Identification of a gene encoding the predicted ribosomal protein L7b divergently transcribed from POL1 in fission yeast Schizosaccharomyces pombe

Véronique Damagnez, Anne-Marie de Recondo and Giuseppe Baldacci*
Laboratoire de Biologie Moléculaire de la Réplication UPR 272, IRSC-CNRS, 7 rue Guy Moquet, 94800 Villejuif BP 8, France

Received November 19, 1990; Revised and Accepted February 7, 1991
EMBL accession no. X54983

ABSTRACT
A 0.85 Kb RNA molecule is transcribed in the region upstream from the 5'-end of the S. pombe POL1 gene encoding the catalytic subunit of DNA polymerase alpha. The nucleotide sequence of the DNA region hybridizing with the 0.85 Kb transcript allowed us to identify an open reading frame coding for a predicted peptide which shows 50% identity with the rat ribosomal protein L7 and which is transcribed divergently from POL1. We have named this gene RPL7b because of the existence in S. pombe of a different sequence, named RPL7, which also codes for a putative protein showing homology with the rat ribosomal protein L7. The RPL7b gene includes a 291 bp-long intron containing the sequences necessary for intron excision and RNA splicing in S. pombe. The precise location of the intron was established by amplification and sequencing of a partial cDNA copy of the mRNA, whereas the initiation site of transcription was determined by reverse transcription of the 5' region of the mRNA. The 320 bp separating the starting methionine codons of RPL7b and POL1 genes should contain the signals necessary for their divergent transcription and regulation. The sequence 5'-AAGACAGTCACA-3', whose primary structure is homologous to a conserved block present in the 5'-untranscribed regions of other S. pombe genes of ribosomal proteins, is located about 50 bp upstream the transcription initiation site of RPL7b.

INTRODUCTION
In the course of our study of genes coding for proteins related to DNA replication in fission yeast Schizosaccharomyces pombe, we have observed the presence of a transcript hybridizing immediately upstream the POL1 gene, which codes for the catalytic subunit of DNA polymerase alpha and is located on chromosome II (1). As we intend to characterize the promoter region of the POL1 gene, we studied the gene corresponding to the upstream hybridizing transcript. The predicted peptide coded for by this gene shows significant homology with the ribosomal protein L7 from rat and mouse cells (2, 3) and with a predicted L7 peptide from Dictyostelium discoideum (4). Thus, we have named this gene RPL7b. Recently, another S. pombe sequence coding for a predicted peptide homologous to the rat L7 protein has been described, and its locus on chromosome I has been named RPL7 (5). In contrast, RPL7b is located on chromosome II and its nucleotide sequence is different.

The synthesis of ribosomes requires the coordinated expression of many genes and involves several steps of protein and RNA transport across the nuclear membrane in eukaryotic cells (6). The analysis of ribosomal proteins in different eukaryotic cells has shown a large conservation of their primary structures, facilitating inter-specific identifications. In budding yeast Saccharomyces cerevisiae approximately 80 ribosomal proteins have been detected by two-dimensional polyacrylamide gel electrophoresis (6). The steady-state amounts of ribosomal proteins are equimolar, and the rate of synthesis is adjusted to the cellular growth rate and is influenced by the carbon source (6). In S. cerevisiae the genes encoding ribosomal proteins seem mainly regulated at the transcription level. An upstream cis-acting sequence fixing the transcription factor TUF has been identified by sequence homology and functional characterization (7). In S. pombe the genes of 4 ribosomal proteins have been described; their 5'-untranscribed regions contain some homologous sequences and one of them (homol d) is particularly well conserved in location and primary structure (8). We have observed a related sequence at a similar position upstream from the site for transcription initiation of RPL7b.

MATERIALS AND METHODS
Strains and culture media
A haploid S. pombe leu1-32 h- strain was grown in 1% yeast extract and 3% glucose at 32°C until exponential phase. Cells were synchronized following published procedures (9). The E. coli strain JM105 was used for the growth of M13-derived phages (10).
Nucleic acids preparation and hybridizations

DNA was prepared from spheroplasts following published procedures (1). RNAs were extracted from cells broken with glass beads. Poly A⁺ RNAs were prepared with a mRNA purification kit according to the manufacturer's instructions (Pharmacia, Uppsala, Sweden).

DNA hybridizations were carried out with 50 ng of ³²P-labeled DNA probes; 10⁶ cpm/ng were usually obtained using a Random Primed DNA labeling kit (Boehringer, Mannheim, RFG); assays were performed at 65°C in 5 × SSC and washes in 0.1 × SSC at 65°C. Northern blot hybridizations were carried out in 5 × SSC and 50% formamide at 48°C; washes were performed in 2 × SSC at 53°C.

Figure 1. Schematic restriction enzyme map of the S. pombe 1.8 Kb DNA region corresponding to RPL7b and flanking sequences. Arrows indicate the transcription direction of POLI and RPL7b genes; the starting points of arrows correspond to the AUG initiation codons. A black box indicates the intron and an asterisk marks the stop codon of RPL7b. The dashed line corresponds to the nucleotide sequence shown in figure 2.

Figure 2. Nucleotide sequence of RPL7b and predicted L7b peptide. Non-coding sequences are indicated in lower cases. Bold characters indicate a motif conserved in the regions upstream from the S. pombe genes of ribosomal proteins. The asterisk indicates the stop codon. Underlined nucleotides in upper cases correspond to the synthetic oligonucleotides used for primer extension and cDNA amplification. Underlined nucleotides in lower cases correspond to the 5' end of the RPL7b mRNA (−13 to −10) and to a putative signal of polyadenylation (1120 to 1125).
DNA sequencing

The restriction fragments indicated in figure 1 were subcloned in M13mp18 and M13mp19 phage vectors such that a set of overlapping single-stranded clones spanned the entire fragment in both directions. The nucleotide sequence of inserts was determined using a Sequenase Version 2.0 DNA sequencing kit (U.S.B., Cleveland, Ohio, USA). The products of sequencing reactions were resolved in denaturing polyacrylamide gels and revealed by autoradiography.

Transcripts analysis

SI Nuclease mapping

Five μg of poly A+ RNA was hybridized to single-stranded DNA from an M13mp18 recombinant phage containing the 1.6 Kb Dral DNA fragment complementary to the RPL7b mRNA. The Dral fragment extends beyond the 5' and 3' limits of the transcribed region (figures 1 and 2). Hybrids were digested with 100 units of SI Nuclease (Boehringer, Mannheim, RFG) at 37°C for 30 minutes, and protected fragments were resolved in a 1.7% agarose gel. After blotting, the nitrocellulose filter was hybridized with the 32P-labeled 1.6 Kb Dral DNA fragment.

 Primer extension analysis of RNA

The 20-mer oligonucleotide 5'-GCTTCATAAGCCTCAGCAGCTCAG-3' complementary to positions 451-470 of the RPL7b coding strand was synthesized (figure 2). 150 μg of total RNA was co-precipitated with the 5' end-labeled synthetic primer and treated with M-MLV reverse transcriptase (GIBCO-B.R.L., France) as described (1). The dNTP mixture was, however, modified: dNTPs were 2mM and ddNTPs were 2,5mM. The products of reverse transcription were resolved in denaturing polyacrylamide gels and revealed by autoradiography.

cDNA amplification

A partial cDNA copy of the RPL7b mRNA obtained as described above (without ddNTPs) was amplified using as primers the oligonucleotide complementary to positions 451-470 (5'-GCTTCATAAGCCTCAGCAGCTCAG-3') and the oligonucleotide identical to positions 43-60 (5'-GCTCCCGAGTCTCTTCTC-3') of figure 2. Standard procedures for cyclic DNA amplification at high temperature were used.

Computer analysis

N.B.R.F. and E.M.B.L. protein data banks were scanned for homology with the predicted L7b peptide using the FASTA program (11).

RESULTS

A partial restriction map of the region spanning the 5'-end of the POL1 gene is shown in figure 1. DNA probes from this region hybridize to the 4.5 Kb POL1 transcript when sequences located 5' to the Hpal site are included (1); on the other hand, probes containing sequences located 3' to the Hpal site detect a transcript of about 0.85 Kb (figure 3). The DNA region coding for the latter transcript was sequenced on both strands. All restriction sites used for subcloning in M13-derived phages were 'sequenced-through'. The primary structure of the region corresponding to the dashed line of figure 1 along with the predicted peptide coded for by the transcribed strand are shown in figure 2. The RPL7b gene is divergently transcribed from POL1 and the two starting methionine codons are separated by 320 bp.

The presence in RPL7b of an intron which is processed in the course of the transcript maturation was detected by SI Nuclease digestion of polyA+ RNAs hybridized to a single-stranded DNA fragment complementary to the transcribed strand of RPL7b (figure 3). The precise location of the intron was established by sequencing an in vitro amplified DNA segment complementary to the RPL7b mRNA (figure 4). The intron contains the sequences corresponding to the S. pombe consensus sequences necessary for RNA splicing (12): 5'-GTAAGT-3' (positions 127-132 in figure 2) at the 5'-end; 5'-CTAAT-3' (positions 403-407 in figure 2) at the 3'-end; 5'-CTAAT-3' (positions 403-407 in figure 2) at the 3'-end.

Figure 3. 1. Autoradiogram of the SI Nuclease digest of S. pombe poly A+ RNA/RPL7b gene hybrids. 5μg of poly A+ RNA was hybridized with the 1.6 Kb single-stranded Dral DNA fragment (figure 1) complementary to the RPL7b mRNA. SI Nuclease-resistant hybrids were revealed with the 32P-labeled double-stranded Dral DNA fragment. Sizes were determined by comparison with the migration of molecular weight standards. The arrowhead marks the position of a band evident after a prolonged period of time. 2. Autoradiogram of S. pombe poly A+ RNAs hybridized with a RPL7b gene probe. 1μg of poly A+ RNA was loaded in the slot and the filter was hybridized with the 32P-labeled 250 bp Sspl DNA fragment of figure 1. Size was determined by comparison with the migration positions of mammalian and S. cerevisiae rRNAs.

Figure 4. A autoradiogram of a partial cDNA copy of the RPL7b mRNA followed by high-temperature amplification. The arrow indicates the base residues separated by 291 bp in the DNA sequence which are joined in the mRNA. Lower T corresponds to position 126 while upper G corresponds to position 418 in figure 2. b. Autoradiogram of 5'-end mapping of RPL7b mRNA. The products of reverse transcription of RNAs primed with a synthetic oligonucleotide were loaded onto a denaturing polyacrylamide gel. Size was determined with the M13mp18 sequence as a migration standard: the band in the slot marked 5' is 191 bases long.
figure 2) at the branch site; 5'-TAG-3' (positions 415-417 in figure 2) at the 3'-end. The attempted hybridization of an intron probe with polyA+ RNAs blotted onto nitrocellulose did not reveal hybridization bands (data not shown). Interestingly, this is one of the longest introns known to date in the S. pombe genome (12).

The transcription initiation site of RPL7b was determined by reverse transcription of the 5'-end of the mRNA using a complementary synthetic oligonucleotide as a primer (figure 4). We also performed the primer extension reaction in the presence of dideoxyribonucleotides to check the specificity of hybridization. We also performed the primer extension reaction in the presence of dideoxyribonucleotides to check the specificity of hybridization.

We determined the relative production of the RPL7b mRNA in the course of the cell cycle using the URA4 mRNA as an internal standard (figure 5). The URA4 gene codes for the orotidine-5'-phosphate decarboxylase (15, 16). It does not show cell cycle-related changes in mRNA quantity, a feature common to the homologous URA3 gene from S. cerevisiae (17, 18). The differences in the hybridization signals of the RPL7b transcript in figure 5 are not repeated at identical points of two successive cell cycles and the RPL7b mRNA shows an identical pattern of variations. Thus, these differences seem to be due to different amounts of RNA loaded in the slots, and we conclude that the level of the RPL7b mRNA does not vary significantly in the course of the cell cycle.

We hybridized an internal RPL7b gene probe with enzyme-restricted S. pombe DNA under stringent conditions to detect other copies of this gene. In addition to single prominent bands fitting well with the restriction map of figure 1, a second faint band is also present in each hybridization profile of figure 6. These results indicate the presence in the genome of S. pombe of another sequence similar to the RPL7b gene.

The alignments of the S. pombe predicted L7b peptide with the ribosomal protein L7 from rat cells and with the predicted L7 peptides from D. discoideum and S. pombe are shown in figure 7. We have aligned the protein sequences by visual inspection in order to optimize identity and to reduce insertions or deletions; we have calculated the identity percentages on the basis of this alignment.
DISCUSSION

The identification of the gene coding for the 0.85 Kb transcript hybridizing to the 5'-region of the POL1 gene is based on several arguments. When the predicted peptide is compared with the protein sequences present in the N.B.R.F. and E.M.B.L. data banks, the only significant homology is with the L7 ribosomal proteins from D. discoideum and rat cells. The alignment of the predicted L7b peptide with protein L7 of rat cells shows 50% identity; with the predicted L7 peptides of D. discoideum and S. pombe 47% and 53% identity, respectively (figure 7). In figure 7 we have not aligned the mouse L7 protein sequence because it shows 96% identity with the rat L7 protein (3). The predicted L7b peptide of S. pombe is 250 amino acids long which corresponds to 28,432 of relative molecular mass. It shows a calculated isoelectric point of 10.15 and contains 50 residues of arginine, lysine, and histidine versus 25 residues of aspartic and glutamic acid. These features indicate a basic ribosomal protein. Only 44 codons are used for coding the L7b protein and their distribution is characteristic of S. pombe highly expressed genes (19).

Back-hybridizations of a RPL7b gene probe with enzyme-restricted DNA of S. pombe under stringent conditions showed the presence of an additional faint hybridization band in each profile (figure 6). These results can be accounted for by the presence of the RPL7 sequence coding for a putative peptide homologous to the rat L7 protein (5). RPL7 and RPL7b share 64% nucleotide identity in the regions 3' to their respective introns (computer alignment not shown). In S. pombe multiple copies of the gene coding for the ribosomal protein K37 have been shown (13). A difference with the majority of known genes encoding ribosomal proteins in fission yeast is the presence of an unusually long intron in RPL7b. This intron could have a specific function in the transcriptional control of RPL7b gene expression. Actually, when the intronless gene encoding the ribosomal protein K37 is inserted in a multicopy expression vector, the increase in transcript concentration is not accompanied by a parallel increase in the protein (13). This indicates that the amount of ribosomal protein K37 is controlled at the translation level, as are the genes of ribosomal proteins in mammalian cells (13). However, in S. cerevisiae the transcriptional control of concentration of ribosomal proteins seems to be the rule, and the presence of introns in the corresponding genes has been related to this form of control (6).

The sequence 5'-AAGACAGTCACA-3' (positions −65 to −54 in figure 2) located about 50 bp 5' to the transcription initiation site of RPL7b, is a feature common to other S. pombe genes of ribosomal proteins. This sequence fits well with the homol d block (8) present at a similar distance upstream from other genes for ribosomal proteins in S. pombe (table 1). This sequence appears to be the unique conserved motif upstream from the genes of ribosomal proteins (computer alignments not shown) and could be functionally equivalent to the RPG box, which is present in the untranscribed leader regions of several genes of ribosomal proteins in S. cerevisiae (7). The RPG box has a transcription enhancer function which is controlled by the transcription factor TUF (7). The homol d block could have a similar role in S. pombe. Actually, the homol d block is present in the opposite orientation upstream from the gene of the ribosomal protein S6. This could suggest an enhancer function (table 1).

Another interesting feature of the sequence comprised between the RPL7 and POL1 genes is the initiation of divergent transcripts in this region. This form of gene organization has been described between genes coding for structural (non-regulatory) or regulatory polypeptides in a variety of organisms (20). Recently, a divergent transcript has been detected upstream from the human gene encoding the proliferating cell nuclear antigen (PCNA), which is an auxiliary protein of DNA polymerase delta (21). In S. pombe some histone genes are divergently transcribed (22). The occurrence of contiguous divergent promoters has been related to a possible cell coordination of transcription of two genes (20). However, the regulation of transcription can be independent and the presence of a ‘strong’ promoter could influence the transcription of a divergent RNA molecule by increasing the local concentration of RNA polymerase (20). In the present case, the steady-state amount of RPL7b mRNA from exponentially growing cells is more abundant than the amount of POL1 mRNA, as judged by the time necessary to detect hybridizations in identical experimental conditions. In addition, the divergent transcripts are regulated differently. In fact, the RPL7b mRNA level does not change in the course of the cell cycle (figure 7), whereas, on the contrary, the POL1 gene transcript shows cell cycle-related variations (paper in preparation).

Thus, the 320 bp separating RPL7b and POL1 genes include: i) the signals necessary for their divergent transcription; ii) the signals allowing the coordinated synthesis of the RPL7b mRNA with those of other ribosomal proteins (e. g. the homol d block); iii) the signals allowing the coordinated synthesis and the cell cycle regulation of the POL1 mRNA with those encoding other components of the S. pombe replisome. In conclusion, we believe that this region is particularity suitable for studying regulatory DNA elements and their interacting transcription factors in S. pombe.

ACKNOWLEDGMENTS

We thank our colleague D. Bouvier for help in cell synchronization experiments, K. Maundrell for the gift of plasmid DNA containing the URA4 gene, A. Dias and J. Tillit for skillful technical assistance. This work was supported in part by grant no. 6704 from A.R.C. to G. B.; V. D. is recipient of a M.R.T. fellowship and A-M. d.R. is recipient of a L.N.F.C. grant.

REFERENCES