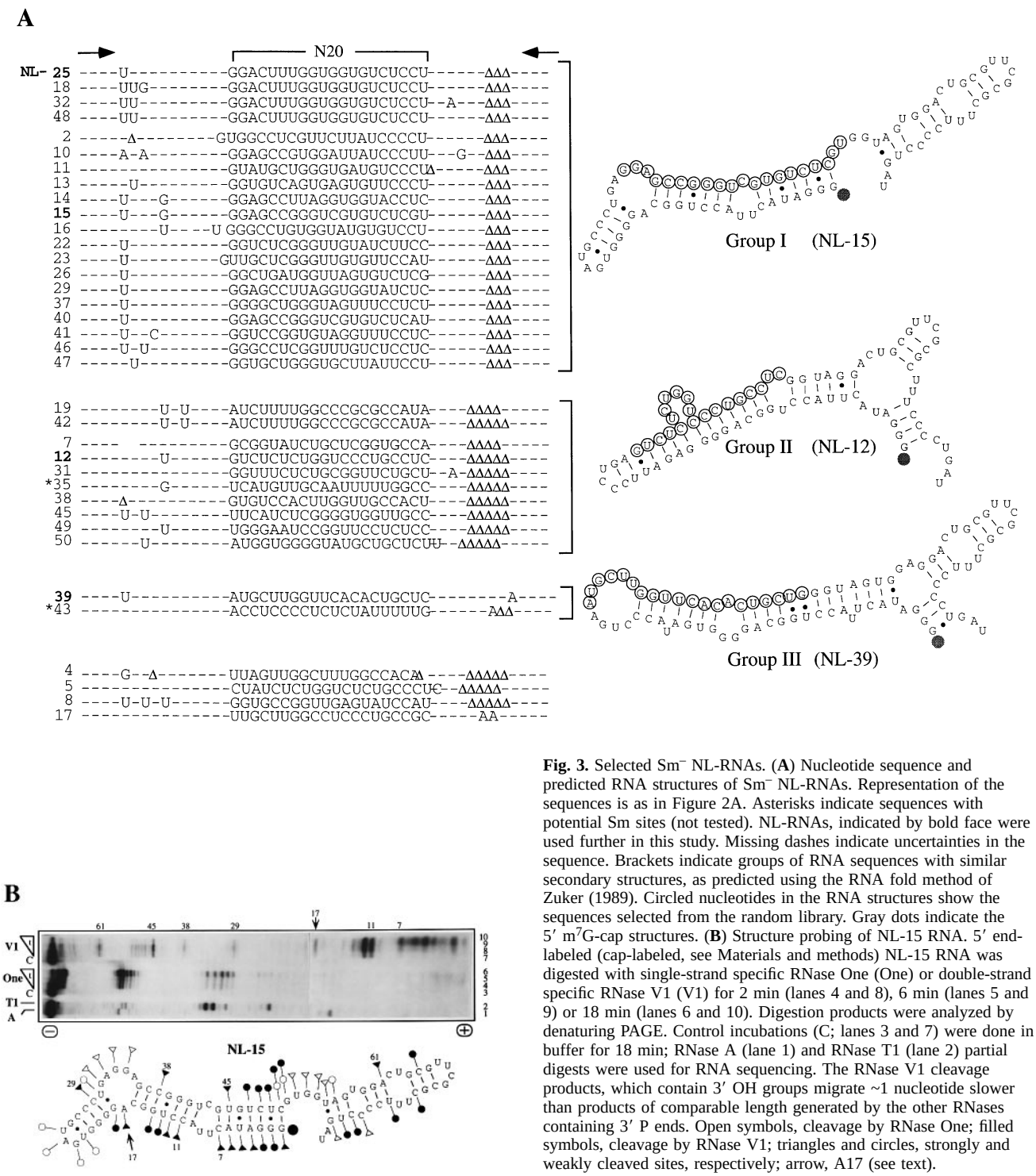

Fig. 3. Selected Sm⁻ NL-RNAs. (**A**) Nucleotide sequence and predicted RNA structures of Sm⁻ NL-RNAs. Representation of the sequences is as in Figure 2A. Asterisks indicate sequences with potential Sm sites (not tested). NL-RNAs, indicated by bold face were used further in this study. Missing dashes indicate uncertainties in the sequence. Brackets indicate groups of RNA sequences with similar secondary structures, as predicted using the RNA fold method of Zuker (1989). Circled nucleotides in the RNA structures show the sequences selected from the random library. Gray dots indicate the 5′ m^7G-cap structures. (**B**) Structure probing of NL-15 RNA. 5′ end-labeled (cap-labeled, see Materials and methods) NL-15 RNA was digested with single-strand specific RNase One (One) or double-strand specific RNase V1 (V1) for 2 min (lanes 4 and 8), 6 min (lanes 5 and 9) or 18 min (lanes 6 and 10). Digestion products were analyzed by denaturing PAGE. Control incubations (C; lanes 3 and 7) were done in buffer for 18 min; RNase A (lane 1) and RNase T1 (lane 2) partial digests were used for RNA sequencing. The RNase V1 cleavage products, which contain 3′ OH groups migrate ~1 nucleotide slower than products of comparable length generated by the other RNases containing 3′ P ends. Open symbols, cleavage by RNase One; filled symbols, cleavage by RNase V1; triangles and circles, strongly and weakly cleaved sites, respectively; arrow, A17 (see text).

adenosine residues. As with the Sm^+ NL-RNAs, almost all of the Sm⁻ NL-RNAs had several nucleotides deleted 3′ of the N20 region (Figure 3A). However, a variant of NL-15 RNA containing the complete sequence of the carrier RNA localized efficiently in the nuclei of oocytes demonstrating that the nuclear localization of the selected RNA was not dependent on the deletion of these nucleotides (data not shown).

Using the RNA M-fold method (Zuker, 1989), we categorized the selected Sm⁻ NL-molecules into three groups according to possible secondary structures (Figure 3A). All of the proposed structures contained strong stems at (group I and II) or near (group III) the 5′ end and most nucleotides derived from the randomized sequence (circled in Figure 3A) were in strongly base-paired regions. The possible structures of four RNAs (NL-4, -5, -8 and -17; bottom of Figure 3A) did not fit any of the three categories; three of these RNAs (NL-4, -5 and -17) were tested individually and showed only inefficient nuclear localization (data not shown); they were not tested further.

The proposed structures of the Sm⁻ NL-RNAs are supported by the digestion pattern of NL-15 RNA (a

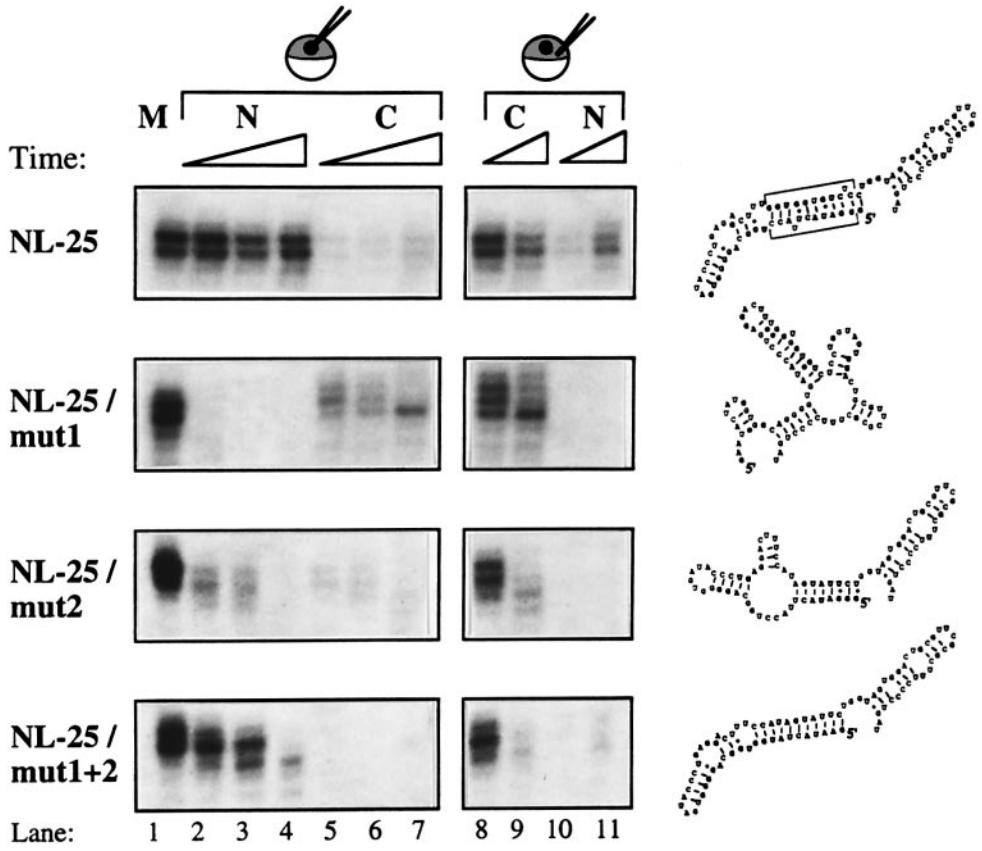

Fig. 4. Nuclear localization and stability of wild-type and mutant NL-25 RNAs. Wild-type (top panel) and mutant m^7G-capped NL-25 RNAs were injected into nuclei (lanes 2–7) or cytoplasms (lanes 8–11) of oocytes. Oocytes were fractionated and RNAs analyzed as in Figure 2B. The RNA secondary structures were predicted as in Figure 3A. Brackets in NL-25 mark the sequences that are altered in the mutant RNAs. NL-25/mut1: 5′-proximal half of the stem mutated; NL-25/mut2: 5′-distal half of the stem mutated; NL-25/mut1+2: both halves mutated (compensatory mutations). M, RNAs prior to injection; N, nuclear RNAs; C, cytoplasmic RNAs.

member of group I) produced by both the single-strand specific RNase One and the double-strand specific RNase V1 (Figure 3B). Cleavage at A17 by RNase V1 might indicate stacking of A17 between the two stems on either side.

The RNAs shown in Figure 3A had been selected by 11 rounds of nuclear injection, followed by a 12th round of cytoplasmic injection to select for RNAs that also can be imported from the cytoplasm. However, a comparison of these RNAs to RNAs that had not been subjected to this last selection step did not reveal any obvious differences in type of sequence (low content of adenosine residues), RNA structure or transport behavior (data not shown).

Importance of the 5′ stem for nuclear localization of Sm⁻ NL-RNAs

To determine whether the structures of Sm⁻ NL-RNAs are important for nuclear localization, we disrupted the 5′ stem of NL-25 RNA by mutagenesis of the DNA template. Sequence alterations in either one side (NL-25/mut1) or the other side (NL-25/mut2) of the stem led to similar decreases in RNA stability and caused cytoplasmic accumulation of the mutant RNAs (Figure 4, second and third panels). In contrast, the RNA with compensatory mutations that resulted in reformation of a strong 5′ stem (NL-25/mut1+2; Figure 4, last panel) again was localized in the nucleus. However, this latter RNA was less stable and less efficiently imported than the original NL-25 RNA. We note that unlike NL-25/mut1+2 RNA, all of the selected Sm⁻ NL-RNAs contain imperfect 5′ stem structures (cf. Figure 3A and data not shown), which might be important for RNA–protein interactions needed for stabilization and nuclear localization of the RNAs (see Discussion). Finally, NL-25/mut1 RNA does not localize in the nucleus even though it contains the sequence selected from the random library, showing that the selected primary sequence alone is not sufficient for retention and import; instead, the sequence probably is important because it contributes to the formation of a specific RNA structure.

The transport behaviors of individual members of the three structural groups supported the importance of the 5′ stem for nuclear localization of Sm⁻ NL-RNAs (Figure 5A). Whereas most molecules of the original RNA pool localized in the cytoplasm (top panel), the *in vivo* selected Sm⁻ RNAs of group I and group II were retained in the nucleus and were imported when injected into the cytoplasm (second and third panels). NL-39 RNA (group III) however, apparently reached an equilibrium between nucleus and cytoplasm 24 h after injection (last panel, compare lanes 4 and 7). We note that the 5′ end of NL-39 RNA is only weakly base-paired (G–U pairing), in contrast to the 5′ ends of RNAs in groups I and II (Figure 3A).

The 5′ end of NL-15 RNA appeared to be masked since neither the intracellular localization nor the stability

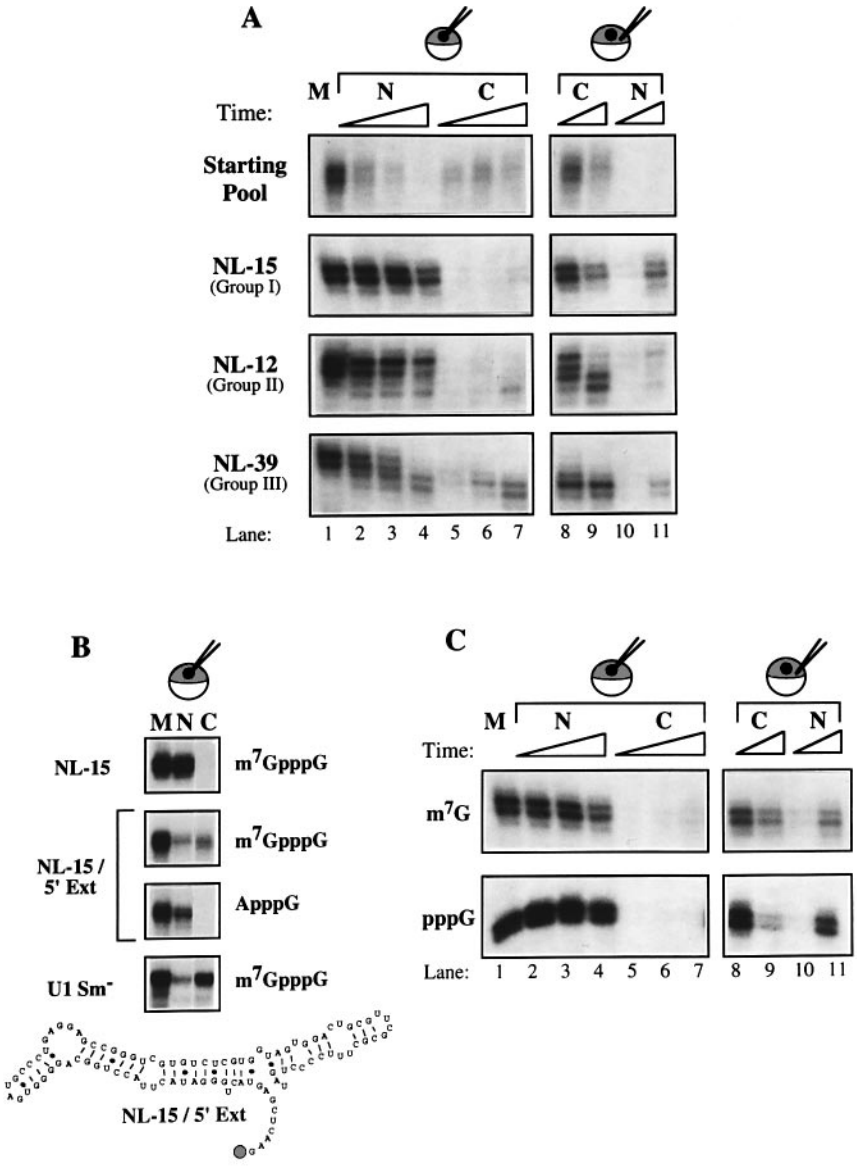

Fig. 5. Nuclear localization of individual Sm^- NL-RNAs and sequestration of the m^7G-cap as an export signal. (**A**) Nucleo–cytoplasmic distribution of m^7G-capped RNAs from the starting pool (round 1 of the selection) and from the three structural groups of selected Sm^- NL-RNAs (round 12, see Figure 3A). Oocyte injection, fractionation and RNA analysis was as in Figure 2B. (B and C) Role of the m^7G-cap as an export signal. (**B**) Nucleo–cytoplasmic distributions of m^7G- and ApppG-capped NL-15 RNA with a 5′ extension (NL-15/5′Ext) were assayed 2 h after nuclear injection and compared with those of m^7G-capped wild-type NL-15 and U1 Sm^- RNAs. M, RNAs prior to injection; N, nuclear RNAs; C, cytoplasmic RNAs. (**C**) Comparison of nuclear localizations of capped (m^7G-) and uncapped (pppG-) NL-15 RNAs. RNAs were analyzed as in (A).

of the RNA was affected by the presence or absence of a 5′-cap (Figure 5C). Also, the capped RNA was poorly precipitated by cap-specific antibodies (data not shown). To test if the proximity of the 5′-cap to the body of the structured RNA interfered with recognition of the m^7G-cap as an export signal, we extended the 5′ end of the RNA with a short unstructured sequence (Figure 5B, NL-15/5′Ext). Cap-specific antibodies could now access the cap and efficiently precipitate the RNA (data not shown). Since the extension should make the cap accessible to other proteins as well, it should allow interaction between the m^7G-cap and proteins of the CBC (Izaurralde *et al.*, 1995) and consequently promote export. As a control, we injected NL-15/5′Ext RNA bearing an ApppG (A-cap) which is a poor substrate for CBC binding. Unlike m^7G-capped wild-type NL-15 RNA (panel 1), m^7G-capped NL-15/5′Ext RNA was exported as efficiently as U1 snRNA (panels 2 and 4); in contrast, the A-capped RNA remained in the nucleus (panel 3). These results indicate that localization of wild-type NL-15 RNA in the nucleus is due, at least in part, to the inability of the 5′ m^7G-cap to interact with CBC. This conclusion is supported by our finding that the efficiency of UV-crosslinking between a component of CBC and the 5′ m^7G-cap of NL-15 RNA is strongly enhanced when the RNA has the 5′ extension (see Figure 8A, below).

Complexes between NL-15 RNA and nuclear proteins

To test whether NL-15 RNA is retained in the nucleus as a consequence of its binding to a nuclear protein, high amounts of unlabeled competitor RNAs were injected to

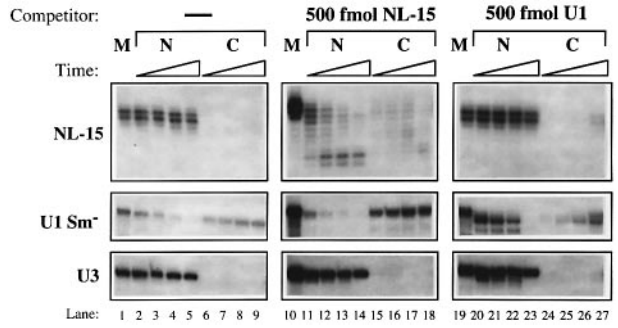

Fig. 6. Saturation of nuclear retention of NL-15 RNA. 1 to 2 fmol of ^{32}P-labeled m^{7}G-capped NL-15, U1 Sm$^-$ and U3 RNAs were co-injected into nuclei of *Xenopus* oocytes in the absence (lanes 1–9) or presence of 500 fmol of unlabeled m^{7}G-capped NL-15 (lanes 10–18) or U1 (lanes 19–27) competitor RNAs. Nucleo–cytoplasmic distribution was tested 1 h (lanes 2, 6, 11, 15, 20 and 24), 2 h (lanes 3, 7, 12, 16, 21 and 25), 4 h (lanes 4, 8, 13, 17, 22 and 26) and 24 h (lanes 5, 9, 14, 18, 23 and 27) after injection as in Figure 2B. M, RNAs prior to injection; N, nuclear RNAs; C, cytoplasmic RNAs.

saturate a potential nuclear retention site(s) (Figure 6). Injection of 500 fmol of NL-15 competitor RNA caused destabilization and cytoplasmic accumulation of labeled NL-15 RNA, but it had no effect on retention of U3 RNA. Likewise, injection of 500 fmol of U1 RNA did not affect nuclear retention of NL-15 RNA but it did saturate export of U1Sm$^-$ RNA (a mutant form of U1 RNA that is exported but cannot be re-imported into nuclei, Mattaj and de Robertis, 1985). Thus, nuclear retention of NL-15 RNA involves a specific and saturable factor(s) that is not required for U3 retention or U1 export.

To learn which nuclear proteins might interact with NL-15 RNA, we incubated the RNA in nuclear extracts and assayed for RNA–protein complex formation by native gel electrophoresis (Figure 7). The complex that formed between NL-15 RNA and a component(s) of the nuclear extract (Figure 7A, lane 2) was specific, since it could be competed efficiently by an excess of unlabeled NL-15 RNA (lanes 3–6), but only poorly by an unrelated RNA (lanes 11–14). U6 RNA also competed for complex formation (lanes 7–10); since U6 RNA binds La protein in nuclei of *Xenopus* oocytes (Terns *et al.*, 1992), we tested whether the complex formed with NL-15 RNA involved La protein. A variety of experiments shows that this is the case (Figure 7B–D): (i) NL-15 RNA did not form a complex in nuclear extracts that had been immunologically depleted of La protein (Figure 7B). (ii) The complex that formed in untreated extracts was supershifted by the addition of anti-La antibodies (Figure 7C, lanes 1–6); the supershift could be reversed by the addition of recombinant human La protein (data not shown). (iii) NL-15 RNA formed a similar complex in extracts made from *Escherichia coli* cells that expressed recombinant human La protein, but not in control *E.coli* extracts; this complex also was supershifted by anti-La antibodies (Figure 7C, lanes 7–10). (iv) NL-15 RNA that was injected in nuclei of *Xenopus* oocytes was coprecipitated with anti-La antibodies (data not shown). (iv) NL-15 RNA competed for a U6 complex formed in nuclear extracts (data not shown). (vi) High levels of NL-15 RNA dramatically accelerated export of hY1 RNA from the nucleus *in vivo* (Figure 7D); because a comparable

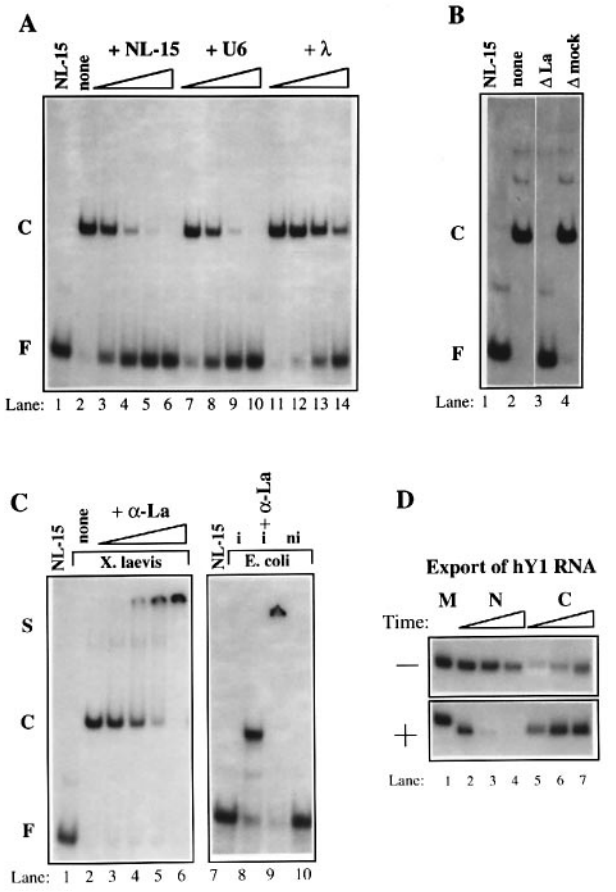

Fig. 7. Complex formation of NL-15 RNA with La protein. (**A**) Formation of complexes in nuclear extracts from *Xenopus* oocytes. 10 fmol of m^{7}G-capped NL-15 RNA was incubated either in buffer (lane 1) or in 0.5 oocyte equivalents of nuclear extract in the absence (lane 2) or presence of 2.5 ng (lanes 3, 7 and 11), 10 ng (lanes 4, 8 and 12), 50 ng (lanes 5, 9 and 13) or 200 ng (lanes 6, 10 and 14) of unlabeled competitor RNA and subsequently analyzed by native PAGE. λ competitor: 172 nucleotides long RNA made from λ DNA. (**B**) Formation of complexes in immunodepleted extracts. 10 fmol of m^{7}G-capped NL-15 RNA was incubated either in buffer (lane 1), in 0.5 oocyte equivalents of untreated nuclear extract (lane 2) or in extract immunodepleted with anti-La antibodies (lane 3) or total human IgGs (lane 4). (**C**) Supershifts of NL-15 complexes with anti-La antibodies. Formation of complexes of m^{7}G-capped NL-15 in nuclear extract from *Xenopus* oocytes (left panel) or extract from *E.coli* (right panel) as in (A). RNA was incubated either in buffer (lanes 1 and 7), in nuclear extract of *Xenopus* oocytes (lanes 2–6) or in extract of *E.coli* cells that have (i; lanes 8 and 9) or have not been induced (ni; lane 10) to express recombinant human La protein. Complexes formed in nuclear extracts from *Xenopus* oocytes were incubated further with increasing amounts of anti-La antibodies (B-103; lanes 3–6). The complex formed in extracts from *E.coli* that contained human La protein was incubated with the amount of anti-La antibodies (B-103) used in lane 6 (lane 9). F, free RNA; C, complex; S, supershift. (**D**) Acceleration of hY1 RNA export by high levels of NL-15 RNA. ^{32}P-labeled hY1 RNA was injected into nuclei of *Xenopus* oocytes in the absence (–; top panel) or presence (+; bottom panel) of 500 fmol of unlabeled NL-15 RNA. Oocytes were fractionated and RNAs analyzed as in Figure 2B.

acceleration of export of hY1 RNA occurs when hY1 RNA is prevented from binding to La protein by mutation (Simons *et al.*, 1996), the effect of NL-15 RNA is probably due to competition of the two RNAs for available La protein. We conclude that NL-15 RNA binds La protein in nuclei (and cytoplasms; see below) of *Xenopus* oocytes.

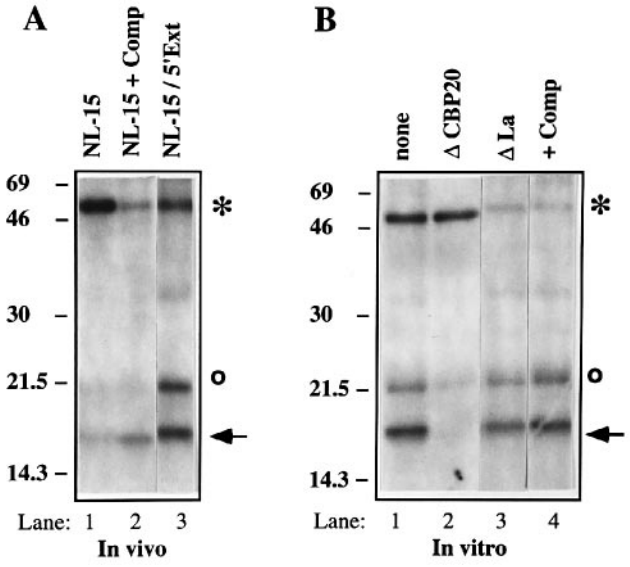

Fig. 8. Crosslinking of proteins to NL-15 RNA by UV light. (**A**) Crosslinking to RNA injected into nuclei of *Xenopus* oocytes. m^7G-cap labeled NL-15 (lanes 1 and 2) or NL-15/5′Ext (lane 3) RNA was injected into nuclei of oocytes. 15–30 min after injection 10 nuclei were isolated, pooled, homogenized and spun for 3 min at 8000 *g*. The cleared supernatant was irradiated on ice for 45 min, treated with RNase A and T1 and fractionated on a 12% SDS–PAGE. (**B**) Crosslinking to RNA incubated in nuclear extract. m^7G-cap labeled NL-15/5′Ext RNA was incubated in nuclear extract of four oocyte-equivalents. UV-crosslinking, RNase treatment and SDS–PAGE as in (A). Molecular size markers are indicated on the left.

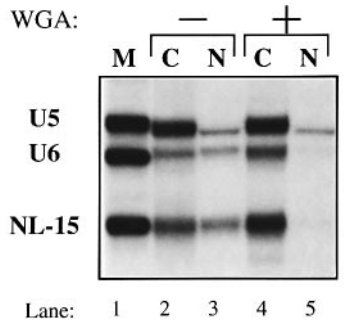

Fig. 9. Import of NL-15 RNA via a protein pathway. Inhibition of import of NL-15 RNA by WGA. The RNA mixture shown in lane 1 (M) was injected into cytoplasms of oocytes that had (+) or had not (–) been pre-injected with WGA. Nucleo–cytoplasmic distribution of the RNAs were analyzed 20 h after RNA injection as in Figure 2B.

However, we cannot exclude the possibility that additional proteins also bind NL-15 RNA.

Crosslinking of NL-15 RNA to nuclear proteins

The binding of NL-15 RNA to La protein was confirmed by transfer of label from the 5′-cap of NL-15 RNA to proteins in nuclei of oocytes upon UV-crosslinking. Several polypeptides were labeled (Figure 8A, lane 1), the most highly labeled of which (indicated by *) had the mobility of La protein (49 kDa). Labeling of this protein was greatly reduced by the co-injection of uncapped NL-15 competitor RNA (lane 2). Surprisingly, the presence of the uncapped competitor RNA also resulted in an increased labeling of the fastest migrating protein (lane 2). Three proteins (including a third protein with an apparent molecular weight of ~22 kDa) were labeled very efficiently when NL-15 RNA carrying the 5′ extension (NL-15/5′Ext, see Figure 5B) was injected (lane 3). These results indicate that the 49 kDa protein is probably La protein, which binds to the NL-15 RNA regardless of its cap status. However, it binds close enough to the cap to be cross-linked to it and to reduce the interaction of the cap with the other proteins. The 5′ extension of NL-15/5′Ext RNA allows the smaller proteins to interact with the cap, even in the presence of the 49 kDa protein. Because of its size and the fact that uncapped competitor RNA did not reduce its labeling (Figure 8B), the smallest protein is very likely to be the small subunit of CBC.

The identities of the proteins were confirmed *in vitro* by immunodepletion of GV extracts prior to RNA addition and UV-crosslinking (Figure 8B). When incubated in untreated extracts, NL-15/5′Ext RNA labeled all three protein (lane 1), but depletion of the extract by anti-CBP 20 antibodies resulted in loss of labeling of the smallest protein (lane 2). Similarly, immunodepletion of La protein from the extracts (lane 3) greatly reduced the labeling of the 49 kDa protein, as did the addition of uncapped competitor NL-15 RNA (lane 4). We conclude that the smallest and largest proteins are CBP 20 and La protein, respectively. The identity of the ~22 kDa protein has not been determined.

The precise location in NL-15 RNA to which La binds has not been determined. However, the 3′ half of NL-15 can be deleted without affecting binding of La to the RNA (data not shown), suggesting an internal sequence and/or structure as the La binding site; not a 3′ uridylate stretch like that to which La binds in U6 RNA. The striking increase in label transfer to CBP 20 upon removal of La protein (Figure 8A, lane 2) indicates that bound La protein inhibits access of CBC to the 5′-cap of NL-15 RNA, either sterically or through stabilization of an RNA structure that masks the cap. Thus, the appearance of NL-15 RNA in the cytoplasm after injection of high levels of the RNA into the nucleus (Figure 6) may result in part from unmasking of the 5′ cap rather than from the saturation of a nuclear retention site.

Proteins involved in nuclear import of selected NL-RNAs

In the final round of the selection procedures RNAs were injected into the cytoplasm but re-isolated from the nucleus, thereby imposing a requirement that the selected RNAs have the capacity to be imported into the nucleus. It is likely that the selected Sm^+ RNAs are imported by the same mechanisms as those used normally for Sm^+ spliceosomal RNPs since these two classes of RNAs undergo similar maturation events in the cytoplasm, such as binding Sm proteins and cap hypermethylation (Figure 2B). In support of this proposal, import of the selected Sm^+ RNA was hardly affected by the lectin WGA (data not shown), an effective inhibitor of import of NLS-containing proteins but not of spliceosomal snRNPs (Fischer *et al.*, 1991).

In contrast, uptake of NL-15 RNA, like that of U6 RNA, was strongly inhibited by injection of WGA under conditions where import of U5 RNA was unaffected (Figure 9). This sensitivity to the lectin indicates that NL-15 RNA, like U6 RNA, probably is imported through

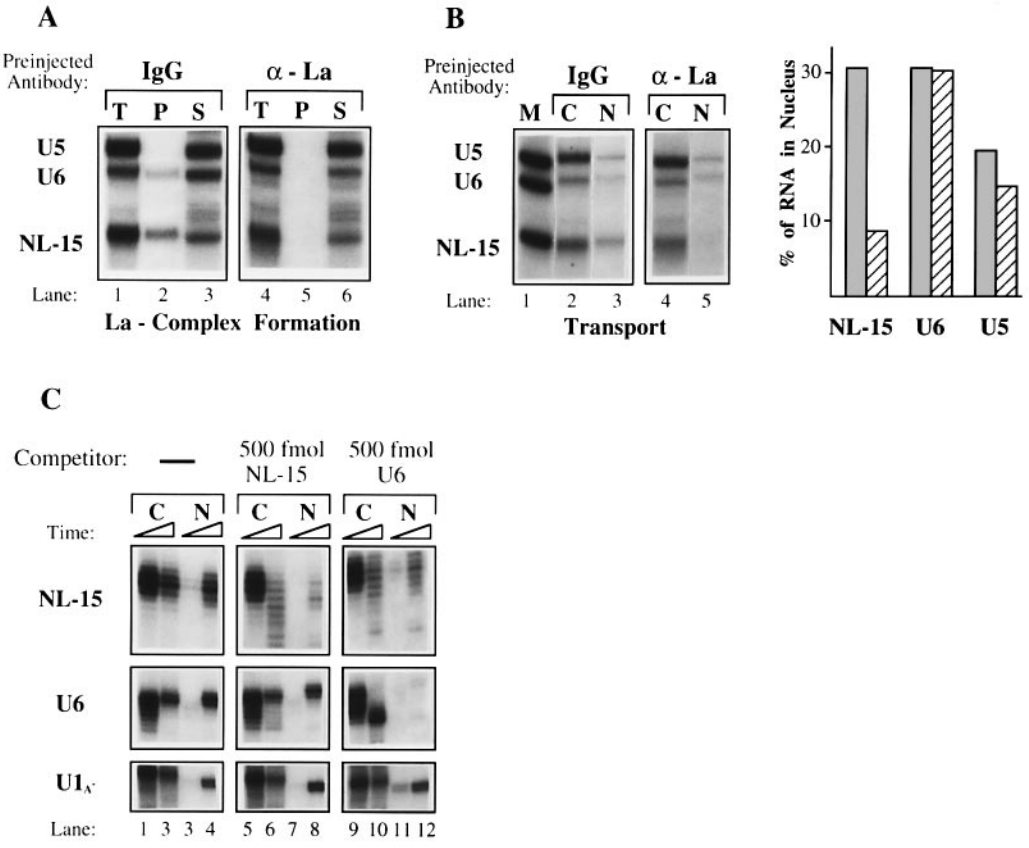

Fig. 10. Protein requirements for nuclear import of NL-15 RNA. (**A**) Blockage of formation of RNA–La complexes by anti-La antibodies. ^{32}P-labeled U5 (m^7G-capped), U6 (γ-mpppG-capped) and NL-15 (m^7G-capped) RNAs were co-injected into cytoplasms of oocytes that had been pre-injected with IgGs from normal human serum (lanes 1–3) or GO anti-La antibodies (α-La; lanes 4–6). Complex formation between La protein and RNAs was tested 9 h after RNA injection by immunoprecipitations of cytoplasmic extracts with B-103 anti-La antibodies. RNAs were prepared from total extract (T), precipitate (P) and supernatant (S) and analyzed by PAGE. (**B**) Blockage of nuclear import of NL-15 RNA by anti-La antibodies. RNAs were injected into oocytes that had been pre-injected either with IgGs (lanes 2 and 3) or GO anti-La antibodies (α-La; lanes 4 and 5) as in (A). Nucleo–cytoplasmic distribution was assayed 9 h after injection as in Figure 2B. M, RNAs prior to injection; C, cytoplasmic RNAs; N, nuclear RNAs. Quantitation of the RNAs in lanes 2–5 was done by PhosphorImager analyses and import in the presence of IgGs (empty bars) or anti-La antibodies (hatched bars) was expressed as [N/(N+C)]×100%. (**C**) Different requirements for nuclear import of NL-15 and U6 RNA. 1 to 2 fmol each of ^{32}P-labeled NL-15 (m^7G-capped), U6 (γ-mpppG-capped) and $U1_{A-}$ (m^7G-capped) RNAs were co-injected into cytoplasms of *Xenopus* oocytes in the absence (lanes 1–4) or presence of 500 fmol unlabeled NL-15 (lanes 5–8) or U6 (lanes 9–12) competitor RNAs. Nucleo–cytoplasmic distributions were assayed 3 h (lanes 1, 3, 5, 7, 9 and 11) and 24 h (lanes 2, 4, 6, 8, 10 and 12) after injection as in Figure 2B. C, cytoplasmic RNAs; N, nuclear RNAs.

its binding to an NLS-containing protein in an energy requiring process. This is supported by our finding that import of NL-15 RNA does not occur when the oocytes are incubated at 4°C (data not shown).

Since La protein can bind to NL-15 RNA in the nucleus, we tested whether the La protein present in the cytoplasm (Peek *et al.*, 1993) might bind NL-15 RNA and influence its nuclear import. After cytoplasmic injection NL-15, and to a lesser extent U6 RNA, could be coprecipitated with anti-La antibodies from the cytoplasm (Figure 10A, lanes 2 and 3). Preinjection of anti-La (but not control IgG) antibodies inhibited the formation of complexes between NL-15 or U6 RNAs and La protein (lanes 4 and 5). The anti-La antibodies reduced the fraction of NL-15 RNA that was imported into the nucleus by ~4-fold (Figure 10B, lanes 3 and 5 and right hand panel), but had no effect on import of U6 and U5 snRNAs. This suggests that import of NL-15 RNA is dependent specifically on complex formation with cytoplasmic La protein, whereas import of U6 (and U5) is not. The failure of high levels of NL-15 or U6 RNA to compete for nuclear import of the other RNA (Figure 10C) also demonstrates that migration of the two RNAs into the nucleus is mediated by different factors.

If La is involved in import of NL-15 RNA, one might expect that high levels of La binding RNAs such as U6 and hY1 would inhibit NL-15 import. However, the 3′ ends of both U6 and hY1 RNAs are removed when they are injected at high levels into the cytoplasm of oocytes (Figure 10C, lane 10 and data not shown), thereby losing their La binding sites as suggested also by Simons *et al.* (1996). Consistent with this, we found that injection of as much as 2 pmol of hY1 RNA in the cytoplasm failed to fully prevent complex formation between La protein and NL-15 RNA (data not shown). Therefore, although both U6 and hY1 RNAs destabilize NL-15 RNA to some extent (Figure 10C, lane 10 and data not shown), they have only a minor effect on NL-15 RNA import.

We propose that NL-15 RNA is imported as a consequence of its ability to bind to cytoplasmic La protein. In that sense, the RNA might use La as an NLS presenting carrier protein (or as a promoter of interaction with a

carrier protein) for its nuclear import. In contrast, import of U6 RNA which is not dependent on interaction with La, probably requires another, yet unidentified protein.

Discussion

The work presented here shows that *Xenopus* oocytes can be used to select RNAs that localize to specific sub-cellular compartments. The selected NLEs had to overcome the inherent rapid export characteristics of the m^7G-capped RNA carrier derived from pre-U1RNA, either by adding signals that would direct import back from the cytoplasm, by inactivating the export signals or by adding nuclear retention signals. Two motifs were isolated, one of which was known from previous work, and the other of which is novel. Undoubtedly, other motifs could be isolated in similar selections, since known NLEs, such as the box D sequences of nucleolar snoRNAs, were not found in this screen.

In vivo selection of functional Sm protein binding sequences

Since Sm protein binding promotes nuclear import of RNAs, we expected that one class of the RNAs selected for nuclear localization would contain Sm protein binding sites. Such a class was enriched by precipitation with anti-Sm antibodies. All of the selected molecules contained the motif AAUUUUUGG, located near the 3′ stem of the carrier RNA (Figure 2A). This consensus sequence strongly resembles other Sm protein binding sites both in sequence and in location, near an RNA stem–loop structure (Jarmolowski and Mattaj, 1993). The ability of the selected Sm sites in the NL-RNAs to function both in transport and cap hypermethylation (Figure 2B) validated the use of *in vivo* selection in the isolation of authentic RNA localization elements. Furthermore the homogeneity of the isolated $A_2U_5G_2$ consensus makes it unlikely that completely unrelated sequences or structures can function efficiently in Sm protein binding. However, the possibility exists that other Sm^+ binding sites were present in the molecules precipitated after round 4, but that these other sites were unable to function effectively enough in either import or stabilization of the RNA to survive all rounds of selection. This may explain why most of the RNAs that were coprecipitated with Sm proteins after four rounds of selection were only poorly imported in round 5 (Figure 1B). We are currently sequencing some of these early Sm^+ NL-RNAs to test the existence of unrelated Sm protein binding sites.

Although the RNA carrier for the randomized sequences was derived from U1 RNA, the Sm sites of the selected NL-RNAs differ from those of most U1 RNAs, in which the U stretch is interrupted by single nucleotide changes (AAUUUGUGG in human, rat, mouse, chicken and bean U1 RNAs; AAUUUCUGG in frog U1 RNAs; Reddy and Busch, 1988, and references therein). We note that the Sm^+ NL-RNAs were not selected for their ability to function in RNA processing. As was shown previously, close agreement with the consensus Sm binding site, while increasing the efficiency of nuclear localization and stability, may interfere with the function of certain Sm^+ RNAs (Grimm *et al.*, 1993). It is unclear whether the consensus found here would affect the ability of U1 RNA to function in pre-mRNA splicing.

Structures of Sm^- RNAs selected for nuclear localization

In contrast to the Sm^+ NL-RNAs, no consensus nucleotide sequence motif could be found in the selected Sm^- NL-RNAs. However, the 5′ regions of all of these RNAs could be folded into structures that contained extended stems, with several bulged nucleotides and/or small loops not conserved in sequence or position (Figure 3 and data not shown). Disruption of the proposed structure by substitution of blocks of nucleotides in NL-25 RNA reduced its nuclear localization and stability, and compensatory substitutions designed to re-establish a double-stranded structure restored these features significantly, but not completely (Figure 4). Because restoration of the nuclear localization and stability of the doubly mutant RNA was incomplete, we suggest that the unpaired nucleotides in the stem region of NL-25 RNA are important for the interaction of this RNA with cellular proteins and for nuclear localization. We note that the selected sequences are particularly low in adenosine residues (Figure 3A) and we speculate that this bias in sequence was caused by conversion of adenosines to inosines in double-stranded stems (Polson and Bass, 1994) during *in vivo* selection.

The selected NLEs of Sm^- NL-RNAs may function in nuclear localization by masking features in the carrier RNA that would otherwise promote export. In particular, the 5′ m^7G-caps of NL-RNAs are located adjacent to the RNA duplexes which may prevent their recognition as an export signal. Although the m^7G-cap structure is not essential for export of pre-snRNAs, it increases the export efficiency of the RNAs (Hamm and Mattaj, 1990; Terns *et al.*, 1993) by providing the binding site for the proteins of CBC (Izaurralde *et al.*, 1995). The importance of an accessible 5′-cap for export of NL-15 RNA is demonstrated by the appearance in the cytoplasm of a variant of NL-15 RNA in which the 5′-cap was at the end of a single-stranded extended region.

Nuclear retention of Sm^- NL-RNAs

The appearance of NL-15 RNA in the cytoplasm also could be promoted by nuclear injection of large amounts of competitor NL-15 RNA (Figure 6), presumably as a consequence of saturation of a limited number of specific retention sites in the nucleus. Several other nuclear RNAs, such as U6 spliceosomal RNA and the snoRNAs U3 and U8, appear to have specific nuclear retention signals (Hamm and Mattaj, 1989; Terns *et al.*, 1993, 1995; Terns and Dahlberg, 1994; Boelens *et al.*, 1995; our unpublished results); the differential saturation of nuclear retention for U3, U8 and U6 RNAs indicates the available number of special retention factors(s) each RNA is limited for (Terns *et al.*, 1995; our unpublished results). However, such experiments cannot distinguish between retention of an RNA as a result of binding to a factor that anchors it to a nuclear structure versus binding to a molecule that masks an export signal.

One nuclear factor that may contribute to the nuclear localization of NL-15 RNA, in part by masking the m^7G-cap export signal, is La protein. NL-15 RNA associates

with this abundant, predominantly nuclear protein, as assayed by gel shift experiments, UV-crosslinking and immunoprecipitations with anti-La antibodies. Moreover, nuclear injection of high levels of NL-15 RNA accelerates the export of wild-type hY1 RNA, an RNA that binds La protein in the nucleus and normally is exported very slowly (Simons *et al.*, 1994; Figure 7D). A mutation in hY1 RNA which removes its La binding site causes the rapid export of the RNA, indicating that La binding is responsible for nuclear retention of hY1 RNA (Simons *et al.*, 1996). Accordingly, we propose that high levels of NL-15 RNA effectively reduce the nuclear pool of free La protein so that most of the hY1 RNA is not bound by La, and thus can be exported rapidly. We conclude that NL-15 RNA binds La protein *in vivo* and that this interaction is involved in the retention of NL-15 in nuclei of *Xenopus* oocytes.

La protein can bind either to uridylate-rich 3′ ends (Stefano, 1984) or to internal sequences (Chang *et al.*, 1994; D.Kenan, personal communication) of RNAs. Because NL-15 RNA contains no 3′ uridylates, and deletion of the 3′ stem–loop of NL-15 RNA does not prevent its binding to La (data not shown), the site in this RNA that is recognized by La protein is probably within the 5′ duplex structure. Proximity of the 5′-cap and the La protein binding site also is indicated by the UV-mediated transfer of label from the cap to bound La protein. By binding close to the cap of NL-15 RNA, La protein apparently interferes with the recognition of the 5′-cap of the RNA by CBC, a nuclear export factor for snRNAs (Figure 8A). This interference could be direct through competition for binding to a site in the RNA or indirect through stabilization of a structure that masks the 5′-cap. The appearance of NL-15 RNA in the cytoplasm in the presence of high levels of NL-15 competitor RNA (Figure 6) thus may reflect saturation of La protein and possibly other nuclear factors, leaving the 5′-cap available for interaction with CBC.

A novel role for La protein as a mediator of RNA import from the cytoplasm to the nucleus

To survive the final round of selection, the NL-RNAs had to have the ability to be imported into the nucleus. Unlike the import of Sm^+ NL-RNAs (which occurs via the snRNP pathway), import of NL-15 RNA was strongly inhibited by the lectin WGA (Figure 9), indicating that the RNA is brought into the nucleus complexed with an NLS-containing protein, as is U6 RNA (Fischer *et al.*, 1991). Furthermore, both the formation of complexes between NL-15 RNA and La in the cytoplasm, and the import of NL-15 RNA into the nucleus were inhibited by cytoplasmic injection of antibodies that recognized La protein. These results suggest that the complex responsible for import of NL-15 RNA either contains La or requires La for its formation.

The failure of other La binding RNAs such as U6 (Figure 10C) and hY1 RNA (data not shown) to compete for nuclear import of NL-15 RNA most likely is due to their inability to interact efficiently with La protein in the cytoplasm. Both U6 and hY1 RNA normally bind La protein through their 3′ uridylate stretch, but both RNAs are trimmed in the cytoplasm when injected at high levels. This results in the loss of their La binding site and hence leads to their inability to compete for La binding. Previously, Simons and co-workers (1994) reported that La protein dissociates from hY1 RNA during or after nuclear export and suggested that this is caused by a 3′ end processing event which removes the La binding site (Simons *et al.*, 1996). Our findings are in agreement with this proposal.

Since nuclear uptake of U6 RNA was unaffected by antibodies to La protein or high levels of NL-15 competitor RNA (Figure 10), we conclude that La protein is not needed for U6 import. Similarly, high levels of poly(ACG), a competent inhibitor for La binding (D.Kenan, personal communication), reduced the import of NL-15 RNA but was without effect on U6 and U1 import (data not shown). Thus, NL-15 and U6 RNAs define two similar but separate RNA import pathways, both of which differ from the pathway used to import Sm^+ snRNPs.

In vivo selection of RNAs from combinatorial libraries

The isolation of functional Sm sites with a strong consensus motif, and the identification of a novel RNA structural element in the selection for NLEs demonstrates the feasibility of using an *in vivo* selection method to isolate RNAs with desired intracellular localization characteristics. The method identifies RNAs within a combinatorial library of molecules that are both stable in the cell and have the selected localization property.

Xenopus laevis oocytes are ideal cells for this type of selection since they can readily be microinjected and fractionated. Moreover, these cells have the capacity to deal with large numbers of molecules, allowing for the use of reasonably large pools of RNAs in the first rounds of *in vivo* selection. We are modifying this selection method to study other mechanisms that contribute to RNA transport and intracellular localization.

The RNAs selected in this study show that several mechanisms can be used, alone or in concert, to localize RNAs in cell nuclei. Likewise, an NLE, such as the La protein binding site, may support nuclear localization in more than one way.

Materials and methods

DNA templates and in vitro transcription

DNA templates for *in vitro* transcription were generated by PCR amplification of RNA coding regions using appropriate primer pairs. Templates used to transcribe U1, U2, U3 and U6 RNAs were described previously (Terns *et al.*, 1993, 1995). The template for U5 RNA was generated by PCR amplification of the *X.laevis* X.l.U5 11H gene (Kazmaier *et al.*, 1987) using a 5′ primer containing the SP6 promoter sequence (5′-GGAATTCGATTTAGGTGACACTATAGAATACTCTG-GTTTCT-3′) and a 3′ primer with a two nucleotide extension to make precursor length U5 RNA (5′-AGTACCTGGTGTGAACCAGGC-3′). The template for hY1 RNA was described previously (Simons *et al.*, 1994). *In vitro* transcription of T7 or SP6 DNA templates was done in 20 µl reactions containing 40 mM Tris–HCl pH 7.9, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 0.1 mg/ml BSA, 10 mM DTT, 2–4 units RNasin, 0.3 mM rATP, rCTP and rUTP, 0.1 mM rGTP plus 20 µCi [α-^{32}P]rGTP (25 pmol) and either 0.5 mM m^7GpppG or ApppG-cap dinucleotide (NEB) or 2 mM γ-mpppG (kindly provided by Ram Reddy); incubation was for 1–2 h at 37°C with 20 units of either T7 or SP6 RNA polymerase. Unlabeled competitor RNAs were prepared in 100 µl reactions containing 80 mM HEPES–KOH [*N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid)] pH 7.5, 16 mM $MgCl_2$, 2 mM spermidine, 40 mM DTT, 2 mM rATP, rCTP, rUTP, 0.2 mM rGTP and 1 mM (m^7GpppG) or 2 mM (γ-mpppG) cap analog. Incubation was for 2 h

with 100 units of RNA polymerase, followed by a second addition of 100 units of RNA polymerase and further incubation for 2 h. All RNA transcripts were purified by electrophoresis in a 8% denaturing polyacrylamide gel and elution in 0.3 M NaCl, 10 mM Tris–HCl pH 7.6, 0.1 mM EDTA and 0.5% SDS.

Oocyte injection and analysis of RNA transport

Nuclei or cytoplasms of intact stage V and VI oocytes from *X.laevis* were injected with 12 nl of H_2O containing 1–10 fmol of ^{32}P-labeled RNAs and where indicated, different amounts of unlabeled competitor RNAs. The injection mixture also contained blue dextran to monitor the accuracy of nuclear injection (Jarmolowski *et al.*, 1994). After incubation at 18°C for different times (see figure legends), oocytes were manually dissected under mineral oil (Lund and Paine, 1990) into nuclear and cytoplasmic fractions. After proteinase K digestion, total RNAs were isolated from each fraction by two extractions with phenol–chloroform (24:1) and ethanol precipitations and purified RNAs were analyzed by electrophoresis in 8% polyacrylamide gels containing 7 M urea.

In vivo selection

The DNA template used to transcribe the pool of RNAs for the first round of selection was prepared by annealing 50 pmol of an 87 nucleotide oligonucleotide (complementary to the RNA sequence shown in Figure 1A) to 250 pmol of a partially overlapping oligonucleotide containing the T7 promoter sequence plus nucleotides 1–19 of the RNA shown in Figure 1A (T7 SELEX: 5′-AATGTCGACTAATACGACTCACTATA-GGGATACTTACCTGGCAGG-3′). After annealing at 60°C, the products were extended with Stoffel enzyme (Perkin Elmer) for 1 h at 60°C. Full-length double-stranded products were purified by electrophoresis in a 6% polyacrylamide gel. For generation of the starting pool of RNAs, 250 ng of the gel-purified template was transcribed with T7 RNA polymerase and RNAs were purified as described above. For the first round of selection, 50 fmol of the experimental RNAs were injected together with 1–2 fmol each of the control RNAs into nuclei or cytoplasms of 50 oocytes. Theoretically, this corresponded to 2.5×10^{12} molecules and thus could contain all of the 1.1×10^{12} different molecules that can be formed from a 20 nucleotide long randomized sequence. After 20–24 h of incubation at 18°C, oocytes were dissected into nuclei and cytoplasms and total RNA prepared from both compartments. Analytical polyacrylamide gels were used to determine the nucleo–cytoplasmic distribution of experimental and control RNAs at each round of selection. Prior to reverse transcription and PCR amplification (RT–PCR), the experimental RNA in the nuclear fraction was size selected and purified by electrophoresis in a 8% polyacrylamide gel containing 7 M urea. Reverse transcription was done in a 20 µl reaction containing 50 mM Tris–HCl pH 8.5, 75 mM KCl, 10 mM DTT, 3 mM $MgCl_2$, 0.5 mM dNTPs, 2–4 units RNasin and 5 µM primer CS536 (5′-ATCAGGGGAAAGCGCGAACGCAGTCC-3′). The mixture was heated for 2 min at 95°C and cooled to 37°C prior to the addition of 1 µl of M-MLV reverse transcriptase (200 units/µl, USB). After incubation for 15 min at 37°C, 80 µl of a mixture containing 1.8 mM $MgCl_2$, 50 mM KCl, 10 mM Tris–HCl pH 9.0, 0.1% Triton X-100, 8 µM primer T7 SELEX and 0.5 units of *Taq* DNA polymerase (Promega) were added and overlayed with mineral oil to prevent condensation. PCR amplification was done using 35 cycles of denaturation (1 min at 95°C), annealing (45 s at 68°C) and extension (1 min at 72°C). RT–PCR products were fractionated by electrophoresis in a 6% polyacrylamide gel and the purified DNA templates were used to transcribe RNA for the next round of selection. The total amounts of RNA used for injection were: 50 (rounds 1–4) or 2–10 (rounds 5–12) fmol per oocyte and the number of oocytes injected were 50 (rounds 1–2), 30 (round 3) or 5–10 (rounds 4–12). The RT–PCR products after the 12th round of selection were re-amplified using a 5′ primer containing a *Hin*dIII restriction site and a 3′ primer containing a *Eco*RI restriction site. These two sites were used for cloning of the PCR products into pGEM-4Z vector (Promega). *Escherichia coli* cells were transfected by electroporation and plasmids were isolated from individual colonies. Inserts were sequenced by the dideoxy termination method using Sequenase version 2.0 (USB).

Structure probing

The substrate for structure probing in solution was *in vitro* prepared NL-15 RNA with a single label in the m^{7}G-cap structure. Unlabeled NL-15 RNA was m^{7}G-capped using guanylyltransferase plus [α-^{32}P] GTP and *S*-adenosylmethionine (SAM) as described (Terns *et al.*, 1995). Cleavage with RNase One (Promega) was done with 0.03 units of enzyme at 22°C in a 100 µl reaction containing 10 mM Tris–HCl pH 7.5, 5 mM EDTA, 200 mM Na-Acetate, 10 µg tRNA and 50 fmol of NL-15 RNA. Cleavage with RNase V1 (Pharmacia) was done with 0.7 units of enzyme at 22°C in a 100 µl reaction containing 10 mM Tris (pH 7.5), 10 mM $MgCl_2$, 50 mM KCl, 10 µg tRNA and 50 fmol of NL-15 RNA. After 2, 6 and 18 min, 25 µl aliquots were removed and the reactions were stopped by the addition of SDS (to a final concentration of 0.1%) and 10 µg yeast RNA. RNAs were prepared immediately by phenol–chloroform (24:1) extraction and ethanol precipitation. Cleavage with RNase A (ICN) and RNase T1 (Calbiochem) was done with 10^{-4} units or 1 unit of enzyme, respectively for 18 min at 55°C in 7 M urea, 1 mM EDTA, 25 mM Na-Acetate pH 7.0 and 10 µg tRNA. For controls, NL-15 RNA was incubated in buffer without enzyme under the respective conditions for 18 min. RNase cleavage products were separated on a 10% polyacrylamide gel containing 8.3 M urea.

Site directed mutagenesis of individual RNAs

NL-25 and NL-15 RNAs were mutagenized by PCR amplification using the following sets of primers. For NL-25/mut1: 5′ primer SP6-U1 5′mut (5′-GAATTCGATTAGGTGACACTATAGAATACTATGGTG-GCAGGGG-3′) and 3′ primer CS536; for NL-25/mut2: 5′ primer T7 SELEX and 3′ primer NL-25/mut2 (5′-ATCAGGGGAAAGCGCG-AACGCAGTCCACTACCAGAATACTATGGAAAGTCCTCAGGG-3′); for NL-25/mut1+2: 5′ primer SP6-U1 5′mut and 3′ primer NL-25/mut2; for NL-15/5′Ext: 5′ primer SP6-NL-RS (5′-GGAATTCGATTTA-GGTGACACTATAGAACTAGAGTACTGGGATACTTACCTGGCA-GGGG-3′) and 3′ primer CS536. PCR products were purified by electrophoresis in a 6% polyacrylamide gel and used for *in vitro* transcription.

Antibodies, immunoprecipitations and immunodepletions

Rabbit polyclonal antibodies against the m^{2,2,7}G- (Bringmann *et al.*, 1983; kindly provided by R.Lührmann) and the m^{7}G-cap (Munns *et al.*, 1982; kindly provided by T.Munns) were used to precipitate deproteinized RNAs, mouse monoclonal antibodies against Sm proteins (mAb Y12, Lerner *et al.*, 1981; kindly provided by J. Steitz) and anti-La antibodies from human patient sera (B-103, GO; kindly provided by D.Kenan and J.Keene) were used to precipitate RNPs from nuclear and cytoplasmic extracts and deplete nuclear extracts of La protein. Anti-CBP20 antibodies (rabbit; Izaurralde *et al.*, 1995; kindly provided by E.Izaurralde and I.Mattaj) were used to immunodeplete nuclear extract of CBC.

Immunoprecipitations were done as described previously (Terns *et al.*, 1992). For the injection of anti-La antibodies, total IgGs were purified from serum GO, essentially as described (Harlow and Lane, 1988). IgGs from 1.2 ml serum were bound to protein A–Sepharose beads in a 1.5 ml column in 100 mM Tris–HCl (pH 8). The column was washed with 10 column volumes of 100 mM Tris–HCl (pH 8) followed by 10 column volumes of 10 mM Tris (pH 8) and IgGs were eluted with 100 mM glycine (pH 3). After the addition of 0.1 volumes of 1 M Tris–HCl (pH 8), the neutralized IgGs were precipitated by the addition of 1 volume of saturated $(NH_4)_2SO_4$ and collected by 30 min of centrifugation at 3000 *g*. The pellet was drained, resuspended in 0.1 ml of PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) and dialyzed against 3× 2000 ml of PBS. IgGs were further concentrated ~8-fold using a microconcentrator (microcon 100; Amicon). Anti-La activity of the concentrated IgGs was tested in separate gel shift experiments (data not shown) and 60 nl (per oocyte) of the solution containing 3 mM DTT and 4 units RNasin/µl were injected into the cytoplasm of each oocyte.

Anti-CBP20 or anti-La (GO) antibodies were coupled to protein A–Sepharose beads and used to immunodeplete nuclear extract from La protein or CBC, respectively. Extract from 50 nuclei was incubated with the respective antibodies for 1.5 h on ice with occasional stirring. The mixes were spun for 10 s and the supernatants used as immunodepleted extracts.

Complex formation and native gel electrophoresis

Nuclear extracts from oocytes were prepared as described (Terns *et al.*, 1995). For complex formation, 10 fmol (1 µl) of m^{7}G-capped NL-15 or γ-mpppG-capped U6 RNA were mixed with 50 ng of 23S rRNA (1 µl) and 8 µl (0.5 oocyte equivalents) of nuclear extract in D_{250} buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 3 mM DTT, 50 mM Tris–HCl, pH 7.6). After incubation for 20 min at 19°C, 2.5 µl of loading solution (50% glycerol, 2.5 mM EDTA, 0.01% bromphenol-blue, 0.01% xylene cyanol) was added and the samples were fractionated immediately in native 6% polyacrylamide gels (30:0.8) in 0.5× TEB (1× TEB: 90 mM Tris, 90 mM boric acid, 2.5 mM EDTA). For supershifts, the samples with preformed complexes (see above) were incubated for 20 min on ice with 1 µl anti-La antibodies B-103 (diluted

in D_{250} and mixed with RNasin) prior to addition of the loading solution. The gels were run at room temperature for ~2 h at 10 V/cm, fixed in a solution containing 10% acetic acid and 20% methanol for 25 min and dried prior to autoradiography.

Escherichia coli extracts containing human La protein

Escherichia coli cells [strain BL21(DE3) pLysS] expressing recombinant human La protein were kindly provided by D.Kenan. Cells were grown in LB in a 50 ml culture; expression of La protein was induced by the addition of IPTG (isopropyl β-D-thiogalactopyranoside) to a final concentration of 0.4 mM. 3 h after induction, PMSF (phenylmethylsulfonyl fluoride) was added to a final concentration of 0.125 mg/ml and cells were harvested by centrifugation at 5000 *g* for 10 min at room temperature. Cells were resuspended in 2 ml of 25 mM Tris–HCl pH 8.0, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM DTT, 100 mM NaCl, 0.125 mg/ml PMSF and quickly frozen in a dry-ice ethanol bath. Cells were thawed in a 37°C water bath in the presence of freshly added PMSF and sonicated for 2× 30 s (Branson sonifier, setting 2). After centrifugation at 12 000 *g* for 30 min at 4°C, the cleared supernatant was collected and fresh PMSF added. The extract was stored at 4°C and used as source for human La protein. Extracts from uninduced cells were prepared as above, except that no IPTG was added. 1 µl extract (diluted 1:16 in D_{250}) was used for complex formation as described above.

UV-crosslinking

10–30 fmol of RNA, labeled only in its 5′ m^7G-cap (Terns *et al*., 1995), were incubated in nuclear extracts as for complex formation (see above) in the absence or presence of unlabeled competitor RNAs as indicated in the figure legends. The extracts were either untreated or depleted from La protein or CBC. After 20 min of incubation, the samples were irradiated with UV light (short wavelength) for 45 min on ice. The samples were then treated with RNase T1 and RNase A for 1 h at 37°C. Proteins were concentrated by the addition of 5 volumes of acetone and centrifugation. The collected proteins were resuspended in SDS loading buffer and fractionated on 12% SDS–polyacrylamide gels. The gels were fixed in 10% acetic acid/20% methanol, dried and exposed for autoradiography. For the *in vivo* UV-crosslinking experiments, the RNA were injected into nuclei of intact oocytes. After 15–30 min of incubation, nuclei were isolated, pooled and homogenized in 50 µl of D_{250}. The homogenate was spun for 3 min at 8000 *g* in a microcentrifuge at 4°C and the cleared extracts were irradiated with UV-light and treated as described above.

Acknowledgements

We thank Michele Barr and Ariane Grandjean for technical assistance. We also thank Ram Reddy for furnishing the γ-mpppG-cap analog and Theodore Munns, Reinhard Lührmann, Joan Steitz and Dan Kenan for providing antibodies used in this study. We thank Dan Kenan and Jack Keene for sending us the *E.coli* strain expressing human La protein. This work was supported by NIH grant GM-30220 to J.E.D. and postdoctoral fellowship 823A-042935 from the Swiss National Science Foundation to C.G.

References

Allison,E.A., North,M.T., Murdoch,K.J., Romaniuk,P.J., Deschamps,S. and LeMaire,M. (1993) Structural requirements of 5S rRNA for nuclear transport, 7S ribonucleoprotein particle assembly, and 60S ribosomal subunit assembly in *Xenopus* oocytes. *Mol. Cell. Biol*., **13**, 6819–6831.

Boelens,W.C., Palacios,I. and Mattaj,I.W. (1995) Nuclear retention of RNA as a mechanism for localization. *RNA*, **1**, 273–284.

Bringmann,P., Rinke,J., Appel,B., Reuter,R. and Lührmann,R. (1983) Purification of snRNPs U1, U2, U4, U5 and U6 with 2,2,7-trimethylguanosine-specific antibody and definition of their constituent proteins reacting with anti-Sm and anti-(U1) RNP antisera. *EMBO J*., **2**, 1129–1135.

Chang,Y.N., Kenan,D.J., Keene,J.D., Gatignol,A. and Jeang,K.T. (1994) Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J. Virol*., **68**, 7008–7020.

Dabauvalle,M.C., Schulz,B., Scheer,U. and Peters,R. (1988) Inhibition of nuclear accumulation of karyophilic proteins in living cells by microinjection of the lectin wheat germ agglutinin. *Exp. Cell Res*., **174**, 291–296.

Davis,L.I. (1995) The nuclear pore complex. *Annu. Rev. Biochem*., **64**, 865–896.

Eckner,R., Ellmeier,W. and Birnstiel,M.L. (1991) Mature mRNA 3′ end formation stimulates RNA export from the nucleus. *EMBO J*., **10**, 3513–3522.

Finlay,D.R., Newmeyer,D.D., Price,T.M. and Forbes,D.J. (1987) Inhibition of *in vitro* nuclear transport by a lectin that binds to nuclear pores. *J. Cell Biol*., **104**, 189–200.

Fischer,U., Darzynkiewicz,E., Tahara,S.M., Dathan,N.A., Lührmann,R. and Mattaj,I.W. (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., Sumpter,V., Sekine,M., Satoh,T. and Lührmann,R. (1993) Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m_3G-cap. *EMBO J*., **12**, 573–583.

Fischer,U., Meyer,S., Teufel,M., Heckel,C., Lührmann,R. and Rautmann,G. (1994) Evidence that HIV-1 Rev directly promotes the nuclear export of unspliced RNA. *EMBO J*., **13**, 4105–4112.

Fischer,U., Huber,J., Boelens,W.C., Mattaj,I.W. and Lührmann,R. (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.

Forbes,D.J. (1992) Structure and function of the nuclear pore complex. *Annu. Rev. Cell Biol*., **8**, 495–527.

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

Grimm,C., Stefanovic,B. and Schümperli,D. (1993) The low abundance of U7 snRNA is partly determined by its Sm binding site. *EMBO J*., **12**, 1229–1238.

Grimm,C., Lund,E. and Dahlberg,J.E. (1995) *In vivo* selection of RNA sequences involved in nucleocytoplasmic RNA trafficking. *Nucleic Acids Symp. Ser.*, **33**, 34–36.

Hamm,J. and Mattaj,I.W. (1989) An abundant U6 snRNP found in germ cells and embryos of *Xenopus laevis*. *EMBO J*., **8**, 4179–4187.

Hamm,J. and Mattaj,I.W. (1990) Monomethylated cap structures facilitate RNA export from the nucleus. *Cell*, **63**, 109–118.

Harlow,E. and Lane,D. (1988) *Antibodies. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Izaurralde,E. and Mattaj,I.W. (1995) RNA export. *Cell*, **81**, 153–159.

Izaurralde,E. and Mattaj,I.W. (1992) Transport of RNA between nucleus and cytoplasm. *Semin. Cell Biol*., **3**, 279–288.

Izaurralde,E., Lewis,J., Gamberi,C., Jarmolowski,A., McGuigan,C. and Mattaj,I.W. (1995) A cap-binding protein complex mediating U snRNA export. *Nature*, **376**, 709–712.

Jarmolowski,A. and Mattaj,I.W. (1993) The determinants for Sm protein binding to *Xenopus* U1 and U5 snRNAs are complex and non-identical. *EMBO J*., **12**, 223–232.

Jarmolowski,A., Boelens,W.C., Izaurralde,E. and Mattaj,I.W. (1994) Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell Biol*., **118**, 1287–1295.

Kazmaier,M., Tebb,G. and Mattaj,I.W. (1987) Functional characterization of *X.laevis* U5 snRNA genes. *EMBO J*., **6**, 3071–3078.

Lerner,E.A., Lerner,M.R., Janeway,C.A.,Jr and Steitz,J.A. (1981) Monoclonal antibodies to nucleic acid-containing cellular constituents: Probes for molecular biology and autoimmune disease. *Proc. Natl Acad. Sci. USA*, **78**, 2737–2741.

Lund,E. and Paine,P.L. (1990) Nonaqueous isolation of transcriptionally active nuclei from *Xenopus* oocytes. *Methods Enzymol*., **181**, 36–43.

Mattaj,I.W. and DeRobertis,E.M. (1985) Nuclear Segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell*, **40**, 111–118.

Michaud,N. and Goldfarb,D.S. (1992) Microinjected U snRNAs are imported to oocyte nuclei via the nuclear pore complex by three distinguishable targeting pathways. *J. Cell Biol*., **116**, 851–861.

Munns,T.W., Liszewski,M.K., Tellam,J.T., Sims,H.F. and Rhoads,R.E. (1982) Antibody-nucleic acid complexes. Immunospecific retention of globin messenger ribonucleic acid with antibodies specific for 7-methylguanosine. *Biochemistry*, **21**, 2922–2928.

Nigg,E.A., Baeuerle,P.A. and Lührmann,R. (1991) Nuclear import-export: In search of signals and mechanisms. *Cell*, **66**, 15–22.

Peek,R., Pruijn,G.J., van der Kemp,A.J. and van Venrooij,W.J. (1993) Subcellular distribution of Ro ribonucleoprotein complexes and their constituents. *J. Cell Sci*., **106**, 929–935.

Pokrywka,N.J. and Goldfarb,D.S. (1995) Nuclear export pathways of tRNA and 40S ribosomes include both common and specific intermediates. *J. Biol. Chem*., **270**, 3619–3624.

Polson,A.G. and Bass,B.L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J.*, **13**, 5701–5711.

Powers,M.A., Forbes,D.J., Dahlberg,J.E. and Lund,E. (1997) The vertebrate GLFG nucleoporin, Nup98, is an essential component of multiple RNA export pathways. *J. Cell Biol.*, **136**, in press.

Reddy,R. and Busch,H. (1988) Small nuclear RNAs: RNA sequences, structure, and modifications. In Birnstiel,M.L. (ed.), *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*. Springer-Verlag, Berlin, Germany, pp. 1–37.

Simons,F.H.M., Pruijn,G.J.M. and Van Venroij,W.J. (1994) Analysis of the intracellular localization and assembly of Ro ribonucleoprotein particles by microinjection into *Xenopus laevis* oocytes. *J. Cell Biol.*, **125**, 981–988.

Simons,F.H.M., Rutjes,S.A., Van Venroij,W.J. and Pruijn,G.J.M. (1996) The interactions with Ro60 and La differentially affect nuclear export of hY1 RNA. *RNA*, **2**, 264–273.

Stefano,J. (1984) Purified lupus antigen La recognizes an oligouridylate stretch common to the 3′ termini of RNA polymerase III transcripts. *Cell*, **36**, 145–154.

Terns,M.P. and Dahlberg,J.E. (1994) Retention and 5′ cap trimethylation of U3 snRNA in the nucleus. *Science*, **264**, 959–961.

Terns,M.P., Lund,E. and Dahlberg,J.E. (1992) 3′ end-dependent formation of U6 small nuclear ribonucleoprotein particles in *Xenopus laevis* oocyte nuclei. *Mol. Cell. Biol.*, **12**, 3032–3040.

Terns,M.P., Dahlberg,J.E. and Lund,E. (1993) Multiple *cis*-acting signals for export of pre-U1 snRNA from the nucleus. *Genes Dev.*, **7**, 1898–1908.

Terns,M.P., Grimm,C., Lund,E. and Dahlberg,J.E. (1995) A common maturation pathway for small nucleolar RNAs. *EMBO J.*, **14**, 4860–4871.

Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.

Vankan,P., McGuigan,C. and Mattaj,I.W. (1990) Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in *Xenopus* oocytes. *EMBO J.*, **9**, 3397–3404.

Yang,H., Moss,M.L., Lund,E. and Dahlberg,J.E. (1992) Nuclear processing of the 3′ terminal nucleotides of pre-U1 RNA in *Xenopus laevis* oocytes. *Mol. Cell. Biol.*, **12**, 1553–1560.

Zapp,M.L. (1992) RNA nucleocytoplasmic transport. *Semin. Cell Biol.*, **3**, 289–297.

Zuker,M. (1989) On finding all suboptimal foldings of an RNA molecule. *Science*, **244**, 48–52.

Received on December 14, 1995; revised on September 30, 1996