Volume 16 Number 12 1988 Nucleic Acids Research

Plant small nuclear RNAs. V. U4 RNA is present in broad bean plants in the form of sequence variants and is base-paired with U6 RNA*

Tamás Kiss, Gábor Jakab, Mária Antal, Zsófia Pálfi, Hedvig Hegyi, Mihály Kis and Ferenc Solymosy+

Institute of Plant Physiology, Biological Research Center, Hungarian Academy of Sciences, Szeged, PO Box 521, 6701 Hungary

Received March 17, 1988; Accepted April 29, 1988 Accession nos X07112 and X07113

ABSTRACT

U4 RNA, which is known to play an indispensable role in pre-mRNA splicing, is present in plant nuclei, has a canonical $m_3^{2,2,7}G$ cap at its 5' end and is associated with U6 RNA in snRNP particles. It occurs in broad bean in the form of a number of sequence variants. Two of these were sequenced: U4A RNA is 154 and U4B RNA is 152 nucleotides long. Sequence similarity of broad bean U4B RNA is 94 per cent to broad bean U4A RNA, 65 per cent to rat U4A RNA, 61 per cent to Drosophila U4A RNA and 50 per cent to snR14, the U4 RNA equivalent of the yeast Saccharomyces cerevisiae. Sequence conservation is much more pronounced in the 5' half of the molecule than in its 3' half. The secondary structure of both variants of broad bean U4 RNA perfectly fits with that of all other U4 RNAs sequenced so far. Nucleotide changes between broad bean U4A and U4B RNAs are restricted to molecular regions that affect the thermodynamic stability of these molecules. A model is proposed for the base pairing interaction of broad bean U4 RNA with broad bean U6 RNA. This is the first report on the structure of a plant U4 RNA.

INTRODUCTION

All eukaryotic cells contain at least six major species of urydilate-rich small nuclear (sn) RNAs (U-RNAs): U1 to U6 RNAs in metazoa [reviewed in (1)] or their equivalents in the yeast Saccharomyces cerevisiae (2, 3, 4, 5, 6, 7). These molecules are capped at their 5' end by $m_3^{2,2,7}G$ [m_3G; cf. (1)], except U6 RNA which has a non-nucleotide cap [cf. (1)], and occur in the form of individual ribonucleoprotein complexes [U-snRNPs; reviewed in (8, 9)], except U4 and U6 RNAs which co-exist in a base-paired state in the same U-snRNP particle [U4/U6 snRNP; (10, 11, 12)]. U-snRNPs are known to play an indispensable role in pre-mRNA splicing within the spliceosome [reviewed in (13)].

Whereas the occurrence of U-RNAs in animals has been known for over twenty years (14), they were only first discovered in plants five years ago by Krol et al. (15). These authors used anti-cap antibodies (16) to specifically precipitate the m_3G-capped U-RNAs from pea nuclei, established

the entire sequence of five variants of pea U5 RNA and partially sequenced U1 and U2 RNAs from their 3' ends. They failed, however, to detect in pea nuclei U6 RNA for obvious reasons (lack of a canonical cap) and U3 RNA as well as U4 RNA for undefined reasons.

Two years later we developed two methods (17) for the isolation of nuclei from plant material. This enabled us to detect, and to establish the partial nucleotide sequence of, U3 (17) and the entire nucleotide sequences of U6 (18) and U2 (19) RNAs from broad bean.

Strangely enough, neither Krol et al. (15) nor we (17, 18) were able to ascertain the presence of U4 RNA in plant nuclei on the basis of expected gel electrophoretic mobility. On the other hand, Tollervey (20) has recently referred to a pea snRNA species that was precipitable with both anti-m_3G and anti-Sm antibodies and hybridized to cloned human U4. The fact that in all eukaryotes studied so far U4 and U6 RNAs, and even their equivalents in Saccharomyces cerevisiae (6), seem to co-exist in U4/U6 snRNP [cf. (8)] and that U6 RNA was shown to be present in plant nuclei (18), prompted us to re-investigate the occurrence of U4 RNA in plants by a combination of different approaches.

In this paper we report that U4 RNA can be detected in broad bean and tomato nuclear extracts by hybridization with a human U4 DNA clone, has a canonical cap structure (precipitation with anti-m_3G antibodies), is associated with U6 RNA in snRNP particles (precipitation with anti-Sm antibodies) and occurs, at least in broad bean nuclei, in the form of a number of sequence variants (mobilities in different gel electrophoretic systems). We established the entire primary and possible secondary structures of two of them and propose a model for base-pairing interaction of broad bean U4 RNA with broad bean U6 RNA.

MATERIALS AND METHODS

1. Reagents and Enzymes

All chemicals were of analytical grade. [Alpha-^{32}P] ATP, [gamma-^{32}P] ATP and [^{32}P] -orthophosphate were from Izinta (Budapest), Nonidet P-40 and Pansorbin from Sigma (St Louis), Protein A-Sepharose CL-4B and RNA ligase from Pharmacia P-L Biochemicals (Uppsala) and polynucleotide kinase from Boehringer (Mannheim).

2. Plant Material

Leaves of broad bean (Vicia faba L.) or tomato (Lycopersicon esculentum Mill cv. K-700) plants grown under ordinary greenhouse conditions and suspension

culture cells from Lycopersicon peruvianum Mill callus tissue kindly provided by Dr. L. Nover (Institut für Biochemie der Pflanzen, Halle) were used.

3. Cell Growth and Labeling Conditions

Lycopersicon peruvianum cells were grown in suspension culture at 25° C in a medium described by Nover et al. (21) and kept in the mid-log phase by passage every third day. A 100 ml culture (approximately 10^6 cells/ml) was labeled with [^{32}P]-orthophosphate (0.37 MBq/ml) for 4 h, after depletion of the phosphate-pool for 12 h in a Nover medium with 1/10 of the original phosphate content.

4. Isolation of Plant Nuclei

Nuclei were isolated from leaf material by Method I and from labeled suspension culture cells, after partial enzymatic digestion, by Method II of Kiss et al. (17).

5. Preparation of Plant Nuclear Extracts

Extracts from labeled nuclei of suspension culture cells of Lycopersicon peruvianum were obtained essentially according to the protocol described by Bringmann et al. (10) for the preparation of high salt nuclear extracts from HeLa cells.

6. Immune Affinity Chromatography of Plant Nuclear Extracts Using Anti-m_3G IgG

Anti-m_3G affinity column was prepared by covalently linking rabbit anti-m_3G IgGs (a kind gift of Dr. P. Bringmann and Professor R. Lührmann, Max-Planck-Institut für Molekulare Genetik, Berlin) to Protein A-Sepharose CL-4B according to Gersten and Marchalonis (22). The affinity column was washed with 5 ml of high salt buffer (40 mM Tris-HCl, pH 8.0, 420 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 0.5 mM dithioerythritol and 0.5 mM phenylmethylsulfonyl fluoride). Then 1 ml of [^{32}P]-labeled high salt nuclear extract was passed over the column five times, followed by washing with 2.5 ml of high salt buffer. The bound material was eluted from the column with 1 ml high salt buffer containing 12 mM m^7G (23).

7. Immunoprecipitation of Plant SnRNPs Using Anti-Sm Mouse Monoclonal Antibody, Y12

Precipitation of antigen-antibody complexes by Protein A-Sepharose was based on the method described by Matter et al. (24) and additional steps of immunoprecipitation were modifications of those by Lerner and Steitz (25) and Mimori et al. (26). Non-specifically binding material in [^{32}P]-labeled plant nuclear extracts was adsorbed with the same volume of a 10 % Pansorbin suspension in IPP buffer (500 mM NaCl, 10 mM Tris-HCl, pH 8.0, 0.05 % Nonidet

P-40) for 20 min at 4° C. In a microfuge tube, 150 µg of anti-Sm mouse monoclonal antibody, Y12 (kindly provided by Dr. K. Parker and Professor J.A. Steitz, Yale University Medical School, New Haven) was adsorbed on 500 µl of 10 % Protein-A Sepharose in PBS buffer (10 mM K-phosphate, pH 8.0, 150 mM NaCl). The binding of IgGs to Protein-A Sepharose was allowed to occur for 2 h at room temperature. Then the beads were washed four times with 500 µl IPP buffer and incubated with 500 µl of the ^{32}P-labeled high salt nuclear extract (previously pre-adsorbed to Pansorbin in IPP buffer) by end-over-end rotation for 2 h at 4° C. Finally the antigen-antibody complexes immobilized on Protein-A Sepharose were washed five times with IPP buffer, and the beads resuspended in 400 µl NET-2 buffer (150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.05 % Nonidet P-40) supplemented with 0.5 % (w/v) SDS and 5 µg E.coli tRNAs.

8. RNA Extraction

RNA from leaf nuclei, from ^{32}P-labeled high salt nuclear extracts as well as from both the flow-through fraction and the m^7G eluate of immune affinity columns, and from immunoprecipitates were phenol-extracted at 65° C as described by Steele et al. (27).

9. Isolation of Rat Nuclear RNA

Rat liver nuclear RNA was isolated according to Reddy et al. (28).

10. Fractionation of RNAs

RNAs were fractionated by polyacrylamide gel electrophoresis under non-denaturing, semi-denaturing or fully denaturing conditions modified after the procedures described by Forbes et al. (29) and Lund and Dahlberg (30), and specified in detail in the figure legends. Elution of RNA from the gels was done as reported earlier (17).

11. Immunoprecipitation of RNAs with Anti-m_3G IgG

This was done as described by Krol et al. (15).

12. Molecular Hybridization

Rat, broad bean and tomato nuclear RNAs fractionated on polyacrylamide gels were transferred to nitrocellulose paper (Schleicher and Schüll BA 85) according to Southern (31) as originally described for DNA transfer, with the following modifications: Urea was washed out by soaking the gel in 10 mM Na-phosphate buffer, pH 7.0, and RNAs were directly blotted onto a nitrocellulose filter using 20 x SSC rather than 10 x SSC. Recombinant p21 plasmid harboring the entire coding region of human U4 RNA (32) was kindly provided by Dr. U. Pettersson (University of Uppsala, Biomedical Center, Uppsala), recombinant pSP65 plasmid containing a 1 kb Msp 1 fragment harbor-

ing the Drosophila U4 RNA gene (33) by Professor R. Reddy (Baylor College of Medicine, Houston) and plasmid pmU6 - 52BE carrying the entire coding region of mouse U6 RNA (34), by Dr. Y. Ohshima (Univ. of Tsukuba, Inst. of Biol. Sci., Sakura-mura, Ibaraki). The entire plasmid was labeled by random priming (35) and hybridized to the RNA blot at 37° C for 16 h in a buffer containing 30 % (v/v) formamide, 5xSSC, 0.1 % SDS, 1 mM EDTA, 10 mM Tris-HCl, pH 7.5, and 1 x Denhardt's reagent [1 x Denhardt's reagent is 0.02 % (w/v) bovine serum albumin, 0.02 % (w/v) polyvinylpyrrolidone and 0.02 % (w/v) Ficoll].

13. Nucleotide Sequence Analysis

Synthesis of [5'-^{32}P] pCp and 3'end-labeling of RNAs were performed according to England et al. (36). Chemical sequencing of U4 RNAs was carried out

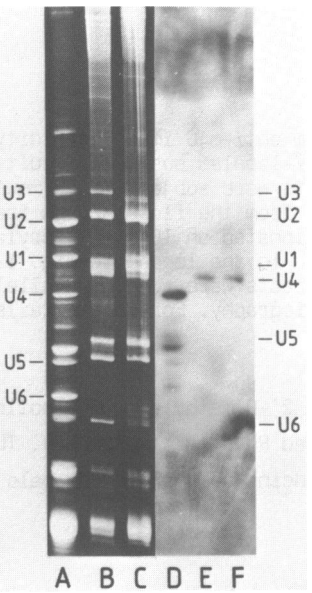

Fig. 1. Detection of plant U4 RNA by molecular hybridization with a human U4 DNA clone. Nuclear RNA extracts from rat (Lane A), broad bean (Lane B) and tomato (Lane C) cells were loaded on a 20 cm long and 1 mm thick 10 % polyacrylamide (19:1, acrylamide/bisacrylamide) gel, run under denaturing conditions at 60° C and 40 V/cm for 2.5 hours in 50 mM Tris-borate buffer (pH 8.3) containing 1 mM EDTA and 8M urea, and stained with ethidium bromide. The RNAs were then blotted onto a nitrocellulose filter and probed with a labeled human U4 DNA clone as described in detail in Materials and Methods. The autoradiograms are shown for rat (Lane D), broad bean (Lane E) and tomato (Lane F) nuclear RNAs. The RNA species in Lane A are identified on the left, and those in Lane B on the right side of the Figure.

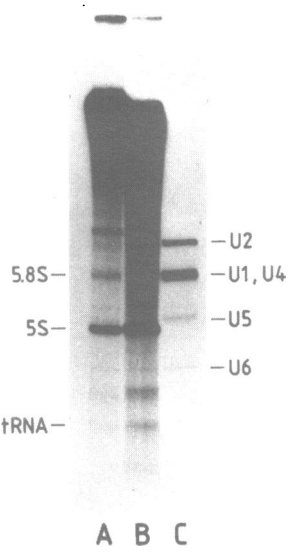

Fig. 2. U-RNAs eluted from anti-m_3^2G immune affinity columns treated with nuclear extracts of [^{32}P]-labeled suspension culture cells of Lycopersicon peruvianum. Nuclear extracts were subjected to immune affinity chromatography and the RNAs extracted from the flow-through fraction (Lane B) and the eluate (Lane C) were fractionated on 10 % polyacrylamide gels under denaturing conditions as described in the legend to Fig. 1. As a reference RNAs from untreated nuclear extracts were run in parallel (Lane A). The RNA bands were visualized by autoradiography. For more details see Materials and Methods.

according to Peattie (37). 2'-O-methylated nucleotides were mapped by limited hydrolysis of end-labeled RNAs in alkali (38). The digests were fractionated on 20 % and 8 % sequencing polyacrylamide gels (39).

RESULTS

1. Detection of U4 RNA in Plant Nuclei

When nuclear RNA from rat liver (Fig. 1A), broad bean (Fig. 1B) or tomato (Fig. 1C) was run on a 10 % denaturing polyacrylamide gel and hybridized with a human (or Drosophila, not shown) U4 DNA clone, positive signals were obtained with all three samples. U4 RNA from both broad bean (Fig. 1E) and tomato (Fig. 1F) nuclei migrated more slowly than that of rat liver nuclei (Fig. 1D) and was positioned in the lower part of the plant U1 RNA region.

2. Purification of Broad Bean U4 RNA

Since excessive contamination by broad bean U1 RNA of eventual U4 RNA ex-

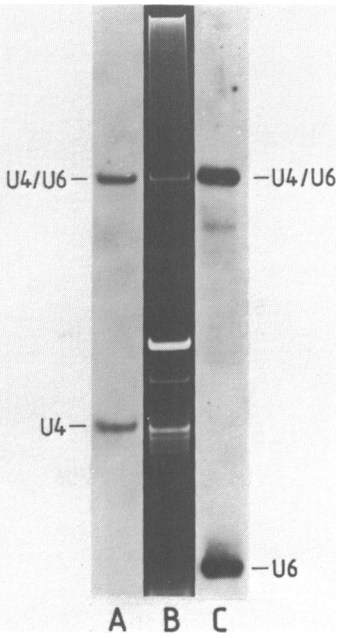

Fig. 3. Detection of base pairing interaction between U4 and U6 RNAs in RNA extracts from broad bean nuclei. A broad bean nuclear RNA extract was loaded on a 20 cm long and 1 mm thick 7 % polyacrylamide (29:1, acrylamide/bisacrylamide) gel, run under non-denaturing conditions at 4° C and 10 V/cm for 5 hours in 50 mM Tris-borate buffer (pH 8.3) containing 1 mM EDTA, and stained with ethidium bromide (Lane B). After blotting, the nitrocellulose filters were probed with a human U4 (Lane A) or a mouse U6 (Lane C) DNA clone, as described in Materials and Methods.

tracts from denaturing preparative gels, onto which total nuclear RNA had been loaded, would make any further analysis of broad bean U4 RNA impossible, we made an attempt to isolate broad bean U4 RNA from its U4/U6 RNA complex expected to occur in plants by analogy with other eukaryotes (6, 11, 12). The presence of U4/U6 RNA complexes in plant nuclei was already suggested earlier by our finding that from plant nuclear extracts U6 RNA was precipitated by anti-m_3G antibodies (Fig. 2) despite the lack of a canonical cap structure in plant U6 RNA (18). Indeed, when a nuclear RNA extract from broad bean was run on a 7 % polyacrylamide gel under non-denaturing conditions a slowly migrating band appeared (Fig. 3B) that contained both U4 RNA and U6 RNA. This was demonstrated by molecular hybridization with either a human U4 (Fig. 3A) or a mouse U6 DNA clone (Fig. 3C) as well as by re-

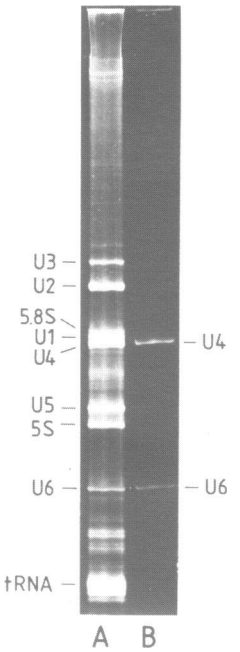

Fig. 4. Resolution of broad bean U4/U6 RNA complex into its components. The broad bean U4/U6 RNA complex eluted from a non-denaturing polyacrylamide gel (see Fig. 3. Lane B) was subjected to polyacrylamide gel electrophoresis under denaturing conditions as described in the legend to Fig. 1., and stained with ethidium bromide (Lane B). As a marker, total nuclear RNA from broad bean was electrophoresed under the same conditions (Lane A) along with the U4/U6 RNA complex.

-running the complex on a 10 % denaturing gel (Fig. 4). Broad bean U4 RNA isolated from similar denaturing gels proved to be capped by m_3G at its 5' end, indicated by its immunoprecipitability with authentic anti-cap antibodies (Fig. 5). This result supported our earlier finding presented in Fig. 2. Since labeled U4 RNA obtained in this way did not prove to be homogeneous enough for sequence analysis it was further separated into three components, each with a slightly different mobility on a 15 % non-denaturing gel (Fig. 6, upper horizontal Lane A). Subsequent fractionation of these components on a 12 % semi-denaturing polyacrylamide gel gave rise (Fig. 6, lower vertical Lanes B-D) to three apparently distinct RNA species (1, 2a and 3a), all with an A residue at the 3' end (not shown), one mixture (2b) of molecular species (both A and G residues at the 3' end, not shown) and a minor component (3b) which was not further analyzed.

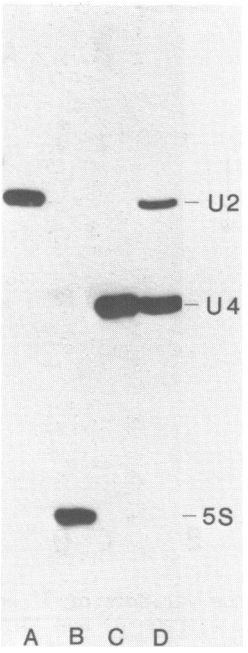

Fig. 5. Precipitation of broad bean U4 RNA with anti-m_3^2G antibodies. The 3' end-labeled U2, 5S and U4 RNAs from broad bean were combined and treated with anti-m_3^2G antibodies as described in Materials and Methods. The precipitable RNAs (Lane D) of the above mixture as well as 3' end-labeled U2 (Lane A), 5S (Lane B) and U4 (Lane C) RNAs as molecular markers were separated on a 10 % polyacrylamide gel according to (39). The RNA bands were visualized by autoradiography.

3. Primary Structure of U4A and U4B RNAs from Broad Bean

The RNAs extracted from bands 1 and 3a (Fig. 6 Lanes B and D) were subjected to chemical sequence analysis (Fig. 7). The sequences of U4A RNA extracted from band 1 and of U4B RNA extracted from band 3a are presented and compared with those of rat, Drosophila and the yeast, Saccharomyces cerevisiae, in Fig. 8. Broad bean U4B RNA is 152, U4A RNA is 154 nucleotides long, and both are longer than their metazoan equivalents. Sequence similarity of broad bean U4B RNA is 94 per cent to broad bean U4A RNA, 65 per cent to rat U4A RNA, 61 per cent to Drosophila U4A RNA and 50 per cent to snR14, the U4 RNA equivalent of the yeast Saccharomyces cerevisiae. Sequence conservation is much more pronounced in the 5' halves of U4 RNAs than in their 3' halves. Post-transcriptional modifications in broad bean U4B and U4A RNAs are confined to the 5' halves of the molecules, as is the case in all other U-RNAs

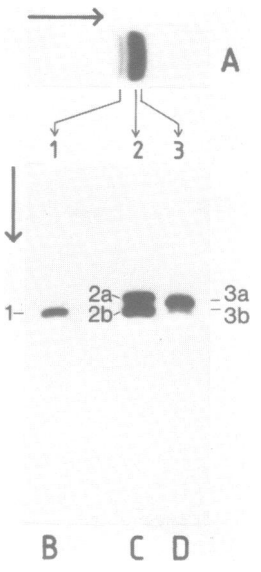

Fig. 6. Separation of sequence variants of 3' end-labeled U4 RNA from broad bean. Broad bean U4 RNA (Fig. 4, Lane B) was eluted from a preparative denaturing polyacrylamide gel run under the conditions as described in the legend to Fig. 1 and was labeled at the 3' end as specified in Materials and Methods. This sample was first loaded on a 20 cm long and 1 mm thick 15 % polyacrylamide (29:1, acrylamide/bisacrylamide) gel, run under non-denaturing conditions at 4° C and 20 V/cm for 18 hours in 50 mM Tris-borate buffer (pH 8.3) containing 1 mM EDTA, and subjected to autoradiography (upper horizontal Lane A). The RNAs from bands 1, 2 and 3 (marked by thin arrows) were then eluted separately, loaded each on a 20 cm long and 1 mm thick 12 % polyacrylamide (19:1, acrylamide/bisacrylamide) gel, run under semi-denaturing conditions at 37° C and 30 V/cm for 3 hours in 50 mM Tris-borate buffer (pH 8.3) containing 1 mM EDTA and 8M urea, and subjected to autoradiography (lower vertical Lanes B, C and D). The directions in which the gels were run are indicated by thick arrows, and the bands in the vertical lanes by numbers.

(except U6 RNA) sequenced so far [cf. (1)], but the modifications do not coincide in position with those in rat U4 RNA, except Am in position 67 of broad bean U4B RNA. However, in both sequence variants of broad bean U4 RNA modifications fall at identical positions in the aligned sequences. The Sm protein binding site (nucleotides 119 to 137 in broad bean U4B RNA) is present near the 3' end of both sequence variants of broad bean U4 RNA, as is the case with the other U4 RNAs. The stretch (nucleotides 58 to 65 in broad bean U4B RNA) shown for HeLa cells to base pair in vivo with U6 RNA (40) is also present in both sequence variants of plant U4 RNA, and is followed

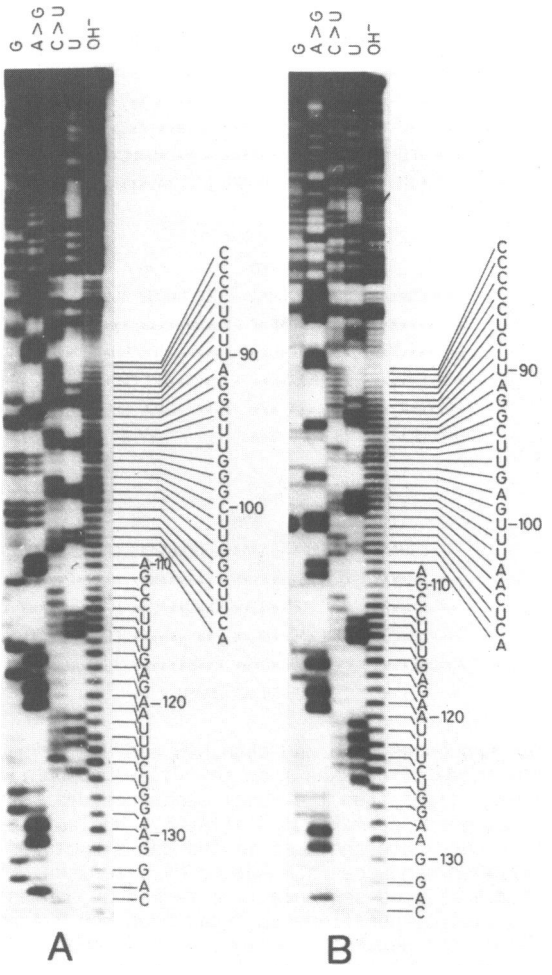

Fig. 7. Sequencing gels for broad bean U4A and U4B RNAs. 3' end-labeled U4A and U4B RNAs of broad bean were eluted from semi-denaturing polyacrylamide gels (cf. Fig. 6., bands 1 and 3a, respectively) and subjected to chemical sequence analysis as described in Materials and Methods. The degradation products were run on sequencing gels. Lane A shows the sequence from nucleotide 83 to 134 of broad bean U4A RNA and Lane B shows the sequence from nucleotide 82 to 133 of broad bean U4B RNA.

downstream (nucleotides 66-76 in broad bean U4B RNA) by another evolutionarily conserved sequence.

4. Possible Secondary Structure of U4A and U4B RNAs from Broad Bean

Fig. 9 presents a possible secondary structure model for broad bean U4B RNA. Differences in the secondary structure of broad bean U4A RNA from that of

Fig. 8. Nucleotide sequences of broad bean U4B and U4A RNAs compared with each other and with those of U4 RNAs from rat (53), Drosophila (54) and of SnR14, the U4 RNA equivalent from the yeast Saccharomyces cerevisiae (6). The sequences are aligned for maximum similarity. Nucleotide residues, methyl groups as well as phosphate residues in the cap structures, which are identical in the aligned sequences are marked by asterisks, whereas residues that are absent in the aligned sequences are marked by dashes. Figures on the top refer to nucleotide positions in broad bean U4B RNA and those in brackets at the end of the sequences designate the number of nucleotide residues in the individual U4 RNA molecules. The positions of methylations in SnR14 are not known and sugar methylation ⌈(m)⌋ of the first A residue in broad bean U4A and U4B RNAs is uncertain. In establishing sequence conservation (bottom line), modifications were disregarded. In the positions with transitions purines were marked by R, pyrimidines by Y and positions with either transversions or gaps in the aligned sequences were left blank. Phosphate residues are indicated only in the cap structure and pseudouridines are marked by F.

U4B RNA are given in insets. It can be seen that the Sm protein binding site (nucleotides 119 to 137), the stretch possibly involved in base pairing with U6 RNA (nucleotides 58 to 65) as well as an evolutionarily conserved sequence (nucleotides 66 to 76) downstream of the above one, are located in single stranded regions and may, therefore, be functional in molecular interactions.

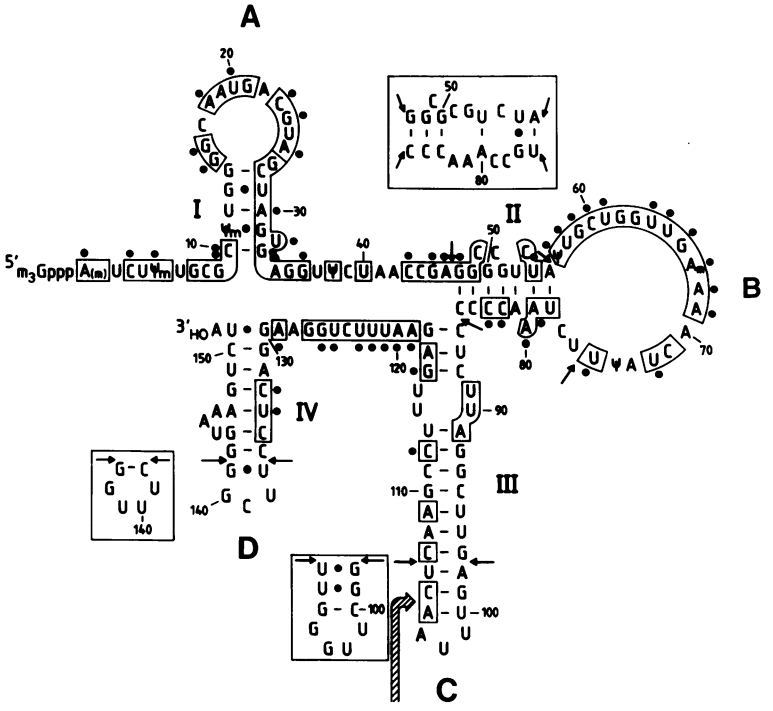

Fig. 9. Possible secondary structure of broad bean U4B RNA. This secondary structure model agrees with that of all metazoan U4 RNAs sequenced so far and is based on the model proposed by Rinke et al. (40) for HeLa U4 RNA. Stems are marked by roman numerals (I to IV) and loops by capital letters (A to D). Nucleotide positions are numbered. The site of insertion of the unique sequence of SnR14 (cf. Fig. 8) into U4 RNA is indicated by a hatched arrow. Regions of broad bean U4A RNA that differ in nucleotide sequence from those of broad bean U4B RNA are displayed in insets and are designated by arrows in both the model and the insets. Regions exhibiting sequence conservation as shown in the bottom line of Fig. 8 are boxed. Evolutionarily highly conserved individual nucleotides not involved in transition are marked by dots. Modifications are not considered.

5. Base Pairing of Broad Bean U4 RNA with Broad Bean U6 RNA

When nuclear extracts of Lycopersicon peruvianum were immunoprecipitated with anti-Sm mouse monoclonal antibody, Y12, all major U-RNAs that are precipitated by this antibody from eukaryotes other than plants (25) appeared in the precipitate (Fig. 10). This result, together with (i) the demonstration of the presence of a U4/U6 RNA complex within the same snRNP particle in HeLa cells (11, 12), (ii) our finding of the precipitability of U6 RNA by anti-cap antibodies from plant nuclear extracts (Fig. 2) and (iii) the sub-

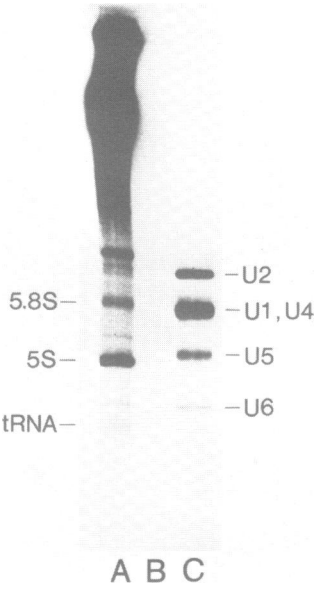

Fig. 10. U-RNAs present in immunoprecipitates obtained by treating nuclear extracts of [^{32}P]-labeled suspension culture cells of Lycopersicon peruvianum with anti-Sm mouse monoclonal antibody, Y-12 (Lane C) or normal serum (Lane B). As a reference non-immunoprecipitated RNAs present in the high salt nuclear extracts are shown (Lane A). The RNA bands were visualized by autoradiography. For more details see Materials and Methods.

stantiation of the occurrence of U4/U6 RNA complexes in broad bean nuclear RNA extracts (Fig. 3), indicate that also in plants U4 RNA is complexed with U6 RNA in the same snRNP particle. This conclusion is supported by a model of base pairing between broad bean U4 and U6 RNAs (Fig. 11).

DISCUSSION

This is the first report on the structural characterization of a plant U4 RNA. The apparent reason for earlier failures to even detect U4 RNA in plants lies in the fact that it is (as exemplified here in the lengths of broad bean and tomato U4 RNAs) somewhat longer (152-154 nucleotide residues) than "standard" size [141-145 nucleotide residues, cf. (1)] approaching the length (158 nucleotide residues) of snR14 (6). Therefore, under denaturing conditions it co-migrates in polyacrylamide gels with (a set of) U1 RNA (variants) and also 5.8S rRNA, thereby impeding visual detection. In addition, it probably occurs in the form of a number of sequence variants in

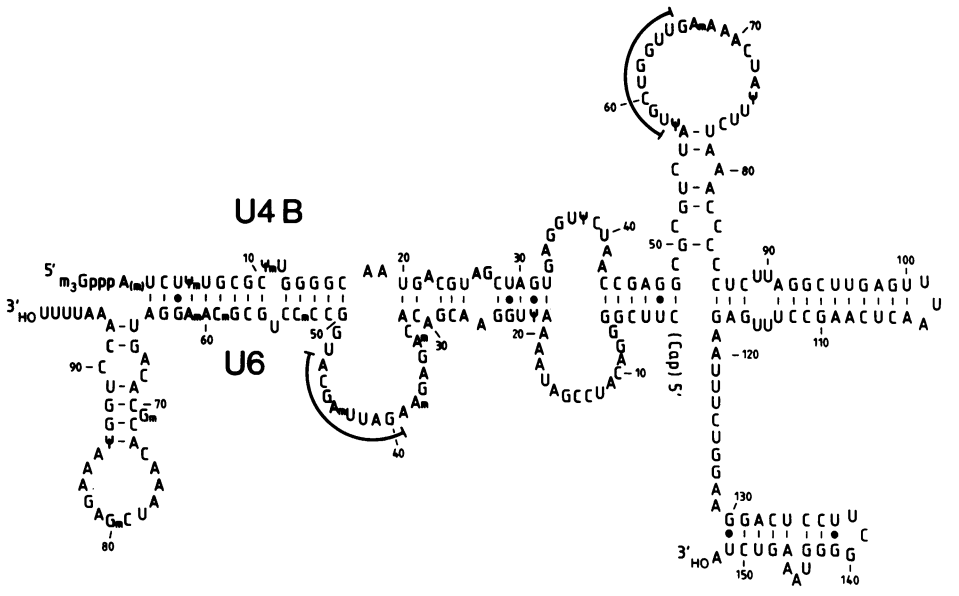

Fig. 11. Possible secondary structure of broad bean U4/U6 RNA complex. The sequence of broad bean U4B RNA was taken from Fig. 8 and that of broad bean U6 RNA from (18). The model essentially agrees with those proposed in (12) for the vertebrate U4/U6 RNA complex and in (41) for the Drosophila U4/U6 RNA complex. The U4/U6 cross-link sites reported in (40) for HeLa cells are indicated with heavy lines drawn above the sequences concerned.

plants, similar to broad bean, and this makes its routine isolation and purification a discouraging endeavor. However, by using a variety of conditions of gel electrophoresis and by taking advantage of its tendency to base-pair with U6 RNA, plant U4 RNA can be isolated in a pure enough state to lend itself to sequencing.

The high degree of sequence similarity of the 5' half of broad bean U4 RNA to U4 RNAs from other eukaryotes was to be expected, because broad bean U6 RNA also shows a very high degree (80 %) of sequence similarity to U6 RNAs from eukaryotes other than plants (18), and these two molecules are apparently base-paired within the same snRNP complex (11, 12). In this interaction it is the 5' half of U4 RNA that was implicated.

A salient point worth discussing, however, is the involvement of individual stretches of the U4 and U6 RNA molecules in base-pairing in vivo. In thermodynamic terms quite extended stretches of U4 RNA and U6 RNA can be assumed to interact (11, 12, 41), whereas results of psoralen cross linking experiments limited the range of such interaction to no more than 8 base

pairs (40). Our findings with plant U4 and U6 RNAs corroborate both models of base pairing. It should be noted that the two types of interaction do not have to be mutually exclusive. There is no a priori reason to suppose that only one type of base pairing is functional. During their operation in the splicing complex U4 and U6 RNAs may interact in multiple ways mediated by other RNA and/or protein factors. In addition, the occurrence of U4 RNA (or rather U4 snRNP) as a distinct entity in nuclei also gets some support from our data, indicating that plant U4 RNA can be fitted into a self-contained secondary structure model which, in turn, is identical with that of all other U4 RNAs sequenced so far, including snR14 (6) to which it shows only 50 per cent sequence similarity. Recently, Konarska and Sharp (42) have shown that U4/U6 snRNP disassembles and assembles during pre-mRNA splicing in vitro.

Determination of the entire sequence of a plant U4 RNA now completes the series of studies aimed at the full structural characterization of the six major plant U-RNAs. Up to now, the entire sequences of bean U1 RNA (43), Chlorella U1 RNA (44), broad bean U2 RNA (19), Arabidopsis U2 RNA variants (45), pea U5 RNA (15), and broad bean U6 RNA (18) have been determined. Data on a partial sequence of broad bean U3 RNA (17) will soon be complemented by the entire sequence of tomato U3 RNA (Kiss et al., to be published).

These studies clearly reveal that all of the structural features of the individual U-RNAs to which a functional role has been attributed in other eukaryotic U-RNAs are present in their plant equivalents. In addition, there are at least three more particular aspects which emerged from structural studies of plant U-RNAs.

(i) In different U-RNAs evolutionarily novel conserved sequences have been revealed that were "masked" in the U-RNAs from phylogenetically closely related species (15, 17, 18, 43, 44).

As far as U4 RNA is concerned we found, by including broad bean U4 RNA into the list of U4 RNAs to be compared, that the most pronounced sequence conservation, in addition to that of the Sm binding site near the 3' end of the molecule, is located in loop B (Fig. 9). The first part of this consensus sequence, UGCURRUU, is involved in base pairing with U6 RNA, and the rest of it, $GAAA(AC)_{0-1}U(A)_{0-1}\frac{U}{A}U\frac{U}{A}$, has been shown by oligonucleotide-targeted degradation to be the major part of an indispensable element needed for the in vitro splicing of pre-mRNA (46-48).

The apparent sequence conservation of loop A reflects potential base

pairing interactions with U6 RNA of evolutionarily very highly conserved primary structure. The same applies in general terms to the entirety of the 5' third of U4 RNA (nucleotides 2 to 48 in broad bean U4B RNA, Fig. 8). Stems III and IV as well as loops C and D (Fig. 9) do not seem to be involved in the above molecular interaction and show very little, if any, sequence conservation. Apparently the function of these regions depends much more on the mere existence of properly located stems and loops of appropriate lengths (in all U4 RNAs stem III is much longer than stem IV) than on their primary structure, and might be species-specific. In the yeast U4 equivalent stem IV and loop D are absent according to the secondary structure model proposed by Siliciano et al. (6), although the 3' terminal sequence of this particular U4 RNA equivalent would permit the construction of a rudimentary stem and loop structure.

(ii) The presence of Domain A including the Sm binding site in all plant equivalents of vertebrate and Drosophila U-RNAs that likewise contain it, points to the high evolutionary conservation of not only the RNA but also of the protein components of the U-snRNPs. Plant U-snRNPs could indeed be precipitated by anti-Sm antibodies [(20, 49) and this paper]. It should be noted, however, that whereas in all U2 and U4 U-RNAs from eukaryotes other than plants the oligopyrimidine track of the Sm protein binding site consists solely of U residues [cf. (1)], one U residue in this run of Us is replaced by C in broad bean U2 (19), in Chlorella U2 (Kiss et al., unpublished results), in broad bean U4 (this paper) RNAs and also in one of the sequence variants of Arabidopsis U2 RNA (45). This nucleotide replacement may be mirrored in some structural differences between the polypeptides of plant U-snRNPs and those of other eukaryotic snRNPs.

(iii) The easily detectable occurrence of sequence variants in plant U5 (15) and U4 (this paper) RNAs suggests that at least some of the major U-RNAs are particularly likely to be transcribed from gene families with slightly different, rather than identical sequences. An interesting aspect of this phenomenon would be to study the developmental control of U-RNA gene transcription in plants. The pioneering work of Forbes, Lund, and Dahlberg, [reviewed in (50)] clearly indicates that in Xenopus the expression of both U1 and U4 RNAs genes (29, 30, 51) and in mouse that of at least U1 RNA genes (52) are under developmental control. Plant systems lend themselves particularly well to studies of this sort. Taking advantage of the preferential synthesis of some specific proteins at certain stages of plant development may offer a unique opportunity to decide whether or not the developmen-

tally controlled accumulation of certain U-RNA sequence variants has anything to do with differential splicing of specific pre-mRNAs.

Another relevant aspect of the presence of U4 RNA sequence variants in broad bean nuclei is our finding that all base exchanges between U4A and U4B RNAs of broad bean are localized to loop-closing stem regions (Fig. 9). These make stems II and III thermodynamically more stable, and stem IV less stable in broad bean U4B RNA, than they are in broad bean U4A RNA. It is tempting to speculate that differential structural flexibility of the above regions in the molecules has something to do with specific roles of U4 RNA sequence variants in molecular interactions during splicing. Sequence data on more sequence variants of different U-RNAs from various organisms may throw some light on the validity of this conjecture.

ACKNOWLEDGEMENTS

We thank Dr. P. Bringmann and Professor R. Lührmann for their kind gift of anti-$m_3^{2,2,7}G$ IgG, Professors U. Pettersson, R. Reddy and Y. Ohshima for providing us with human U4, Drosophila U4 and mouse U6 DNA clones, respectively, Dr. K. Parker and Professor J.A. Steitz as well as Dr. I. Pettersson for sending us the mouse Y12 monoclonal antibodies, Dr. L. Nover for the Lycopersicon peruvianum callus tissue and last but not least Professor C. Guthrie and members of her group for helpful discussions. Thanks are due also to Mr. B. Dusha for preparing the photographs, Miss G. Kaliczka for drawing the figures and Mrs. M. Doktor for typing the manuscript. This work was supported by Grants KKA 8/1/400/2/85/1/118, OTKA 564, and OKKFT(Tt)/1986 to F.S. from the Hungarian Academy of Sciences.

*Parts I, II, III and IV are references (17), (18), (19) and (44), respectively.

+To whom correspondence should be addressed.

REFERENCES

(1) Reddy, R. and Busch, H. (1988) in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, Birnstiel, M.L. Ed., Springer Verlag, Berlin, pp. 1-37.
(2) Ares, M., Jr (1986) Cell 47, 49-59.
(3) Hughes, J.M.X., Konings, D.A.M. and Cesareni, G. (1987) EMBO J. 6, 2145-2155.
(4) Kretzner, L., Rymond, B.C. and Rosbash, M. (1987) Cell 50, 593-602.
(5) Siliciano, P.G., Jones, M.H. and Guthrie, C. (1987) Science 237, 1484-1487.
(6) Siliciano, P.G., Brow, D.A., Roiha, H. and Guthrie, C. (1987) Cell 50, 585-592.

(7) Patterson, B. and Guthrie, C. (1987) Cell 49, 613-624.
(8) Lührmann, R. (1988) in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, Birnstiel, M.L. Ed., Springer Verlag, Berlin, pp. 71-99.
(9) Riedel, N., Wolin, S. and Guthrie, C. (1987) Science 235, 328-331.
(10) Bringmann, P., Rinke, J., Appel, B., Reuter, R. and Lührmann, R. (1983) EMBO J. 2, 1129-1135.
(11) Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H. and Lührmann, R. (1984) EMBO J. 3, 1357-1363.
(12) Hashimoto, C. and Steitz, J.A. (1984) Nucleic Acids Res. 12, 3283-3293.
(13) Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Krämer, A., Frendewey, D. and Keller, W. (1988) in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, Birnstiel, M.L. Ed., Springer Verlag, Berlin, pp. 115-154.
(14) Muramatsu, M., Smetana, K. and Busch, H. (1963) Cancer Res. 23, 510-518.
(15) Krol, A., Ebel, J.P., Rinke, J. and Lührmann, R. (1983) Nucleic Acids Res. 11, 8583-8595.
(16) Lührmann, R., Appel, B., Bringmann, P., Rinke, J., Reuter, R., Rothe, S. and Bald, R. (1982) Nucleic Acids Res. 10, 7103-7113.
(17) Kiss, T., Tóth, M. and Solymosy, F. (1985) Eur. J. Biochem. 152, 259-266.
(18) Kiss, T., Antal, M. and Solymosy, F. (1987) Nucleic Acids Res. 15, 543-560.
(19) Kiss, T., Antal, M. and Solymosy, F. (1987) Nucleic Acids Res. 15, 1332.
(20) Tollervey, D. (1987) J. Mol. Biol. 196, 355-361.
(21) Nover, L., Kranz, E. and Scharf, K.D. (1982) Biochem. Physiol. Pflanzen 177, 483-499.
(22) Gersten, D.M. and Marchalonis, J.J. (1978) J. Immunol. Methods 24, 305-309.
(23) Bringmann, P. and Lührmann, R. (1986) EMBO J. 5, 3509-3516.
(24) Matter, L., Schopfer, K., Wilhelm, J.A., Nyffenegger, T., Parisot, R.F. and De Robertis, E.M. (1982) Arthritis Rheum. 25, 1278-1283.
(25) Lerner, M.R. and Steitz, J.A. (1979) Proc. Natl. Acad. Sci. USA 76, 5495-5499.
(26) Mimori, T., Hinterberger, M., Pettersson, I. and Steitz, J.A. (1984) J. Biol. Chem. 259, 560-565.
(27) Steele, W.J., Okamura, N. and Busch, H. (1965) J. Biol. Chem. 240, 1742-1749.
(28) Reddy, R., Li, W.-Y., Henning, D., Choi, Y.C., Nohga, K. and Busch, H. (1981) J. Biol. Chem. 256, 8452-8457.
(29) Forbes, D.J., Kirschner, M.W., Caput, D., Dahlberg, J.E. and Lund, E. (1984) Cell 38, 681-689.
(30) Lund, E. and Dahlberg, J.E., (1987) Genes Dev. 1, 39-46.
(31) Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
(32) Bark, C., Weller, P., Zabielski, J. and Pettersson, U. (1986) Gene 50, 333-344.
(33) Saba, J.A., Busch, H., Wright, D. and Reddy, R. (1986) J. Biol. Chem. 261, 8750-8753.
(34) Ohshima, Y., Okada, N., Tani, T., Itoh, Y. and Itoh, M. (1981) Nucleic Acids Res. 9, 5145-5158.
(35) Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
(36) England, T.E., Bruce, A.G. and Uhlenbeck, O.C. (1980) Methods in Enzymology 65, 65-74.
(37) Peattie, D.A. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 1760-1764.

(38) Donis-Keller, H., Maxam, A. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527-2538.
(39) Sanger, F. and Coulson, A.R. (1978) FEBS Lett. 87, 107-110.
(40) Rinke, J., Appel, B., Digweed, M. and Lührmann, R. (1985) J. Mol. Biol. 185, 721-731.
(41) Das, G., Henning, D., Busch, H. and Reddy, R. (1987) J. Biol. Chem. 262, 1187-1193.
(42) Konarska, M.M. and Sharp, P.A. (1987) Cell 49, 763-774.
(43) van Santen, V.L. and Spritz, R.A. (1987) Proc. Natl. Acad. Sci. USA 84, 9094-9098.
(44) Kiss, T., Antal, M., Hegyi, H. and Solymosy, F. (1988) Nucleic Acids Res. 16, 2734.
(45) Vankan, P. and Filipovicz, W. (1988) EMBO J. 7, 791-799.
(46) Berget, S.M. and Robberson, B.L. (1986) Cell 46, 691-696.
(47) Chabot, B. and Steitz, J.A. (1987) Mol. Cell. Biol. 7, 281-293.
(48) Krämer, A. (1987) J. Mol. Biol. 196, 559-573.
(49) Tollervey, D. and Mattaj, I.W. (1987) EMBO J. 6, 469-476.
(50) Dahlberg, J.E. and Lund, E. (1988) in Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, Birnstiel, M.L. Ed., Springer Verlag, Berlin, pp. 38-70.
(51) Lund, E.,Bostock, C.J. and Dahlberg, J.E. (1987) Genes Dev. 1, 47-56.
(52) Lund, E., Kahan, B. and Dahlberg, J.E. (1985) Science 229, 1271-1274.
(53) Reddy, R., Henning, D. and Busch, H. (1981) J. Biol. Chem. 256, 3532-3538.
(54) Myslinski, E., Branlant, C., Wieben, E.D. and Pederson, T. (1984) J. Mol. Biol. 180, 927-945.