The 26S rRNA binding ribosomal protein equivalent to bacterial protein L11 is encoded by unspliced duplicated genes in *Saccharomyces cerevisiae*

M.G.Pucciarelli, M.Remacha, M.D.Vilella and J.P.G.Ballesta*
Centro de Biología Molecular (CSIC-UAM), Universidad Autónoma, Canto Blanco, 28049 Madrid, Spain

Received May 17, 1990 Revised and Accepted June 21, 1990 EMBL accession nos X51519, X51520

ABSTRACT

Transformant phages expressing L15, a yeast ribosomal protein which binds to 26S rRNA and interacts with the acidic ribosomal proteins, were isolated by screening a yeast cDNA expression library in λgt11 with specific monoclonal antibodies. Using yeast DNA HindIII fragments that hybridize with the cDNA insert from the L15-expressing clones, minilibraries were prepared in pUC18, which were afterward screened with the same cDNA probe. In this way, plasmids carrying two different types of genomic DNA inserts were obtained. The inserts were subcloned and sequenced and we found a similar coding sequence in both cases flanked by 5′ and 3′ regions with very low homology. Sequences homologous to the consensus TUF-binding UAS boxes are present in the 5′ flanking regions of both genes. Southern analysis revealed the presence of two copies of the L15 gene in the *Saccharomyces cerevisiae* genome, which are located in different chromosomes. The encoded amino acid sequence corresponds, as expected, to protein L15 and shows a high similarity to bacterial ribosomal protein L11.

INTRODUCTION

The large bacterial ribosomal subunit has, among other structural features, a typical lateral protuberance, the L7/L12 stalk, that plays an important role in the interaction of the supernatant factors with the ribosome during protein synthesis (1). The stalk is made of two dimers of the acidic proteins that, interacting with protein L10, form a pentameric complex able to bind directly to the 23S rRNA (2,3).

On the other hand, protein L11 interacts with the pentamer (L7/12)4-L10 in solution (4), binds to an almost identical rRNA site (2) and affects its interaction with the ribosome (5).

The proteins L7/L12, L10 and L11 are easily removed from the bacterial ribosome by ammonium-ethanol washing (5), leaving core particles that preserve some of the ribosomal activities (6).

By treating eukaryotic ribosomes in similar conditions, a family of acidic proteins, structurally and functionally equivalent to L7/L12 are released (7–9). In addition, a basic protein from the large ribosomal subunit is recovered (10,11). In *Saccharomyces cerevisiae* the acidic protein family is formed by three major phosphoproteins, L44, L44′ and L45 (12) and a fourth minor component of similar size (13). The additional ethanol-extractable basic polypeptide was identified as protein L15 (9).

The yeast protein L15 has been shown to interact with the acidic ribosomal proteins in solution forming an association that, although less stable, is reminiscent of the complex formed by bacterial proteins L10 and L7/L12 (14). At the same time, this protein binds to the 26S rRNA at a highly conserved site equivalent to the 23S rRNA region where proteins L10 and L11 bind in the bacterial ribosome (15). Previously, yeast protein L15 and bacterial protein L11 were shown to be functional and immunological equivalents (11).

All these data suggested the possibility that the yeast protein L15 might play the structural role that proteins L10 and L11 play in bacteria, forming a complex with the acidic proteins which interacts with the rRNA (14).

A larger eukaryotic ribosomal component, protein P0, that also seems to be structurally related to the acidic proteins has been reported by different authors (16–18). P0 has a molecular mass of about 30 kDa (40 kDa in mammals) and has been reported to possess the highly conserved carboxyl terminal amino acid sequence typical of all the eukaryotic acidic proteins, but contrary to them, is not present in the cytoplasmic pool (19). Based on a limited amino acid sequence homology, protein P0 has been proposed to be equivalent to bacterial protein L10 (20) but in contrast with this protein, P0 is not released from the ribosome by ammonium-ethanol (10,16). There are no biochemical studies on the role of this protein, but it would be interesting to study its functional relationship to protein L15.

In order to analyze in detail the structure and function of L15, cloning of the gene encoding this protein was attempted and the results are presented in this report.

* To whom correspondence should be addressed
MATERIALS AND METHODS

Yeast, bacteria phage and plasmids

*Saccharomyces cerevisiae* Y166 was used throughout this work. Cells were grown in 1% yeast extract, 2% peptone and 2% glucose.

The cDNA libraries in phage λ gt11 (21), a gift from Dr. Altman and Dr. Trachsel (Institut für Biochemie und Molekularbiologie, Bern), were expanded in *Escherichia coli* Y1088. *E. coli* Y1090 was used for plating recombinant phages. Subcloning was carried out in plasmid pUC18 (22) using *E. coli* JM83 (23) for transformation.

Minibanks of HindIII fragments were prepared by extracting the selected bands from low melting point agarose gels with CETAB (24) and inserting them into the corresponding site of plasmid pUC18.

Enzymes

Restriction endonucleases, T4 DNA ligase, Klenow *E. coli* DNA polymerase I fragment, etc., were purchased from either Boehringer Mannheim (West Germany), New England Biolabs (USA) or Amersham (U.K.).

Antibodies

Hybridomas were obtained by fusing spleen cells from immunized Balb/c mice and X63/Ag8.653 mouse myeloma cells (25) according to standard procedures (26). Monoclonal specificity was determined by ELISA and immunoblotting.

Screening of recombinant phages

The expression cDNA library was screened with specific antiprotein L15 antibodies as described (27) using peroxidase-linked anti-mouse sera (Nordic Immunology, The Netherlands) to detect bound IgGs.

Recombinant DNA techniques

DNA preparation, restriction enzyme digestions, agarose gel electrophoresis and ligation of DNA fragments were carried out according to standard techniques (28).

DNA inserts in pUC18 were sequenced by the dideoxy chain termination method (29) using universal primers and complementary oligonucleotides synthesized in an Applied Biosystems 381A Synthesizer.

Radioactive labeling of DNA was performed with either (γ-32P)ATP and T4 polynucleotide kinase or (α-32P)dCTP and the Klenow DNA polymerase I fragment (28).

Electrophoresis methods

Pulse field electrophoresis of yeast DNA was carried out in a 55 cm double inhomogeneous PFG apparatus for 40 h at 300 V with 100 sec pulses and 90° field orientation as described (30).

Southern blots

DNA from agarose gel was transferred to nitrocellulose and the paper was treated for 1 h at room temperature with 0.5 M NaOH and 1.5 M NaCl for DNA denaturation, followed by neutralization in 1M Tris-HCl, pH 7.7 and 1.5 M NaCl and finally kept at 45°C for 8 h in SSPE×5 (28), Denhardt’s x5 (28), 100 μg/ml denatured salmon sperm DNA and 50% formamide. The denatured 32P-labeled probes were then added and the incubation was prolonged for 16 h. The papers were then washed at 45°C for 30 min first in SSPE×2 and 0.1% SDS then in

![Fig. 1. Southern blots of yeast DNA digested with Hind III (1) and EcoRI (2) and resolved by agarose gel electrophoresis. Blots were hybridized with either the cDNA insert from a positive phage clone (A), a DNA probe from the L15a gene (B) or a DNA probe from L15b gene (C). SSPE×1 and 0.1% SDS and finally twice in SSPE×0.1 and 0.1% SDS.](image)
the small EcoRI-HindIII DNA fragments from both clones (Fig. 2) were used as probe on Southern blots of yeast DNA restricted with HindIII and EcoRI. Only two bands were found to hybridize with the probes in the DNA treated with both restriction enzymes (Fig. 1B and C), confirming the existence of only two copies of the protein L15 gene. As expected, a relatively high degree of cross-hybridization is detected but the probes hybridize more strongly with the homologous copy.

**Chromosome mapping of the protein L15 gene**

Using an EcoRI-HindIII fragment from the protein L15b gene (Fig. 2) as a probe, a Southern blot of yeast chromosomes resolved by pulsed field electrophoresis was performed (Fig. 3). Two bands showing substantially different intensity were found to hybridize with the DNA probe. The stronger signal corresponded obviously to the L15b gene copy, homologous to the probe, and the weaker spot to the L15a gene. Both hybridizing bands are, however, made of two chromosomes (33); chromosomes IV and XII are present in the slower spot, carrying the L15b gene, and chromosomes V and VIII in the faster one, carrying the L15a gene.

**Sequencing of the genomic DNA inserts**

The DNA inserts in pUC18 were sequenced in both directions by the dideoxy chain termination method, starting from the end proximal to the gene position, using the universal primer first and then synthetic complementary oligonucleotides, according to the strategy indicated (Fig. 2).

Fig. 4 shows the sequence of about 1330 nucleotides in both inserts covering the complete coding region, the part of the 3' flanking region present in the cloned fragments and around 500 nucleotides from the 5' flanking region. A number of significant structural features in the sequence are summarized in Table 1. The deduced amino acid sequence of a 16931.63 daltons protein that can be identified as L15 based on the previously reported amino acid sequence of a short amino terminal fragment of this protein (34), is also shown. The relative synonymous codon usage in the expression of both copies (not shown), estimated using the GCG codon frequency program, corresponds to a highly expressed gene according to the statistical analysis of 110 yeast genes performed by Sharp et al. (35).

**Analysis of protein L15 amino acid sequence**

The amino acid sequence of protein L15 deduced from the nucleotide sequence in the cloned genes has been compared with
the sequence of other ribosomal proteins from yeast and bacteria using the Bestfit and Fasta programs. Considering as background average values obtained from 10 unrelated non ribosomal proteins from bacteria and yeast, the results indicated the existence of a significant similarity between protein L15 and bacterial protein L11 (Table 2); the comparison parameters are, however, scarcely higher than the background when protein L15 is matched with bacterial protein L10. The similarity of proteins L15 and L11 along extended parts of their sequence is shown by the long diagonal runs of matches in the dot-plot diagrams of the comparison data obtained by the Compare program (Fig. 5).

The ratio and percent identity values found when protein L15 is compared to the acidic proteins is also higher than the background. The significance of this similarity increases substantially when the highly conserved carboxyl terminal part of the acidic proteins is excluded from the comparison. Interestingly, protein L44' differs from the other acidic proteins in having lower similarity to protein L15.
TABLE 1

Structural features in the sequence of protein L15 genes

<table>
<thead>
<tr>
<th>L15a gene</th>
<th>L15b gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>TUF-UASrpg</td>
<td></td>
</tr>
<tr>
<td>-528 ACATCCGAACACCA</td>
<td>-514 -503 ACATCCGTGCACCTG</td>
</tr>
<tr>
<td>*-499 ACCACCAACACATAT</td>
<td>*-474 ACAACCCAAACATC</td>
</tr>
<tr>
<td>-384 TCCGCCGAACATCC</td>
<td>-397 -167 ACACCAGGACAGGT</td>
</tr>
<tr>
<td>-298 CCGTCCGTACAGTT</td>
<td>-285 -111 ATGGCGGTACATTA</td>
</tr>
<tr>
<td></td>
<td>+92 AGATCGGTCCATTG</td>
</tr>
<tr>
<td></td>
<td>+517 AAATCCATTCATA</td>
</tr>
<tr>
<td>SUF-UASrpg</td>
<td></td>
</tr>
<tr>
<td>-135 ATCTTTCTCAATG</td>
<td>-123</td>
</tr>
<tr>
<td>-329 GTGACTTTTATCT</td>
<td>-317</td>
</tr>
<tr>
<td>-418 ATCATAAACAGG</td>
<td>-406</td>
</tr>
<tr>
<td>Direct repeats</td>
<td></td>
</tr>
<tr>
<td>-493 AAAGAGACAATGTAATG</td>
<td>-477 -472 TATTTTGCTT</td>
</tr>
<tr>
<td>-480 AATCAGACAATGCTAGT</td>
<td>-464 -448 TATTTTGCTT</td>
</tr>
<tr>
<td>-402 TTGTACCTACAAGGCC</td>
<td>-387 -303 TACTTCGCTT</td>
</tr>
<tr>
<td>+37 TTGTACCTAAGGCC</td>
<td>+52</td>
</tr>
<tr>
<td>-128 AAATTATTCAAA</td>
<td>-117</td>
</tr>
<tr>
<td>+323 AAATTATGAAA</td>
<td>+334</td>
</tr>
<tr>
<td>Inverted repeats</td>
<td></td>
</tr>
<tr>
<td>-539 CTACCCTGTT</td>
<td>-530</td>
</tr>
<tr>
<td>+227 CTACCCTGTT</td>
<td>+218</td>
</tr>
<tr>
<td>-456 ACTATAACAGTATCA</td>
<td>-442</td>
</tr>
<tr>
<td>+237 TTTGCTATTACTGCTT</td>
<td>+253</td>
</tr>
<tr>
<td>Complementary inverted</td>
<td></td>
</tr>
<tr>
<td>-533 TGTCTTAACATCG</td>
<td>-521 -508 TTTAATACATCCG</td>
</tr>
<tr>
<td>-493 ACAACTGTGGGT</td>
<td>-505 -478 GCCTTGTAGGGT</td>
</tr>
<tr>
<td>-62 ACCTTTAAGGA</td>
<td>-52 -484 GTTTGCTTGTTGATTTT</td>
</tr>
<tr>
<td>-37 TGGGTTTCTCTTCTTCTT</td>
<td>-46 -375 TAGACCCCAATACG</td>
</tr>
<tr>
<td>-193 AAATCCCGGAAGAAATTTT</td>
<td>-175 -478 GTTGCTATTTT</td>
</tr>
<tr>
<td>-35 TCTGGGGTTTTCTTTCAAG</td>
<td>-53 -419 TAACAAATAAAGTG</td>
</tr>
</tbody>
</table>

* complementary strand.
DISCUSSION

Protein L15 is encoded by a duplicated gene and, like many other yeast ribosomal protein genes (30,36,37), both gene copies are located in different chromosomes. Contrary to most ribosomal protein genes, however, the presence of introns in the L15 gene coding region has not been found. Although the existence of intervening sequences upstream from the initiation codon, reported in two yeast rp-genes (38,39), cannot be totally excluded, this is highly improbable since the consensus splice junctions and the highly conserved TACTAAC box required for splicing in yeast (40,41) are not present in any of the two gene copies.

The sequence homology in the coding region of both genes is elevated, although a relatively high proportion of third base differences in the reading frame are found which, however, do not affect the amino acid sequence. The homology in the non-coding regions is very low.

The promoter region of both protein L15 genes fits partially into the standard model for rp genes (42). As is usual in yeast promoters (43), canonical TATA boxes are not present in the region −40 to −100 nucleotides from the initiator codon, although related boxes can be found. In this sense it is interesting to find that the closely related sequences, TACATT and TACTATTA are present in that region in both cases. The TC-rich stretch upstream from the TATA box, usually present in yeast rp-genes is found only in the L15b gene.

A number of nucleotide sequences showing a high similarity to the consensus upstream activating sequences (UASrpg) specific for the general transacting protein factor TUF/RAP1/GRF1 (44,45) can be found in both genes (Table 1). These UAS are present in most rp-genes and other highly expressed genes (45).

The UAS showing a higher homology quality are located at approximately 500 bp from the initiation codon, one of them in the complementary strand. A second set can be detected closer (200 to 300 bp) to the ATG. In the L15b copy, UASrpg-like sequences are found even in the coding region.

A few yeast rp-genes, namely those encoding L3 (46) and S33 (47), have a different type of UAS which form part of the binding site for factors SUF, TAF and the ARS/silencer factor GFI (48).

Interestingly, nucleotide boxes matching the consensus sequence for this UAS with reasonable homology can also be found in the protein L15b gene copy (Table 1). To know whether they have any functional meaning requires additional experimental data.

Table 2. Amino acid sequence comparison of protein L15 and other yeast and bacterial ribosomal proteins

<table>
<thead>
<tr>
<th>proteins</th>
<th>BESTFIT ratio</th>
<th>identity</th>
<th>gaps</th>
<th>FASTA Score</th>
<th>init1</th>
<th>opt</th>
</tr>
</thead>
<tbody>
<tr>
<td>yL15-ecL11</td>
<td>0.450</td>
<td>22.9</td>
<td>3</td>
<td>21</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>yL15-ecL10</td>
<td>0.317</td>
<td>14.4</td>
<td>4</td>
<td>18</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>yL15-ecL7</td>
<td>0.446</td>
<td>22.2</td>
<td>3</td>
<td>19</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>yL15-ylA5</td>
<td>0.396</td>
<td>25.9</td>
<td>4</td>
<td>21</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>yL14-ylL44</td>
<td>0.404</td>
<td>18.1</td>
<td>3</td>
<td>21</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>yL15-ylL44'</td>
<td>0.370</td>
<td>15.0</td>
<td>2</td>
<td>20</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>yL15-ylL45(A)</td>
<td>0.441</td>
<td>25.0</td>
<td>2</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>yL15-ylL45(A)</td>
<td>0.449</td>
<td>24.6</td>
<td>3</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>yL15-ylL45(A)</td>
<td>0.376</td>
<td>11.9</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>yL15-ylPO</td>
<td>0.383</td>
<td>21.2</td>
<td>6</td>
<td>21</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>averageb</td>
<td>0.340</td>
<td>17.2</td>
<td>4.7</td>
<td>17</td>
<td>26</td>
<td></td>
</tr>
</tbody>
</table>

a; the comparison has been performed excluding the 30 last amino acids in the carboxyl end of the acidic proteins.

b; the data represent the average values from comparing protein L15 with 5 bacterial and 5 eukaryotic non-ribosomal proteins of similar size.

Fig. 5. Dot-plot of data obtained comparing protein L15 to bacterial proteins L11 (left) and L10 (right) The GCG Compare program, using values of 20 and 9 for the window and the stringency respectively, was used.
Other structural features of both genes are summarized in Table 1. It is interesting to note the presence of sequences with potential capacity for forming secondary structures, in the same region where the UASrpg are found.

The amino acid sequence encoded in both copies of the protein L15 gene is identical and coincides with a partial amino terminal sequence previously reported for this protein (34). It was noted, however, that the published sequence starts at the alanine residue in position 16 suggesting that the protein might be processed after translation. Alternatively, the published sequence might correspond to a degradation product. In fact, the second possibility is probably correct since our attempts to sequence purified protein L15 using automated Edman degradation have failed indicating that the first amino acid residue is probably blocked. Moreover, the only two tyrosine residues in the sequence, located in positions 12 and 14, should be missing in the processed protein; nevertheless, the previously reported amino acid composition of purified protein L15 indicated the presence of this amino acid in the expected proportion (11).

The immunological and functional homologies between protein L15 and bacterial protein L11 previously reported (11,15) are supported by a substantial amino acid sequence similarity. There are some stretches near the carboxyl terminal where the similarity is specially high and they might be involved in the interaction of both proteins with the same region of the large rRNA (15).

As commented above, protein L15 also functionally resembles bacterial protein L10 in forming a complex with the acidic ribosomal proteins (14). However, the overall structural similarity between L15 and L10 is not significant, suggesting that the interaction site with the acidic ribosomal proteins has evolved divergently. This is not surprising considering the differences in amino acid sequences between the eukaryotic and bacterial acidic proteins (20). Alternatively, the interaction of L15 with the yeast acidic proteins might be a specific feature of the eukaryotic systems not related to the L10 function.

The significant amino acid sequence similarity between the acidic proteins L44 and L45—but not L44'—and protein L15 is very interesting. This observation is in agreement with results indicating that protein L15 is able to mask the monoclonal specific epitopes in proteins L44 and L45 but not in protein L44' (B. Ortiz and J.P.G. Ballesta, unpublished results). These data suggest that in the ribosome, protein L15 might interact indistinctly with proteins L44 and L45, which are functionally equivalent (49), but not with L44'. In fact, the functional and structural peculiarities of protein L44', that contrary to L44 and L45 is unable to form dimers in solution (12) and is absent from the cytoplasmic pool of acidic proteins (M.D. Villela, M. Remacha and J.P.G. Ballesta, unpublished results), suggest a different role for this protein, more in line with that of standard ribosomal proteins. This is, in addition, supported by the structure of the protein L44' gene, that unlike L44 and L45, has an intron and standard UASrpg (50,51), resembling the typical ribosomal protein gene model (42).

ACKNOWLEDGEMENTS

We thank M.C. Fernández Moyano for expert technical assistance. This work has been supported by institutional grant to Centro de Biología Molecular from Fundación Ramón Areces and by grant PB0450 to J.P.G.B. from Dirección General de Investigación Científica y Técnica (Spain).

REFERENCES