Volume 17 Number 22 1989 Nucleic Acids Research

Structure of a ribosomal protein gene in *Mucor racemosus*

Luis Sosa+, William A.Fonzi* and Paul S.Sypherd

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, CA 92717, USA

Received July 5, 1989; Revised and Accepted October 17, 1989 EMBL accession no. X15582

ABSTRACT

As an extension of our analysis of the translational apparatus of *Mucor racemosus* we have isolated a gene encoding a ribosomal protein of *Mucor*. Based on a method developed for *S. cerevisiae*, we identified by hybrid selection and *in vitro* translation a λ-Charon 4A clone containing the genomic copy of a *Mucor* ribosomal protein. The gene consisted of two exons of 57 and 387 nucleotides. The two exons were separated by an 131 nucleotide intron. The processed transcript was 714 nucleotides in length and contained a 25 nucleotide untranscribed leader and an 114 nucleotide untranscribed 3'-end. The protein predicted from the nucleotide sequence contained 148 amino acids and exhibited 61% identity with the S19 ribosomal protein of *Xenopus laevis*. The promoter region of the gene contained sequences highly homologous to the *RPG* and *Homol1* promoter elements found in *S. cerevisiae*. Southern blot analysis indicated that the *Mucor* genome contains three copies of this gene.

INTRODUCTION

The fungus *Mucor racemosus* serves as a microbial model for cellular morphogenesis, by possessing several distinct morphogenetic pathways: sporulation, spore germination, and yeast-hyphal transition. The yeast-hyphal transition, or dimorphic character, has received considerable attention in the past, as attempts have been made to identify the biochemical and physiological correlates of the process. Studies in this laboratory have shown that the yeast to hyphae transition, which occurs by shifting the organisms from CO_2 to air, is accompanied by marked changes in the translation rate (1), and the specific activity of the elongation factor EF-1α (2). Further, we have cloned the gene for EF-1α and shown that the protein is coded by a family of three genes (3,4) all of which are transcribed in all phases of growth.

In continuing the analysis of the translation system, we have focused on the regulation of ribosome synthesis, and particularly, the coordinate regulation of EF-1α and ribosomal proteins. In this paper we report the cloning of a ribosomal protein gene, and show its extensive homology with protein S19 from *Xenopus*, as well as the conservation of presumed regulatory sequences between the ribosomal protein gene and the genes for EF-1α, both in *Mucor* and *Saccharomyces cerevisiae*. Finally, the data also demonstrate that the S19 ribosomal protein analogue of *Mucor* is, like EF-1α, coded by a 3-gene family.

MATERIALS AND METHODS

Organism and Culture Conditions

Sporangiospores of *M. racemosus* (*Mucor lusitanicus*) ATCC 1216B were prepared and stored as described previously (3,5). Sporangiospores were germinated in YPG medium

(1% Bacto-peptone, 0.3% yeast extract, 2% glucose, pH 4.5) by shaking on a rotary shaker water bath at 28°C. The culture was sparged with >2 volumes of sterile air per volume of culture per minute. Germlings were collected by filtration when the germ tubes reached a length five times the diameter of the swollen spores. Yeast cells were produced from sporangiospores germinated in YPG and grown until mid-log phase by shaking at 28°C in an atmosphere of 100% CO_2 (0.2 volumes of sterile CO_2 per volume of culture per min). Cultures of yeast cells were induced to undergo morphogenesis by transferring them to air and continuing to shake the culture for 3 h. At this time 80% of the cells had germ tubes. Cells were collected by filtration and frozen immediately in liquid N_2. The nucleic acids were extracted as described below.

Bacterial Strains and Cloning Vectors

E. coli strain, JM84, was the host strain for plasmids pUC9, pUC18 and pUC19; *E. coli* JM105 and JM109 were the hosts for the phages M13mp10, M13mp18 and M13mp19. Plasmid DNA was isolated as described in (6). Phage M13mp derivatives were grown, and the DNA purified, as recommended by BRL, Inc., Gaithesburg, MD.

Construction and Screening of an M. racemosus Gene Bank in Bacteriophage λ

A *Mucor* genomic bank was constructed in the vector λ Charon 4A using DNA fragments from a partial digestion with *EcoRI*, as described by Maniatis *et al.* (6). Recombinant phages were screened by making pools of three independent clones and binding DNA isolated from the pools to nitrocellulose filters. These were used for hybrid-selection of mRNAs (7,8) which were eluted from the filters (6), and translated *in vitro* using a rabbit reticulocyte lysate (BRL) in the presence of [^{35}S]methionine (NEN). The translation reactions were precipitated and washed with methanol and chloroform according to the procedure by Wessel and Flugge (9). The translation products were analyzed by electrophoresis as described in the following section.

Isolation and Electrophoresis of Ribosomal Proteins

Ribosomal proteins were isolated from the mycelia obtained from one liter of YPG medium using the procedure described by Larsen and Sypherd (10). Protein extracted by acetic acid-Mg^{++} was dialyzed overnight in 3 l of 1 M acetic acid at 4°C with two changes of buffer. The dialyzate was lyophilized and the dried ribosomal proteins were stored at −70°C. The protein pellet was suspended in 8 M urea prior to electrophoresis. The bidimensional electrophoresis system utilized to separate the ribosomal proteins was performed according to Otake and Osawa (11). For the first dimension, the tube gels were run 4.5 h at 200 volts. The second dimension was run on a 15% polyacrylamide-SDS gel as described by Hames (12).

Southern and Northern Analysis of Nucleic Acids

DNA was isolated as described by Cihlar and Sypherd (13) and total RNA of *Mucor* was extracted as described by Linz *et al.* (3). 1 to 5 μg of *Mucor* DNA digested with different endonucleases and resolved by electrophoresis through agarose gels was transferred to nitrocellulose filters by the procedure of Southern (14). RNA for Northern analyses was transferred to nitrocellulose filters by the procedure of Thomas (15). The blots were hybridized with agarose gel-fractionated restriction fragments labelled using the random primer technique of Feinberg and Vogelstein (16) or with probes prepared by primer extension labelling of fragments cloned in M13 phage. Nitrocellulose filter hybridization conditions were essentially those of Maniatis *et al.* (6). Filters were exposed to Kodak XAR-film at −70°C with a Cronex Lightning-Plus intensifier screen for 4−24 h.

DNA Sequencing

The nucleotide sequence was determined by the dideoxy chain termination procedure of Sanger and Coulson (17). The sequencing reactions were performed using [1α-^{32}P]dATP for labeling and utilizing a Sequenase™ kit from United States Biochemicals according to the instructions of the manufacturer.

Mapping of 5' Start of Transcription

The primer extension mapping technique described by Hu and Davidson (18) was used. The modifications were: 10μmg of total cellular RNA was hybridized with 100−50 ng of ssDNA of an M13 clone of the 0.5 kb *HindIII* fragment (see Fig. 2), containing the transcriptional start site. The DNA-RNA hybrid was annealed with the universal sequencing primer and the DNA synthesis reaction was run 15 minutes in the presence of [α-^{32}P]dATP or [α-^{32}P]dGTP, followed by a 15 min incubation with cold dNTP. The radioactive DNA was resolved in an 8% polyacrylamide-8 M urea gel and visualized by autoradiography.

Mapping of 3' End of Transcription

Labelled probe for 3'-end mapping was generated by the technique of Calzone *et al.* (19). The complementary strand of an M13mp19 clone of the 0.69 kb *PvuII-HindIII* fragment (Fig. 2) was synthesized by primer extension in the presence of [^{32}P]dATP or [^{32}P]dGTP. The duplex DNA was digested with *HindIII* , denatured with NaOH, and fractionated by agarose gel electrophoresis. The labelled band was localized by autoradiography and then extracted by electrophoretic transfer to DE81 paper. After elution from the DE81, the DNA was concentrated by precipitation with ethanol in the presence of 0.3M sodium acetate.

Hybridization and S1 nuclease treatment were conducted as described by Ausubel *et al.* (20). The protected fragments were resolved on a 8M urea-6% polyacrylamide sequencing gel and visualized by autoradiography .

Computer Analysis of RPG Sequence

The RPG sequence was analyzed with the DNA Inspector II program on a Macintosh Plus computer. A protein and gene comparison search was made using the University of Wisconsin Genetics Computer Group (UWGSG) program.

RESULTS

Identification of a λ Clone Containing the Ribosomal Protein Gene

To isolate ribosomal protein genes of *Mucor* we employed the scheme developed by Woolford *et al.* for the isolation of ribosomal protein genes of *S. cerevisiae* (8). Their protocol was based on the fact that a large fraction of ribosomal proteins are of relatively low molecular weight and are encoded by small mRNAs found predominantly on small polysomes. A library of *Mucor* genomic DNA was prepared in λ Charon 4A and independent isolates were pooled in groups of three. DNA prepared from the pools was bound to nitrocellulose filters and hybridized to poly A^+ RNA recovered from the low molecular weight portion of a sucrose gradient (21). The hybridized mRNA was eluted and translated *in vitro* with a rabbit reticulocyte lysate. The *in vitro* translation products were identified by autoradiography after two-dimensional electrophoresis in polyacrylamide gels. One pool of three clones hybridized to a message which, when translated in a reticulocyte lysate, yielded a small basic protein that appeared to comigrated with a ribosomal protein. The clone which hybridized with this particular mRNA was identified and shown to contain a 13kb insert.

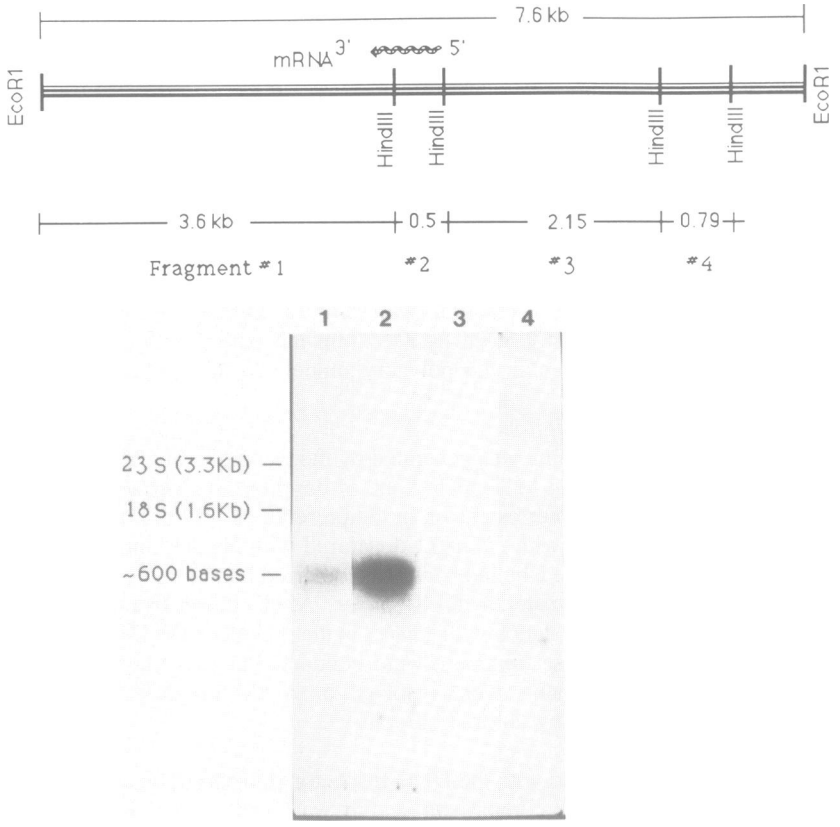

Figure 1. Localization of the transcribed region. The diagram in upper portion of the figure depicts the location of restriction endonuclease sites within the 7.6 kb *EcoRI* fragment obtained from the λ-Charon 4A clone. Each of the indicated fragments was isolated and labelled by the random primer method. The labelled probes were hybridized with a nitrocellulose blot of total *Mucor* RNA fractionated on a 1% agarose-formaldehyde gel. Lanes 1,2,3, and 4 were probed with the correspondingly numbered restriction fragment. Values to the right of the Northern blot indicate the position of the ribosomal RNA bands and the estimated size of the hybridizing band.

Localization of the Transcribed Region

The insert of 13kb was cut with *EcoRI* into two DNA fragments of 7.6 and 5.4 kilobases. Each fragment was labelled *in vitro* and used to probe a Northern blot of *Mucor* RNA. Autoradiography showed that the 5.6 kb fragment hybridized with a message of ~300 nucleotides, while the fragment of 7.6 kb hybridized with an abundant message of approximately 600 nucleotides. The latter message was sufficiently large to code for a protein of the size identified on the polyacrylamide gels. The 7.6 kb *EcoRI* fragment was digested with *HindIII*, and each of the four resulting fragments was used to probe a Northern blot (Fig.1). Both the 0.5 kb *HindIII-HindIII* fragment and the adjacent 3.6 kb *HindIII-EcoRI* fragment hybridized with the 600 nucleotide RNA, the former hybidizing more strongly. This experiment localized the ribosomal protein gene to the 0.5 kb *HindIII* fragment and the adjacent 3.6 kb *HindIII-EcoRI* fragment. Subsequent restriction site

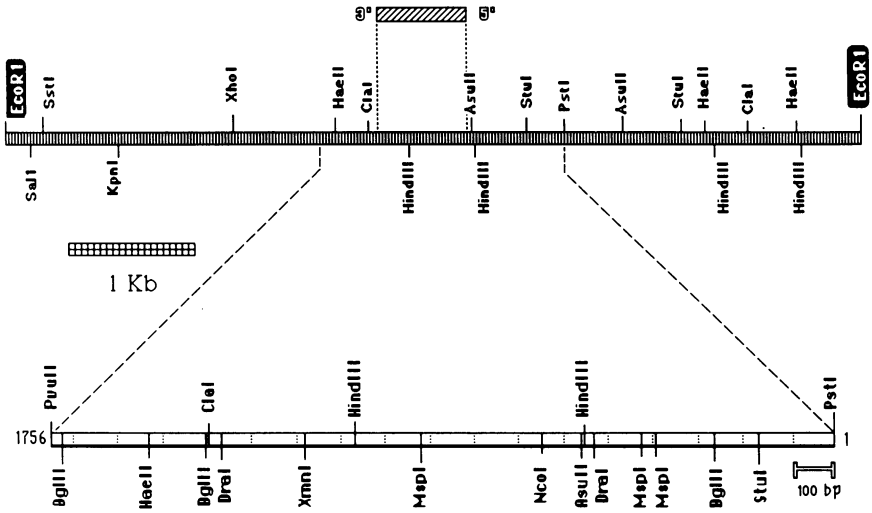

Figure 2. Restriction map. The diagram depicts the location of endonuclease restriction sites within the 7.6 kb *EcoRI* fragment containing the *Mucor* RPG gene.

mapping identified a 1.7 kb *PvuII-PstI* fragment encompassing the transcribed region. This fragment was isolated and subcloned for DNA sequence determination. Figure 2 displays a restriction site map and the various subclones used to sequence the region. The nucleotide sequence is shown in Fig.3.

Transcript Mapping

The direction of transcription was determined by cloning the 0.5 kb *HindIII* fragment containing the coding region (Fig. 2)in both orientations in M13mp10. A complementary DNA for each clone was synthesized in the presence of [^{32}P]dGTP and hybridized to Norther blots. The results showed that transcription proceeded from the *PstI* site toward the *PvuII* site, as depicted in Fig. 2 (data not shown).

The transcription initiation site was determined using the T4 primer-extension technique (18) which is based on the ability of an RNA-DNA hybrid to block procession of T4 polymerase. When total RNA was hybridized with ssDNA from an M13mp10 clone of the 0.5 kb *HindIII* fragment, two primer-extension products were obtained (Fig.4A and 4B). The shorter of the two products 60 nucleotides in length and presumably represented the site of transcription initiation. This result indicated that transcription initiates at a 'C' residue located at nucleotide −26 relative to the translational start site as indicated in Fig.3.

The second primer-extension product was 274 nucleotides in length. The end point of this product correlated with a presumptive splice site within the sequence and likely represents the 3′ boundary of the proposed intron (Fig.3).

The 3′ end of the transcribed region was located by S1 nuclease mapping. A labelled, complementary probe was prepared by primer-extension using ssDNA from an mp19 clone of the 0.69 kb *PvuII-HindIII* fragment (Fig. 2). The probe was hybridized with total *Mucor* RNA and treated with S1 nuclease, then analyzed in a 8M urea, 6% polyacrylamide gel. The autoradiograph, presented in Fig.5, showed the protection of two fragments, 258 and 260 nucleotides in length. This indicated that transcription of the gene ends approximately 714 nucleotides 3′ of the transcriptional start site. The processed form of the transcript.

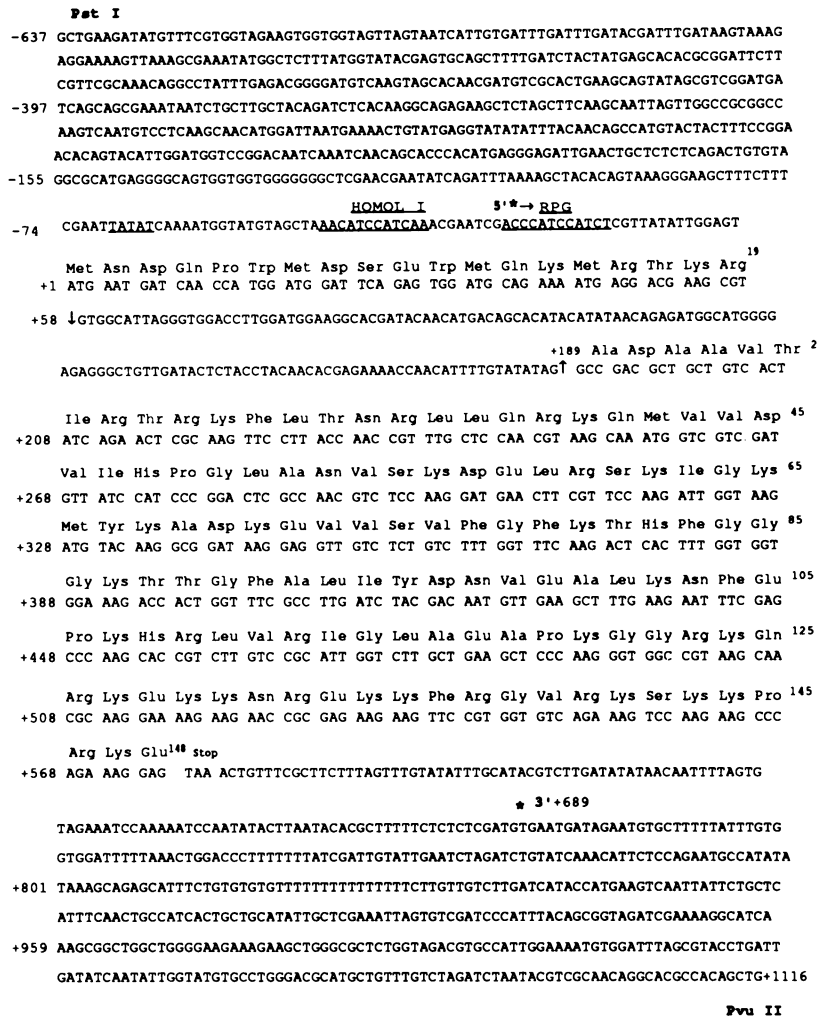

Figure 3. Nucleotide sequence of *Mucor* ribosomal protein gene. The figure shows the complete nucleotide sequence of the *PstI-PvuII* restriction fragment. The numbers to the left of the sequence indicate the nucleotide position relative to the translational start site of the gene. The asterisk above nucleotides −25 and +689 indicate the sites of initiation and termination of transcription respectively. The arrows at positions +58 and +189 indicate the 5′ and 3′ splice boundaries respectively. The *Homol1* and *RPG* consensus sequences and a presumptive 'TATA' box are underlined.

with removal of the 131 base intron (Fig.3), would be a total of 583 nucleotides in length, which agrees well with the estimate of 600 nucleotides determined from agarose gel electrophoresis (Fig.1).

Analysis of the Nucleotide Sequence

Examination of the sequence within the transcribed region revealed no continuous open reading frames starting from any of the several ATG codons within the 5′ end of the

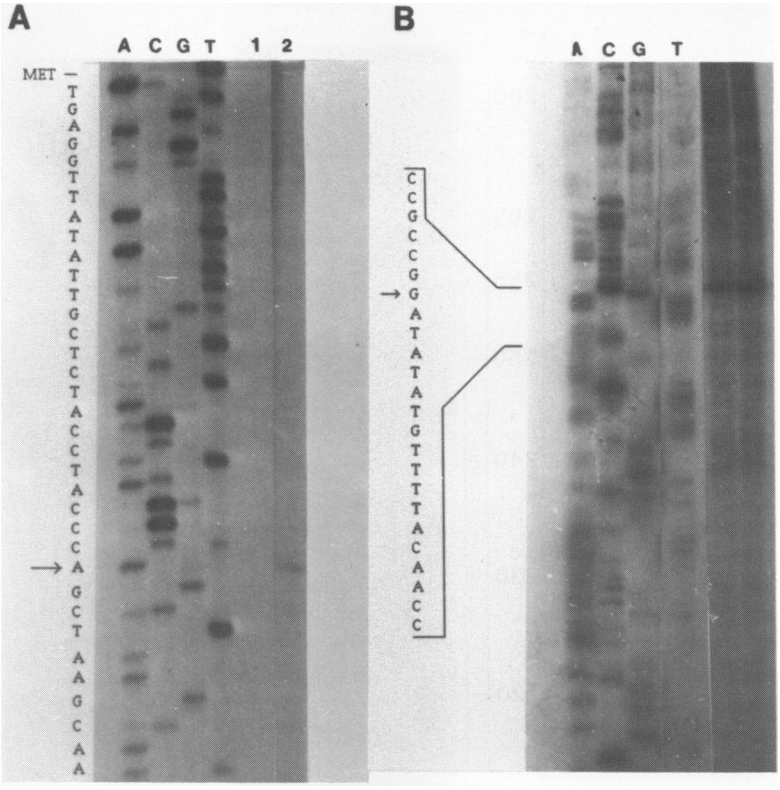

Figure 4. Mapping of the 5′-terminus of the transcript. The transcriptional initiation site was mapped by the procedure of Hu and Davidson (18). An M13 clone of the 0.5 kb *HindIII* fragment (see Fig. 2) was hybridized with total RNA and then with universal sequencing primer. The primer was extended with T4 polymerase up to the position of hybridization with the mRNA. The resulting product was fractionated on an 8% acrylamide sequencing gel. A dideoxy sequencing ladder of the same fragment was run in parallel. Panel A shows the lower part of the gel and panel B shows the upper portion of the gel. The nucleotide sequence is indicated on to the left of the gel. Lanes 1 and 2 are without and with RNA respectively.

transcribed region. The first ATG codon, embedded within the sequence AGT**AGTA**, exhibited the highest degree of homology with eukaryotic translational start sites (22,23) and constitutes the presumptive site of translational initiation.

The presence of an intron was suggested by the lack of a continuous open reading frame in the sequence. There was no clear 5′ consensus splice junction within or following the first open reading frame, however, the process of elimination indicated only one feasible splice site that would align the first and second exons in the proper reading frame. This splice junction was located 19 codons after the first ATG codon within the sequence CGT↓**GTGG**. The transcript mapping data presented in the previous section suggested a 3′ boundary of the presumptive intron at nucleotide +189. The sequence at the 3′ splice site, TTGTATATAG/**GC**, is clearly related to the consensus splice sequence Py6NAG (24). Following this splice site the open reading frame extends for 129 codons. When the predicted protein product of the *Mucor* gene was compared to the NBRF data bank

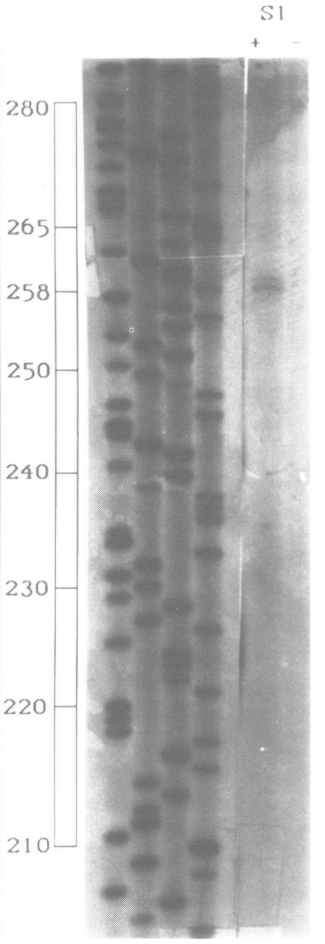

Figure 5. Mapping of the 3'-terminus of the transcript. The 3'-end of the transcript was mapped using S1 nuclease as described in 'Materials and Methods'. The products were fractionated on a 6% acrylamide sequencing gel. A dideoxy sequencing ladder was run in parallel for size comparison. Numbers to the left of the gel indicate the length of the indicated fragments in nucleotides.

the sequence was found to be highly homologous to the amino acid sequence of the ribosomal protein S19 of *Xenopus laevis* (25). A comparison of the sequences is shown in Fig.6. The promoter region of the gene contained a putative TATA box with the sequence AA**TTATAT**CA 38 bp prior to the transcription start site (Figure 3). Unexpectedly , precise matches to the so called UAS**rpg** boxes; *Homol1* (26) and *RPG* (27), reported for genes of the translation apparatus of *S. cerevisiae*, were present at −21 and −1 nucleotides from the transcription initiation site (Table 1 and Figure 2). The TCCA motif which we previously found in the TEF-1 and TEF-3 genes of *Mucor* (4) forms part of the RPG box. There was no polyadenylation consensus site in the 3' flanking sequences, but a related AATATA sequence is located 26 bases before the end of the transcript.

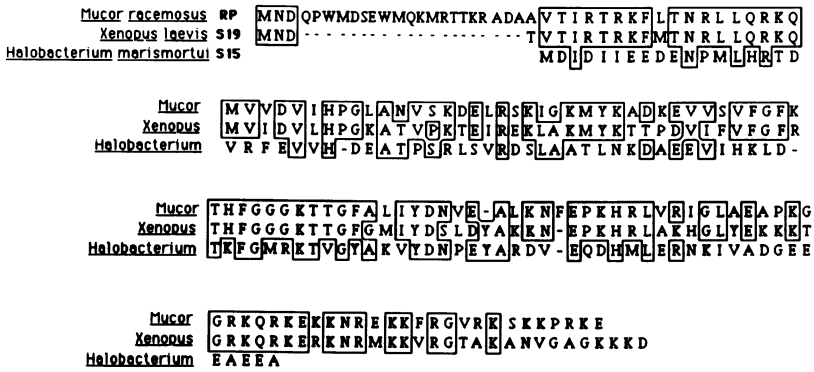

Figure 6. Comparison of amino acid sequences. The protein predicted from the nucleotide sequence of the *Mucor* ribosomal protein gene was aligned with the sequences for *Xenopus* S19 protein and *Halobacterium* S15 protein. Identical residues are enclosed.

Southern Analysis

Southern analysis was performed to determine if the ribosomal protein gene was a single or multi-copy gene. As seen in Fig.7A, digestion with any of several restriction endonucleases resulted in the appearance of three hybridizable bands. A second hybridization using the 5′ flanking region of the cloned gene revealed only a single band (Fig.7B). The results indicated that the genomic location of the cloned gene is distinct from that of the other two hybridizing DNA fragments and suggested that the other hybridizing genomic

Table 1. Comparison of HOMOL1 and RPG sequences.

Gene	HOMOL I	RPG Box	Organism	Ref.
L25	NOT FOUND	ACCCGGACATCT (-329)	Candida utilis	28
K37	AAAATCTATGCA (-224)	GATCATCAATTA (-311)	S. pombe	29
S10-1	TGCATCCGTACA (-325)	CTCCATACATCT (-305)	S. carlsbergensis	27
L16-2	AACATCCAACCA (-385)	ACCCATACCGTT (-359)	Saccharomyces cerevisiae	27
RP51-2	AACATCCATACA (-263)	ACCCATACATTT (-299)	-	27
TEF-1	AACACCCAAGCA (-322)	NOT FOUND	-	30
TEF-2	CACATTTATACA (-418)	ACCCACACATTT (-425)	-	30
RP39A	TTAAACATCCGTACAACGAG	AACCCATCATTACTTT (-221)	-	31
RP Mucor	CTAAACATCCATCAAACGAATCG	ACCCATCCATCTCGTT (-10)	Mucor racemosus	This work
TEF-1	NOT FOUND	CATCCATCCAT (-33,-45)	-	4
TEF-3	CATCCATCAAAC (-28)	CATCCATCCAT (-76,-40)	-	4
CONSENSUS	AACATC CG/TA T A/G CA (-200/-800)	ACCCATACAT TT/CA (-200/-500)		32

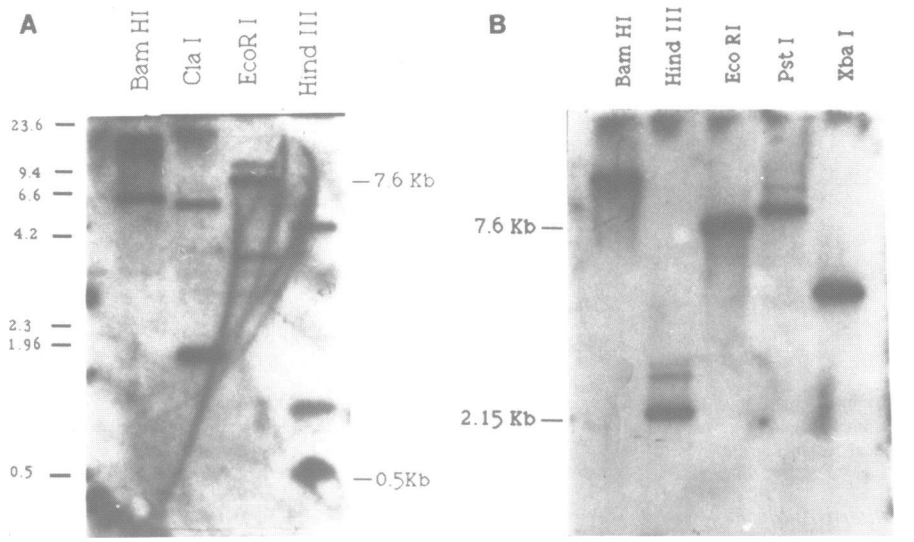

Figure 7. Southern hybridization analysis of genomic DNA. Ten μg of *Mucor* DNA was digested with the indicated enzymes and fractionated on an agarose gel. The gels were blotted to nitrocellulose and hybridized with either the 0.5 kb *HindIII* fragment containing the coding region of the gene (Panel A) or with the *PvuII-HindIII* fragment containing sequences 5′ of the coding region (Panel B). Numbers to the left of Panel A indicate the size and positions of *HindIII* restricted λ DNA . Numbers to the right of Panel A indicate the size and position of expected *EcoRI* and *HindIII* restriction fragments based on the results obtained from the initial isolate. Similarly, numbers to the left of Panel B indicate the size and position of expected restriction fragments.

fragments represent at least one, and possibly two, additional copies of the gene or pseudogene counterparts.

DISCUSSION

Our isolation of a ribosomal protein gene from *M. racemosus* represents the first such isolation from any Zygomycete and was based upon a method used originally for *Saccharomyces* (8). The isolated gene was transcribed into a 714 nucleotide primary transcript containing an 131 nucleotide intron. The presence of an intron near the 5′ end of the gene is comparable to the findings with most ribosomal protein genes from *Saccharomyces cerevisiae* (33,34,35), and with the occurrence of an intron near the 5′ end in two of the three EF-1α genes from *Mucor* (4). The intron in the ribosomal protein gene divides the protein sequence into two domains, the first 19 amino acids being neutral, and the remainder of the protein being decidedly basic, 15 arginine residues and 24 lysine residues out of 129.

The *Mucor* protein was found to be highly homologous with the protein S19 from *Xenopus laevis* (Fig.6). The two proteins shared 61% identity, and an even higher degree of homology if conservative amino acid changes are permitted. As seen in Fig. 6, the *Mucor* protein possesses an extended amino terminal peptide compared to the *Xenopus* protein. As shown previously (36) for the S19 protein of *Xenopus*, there is also a surprising degree of homology (ca. 30%) between the S19 protein and the S15 protein of the archaebacterium *Halobacterium marismortui*. The strong conservation of ribosomal protein sequences among the phyla has been noted by others (37,38).

The codon usage for the *Mucor* ribosomal protein has a bias similar to that reported for the EF-1α genes of *Mucor* (data not shown), showing once again the preference of codon usage in major, constitutively expressed genes as has been emphasized by others (39). This codon usage is, however, species specific, since the codon usage in the proteins of the *Mucor* translation systems are different from that of *S. cerevisiae* (40).

An interesting similarity was found between the promoter region of the *Mucor* ribosomal protein gene and the promoters of *S. cerevisiae* genes encoding proteins of the translational apparatus. The two consensus sequences found in yeast, the *Homol1* (26) and the *RPG* boxes (27), are reported to be involved in regulating gene activity (41,42). A factor called *TUF* has been implicated as a transcriptional activator acting at these sites (30,43). The *Homol1* and *RPG* sites are clearly evident in the *Mucor* sequences and exhibit a high degree of homology with the *Saccharomyces* sequences (Table 1). They do, however, differ considerably in location. The *Mucor Homol1* box is located only 10 nucleotides 5′ of the transcriptional start site, while the *RPG* box overlaps the transcriptional start site. In *Saccharomyces* , these sequences are located much further up-stream. Experiments with yeast have shown that a separation of at least 100 nucleotides from the start site is necessary for the activation of a gene under *RPG* control (32). Functional analysis of the *Mucor* promoter will be necessary to determine if these homologous sequences are of significance. Another element involved in the activation of ribosomal protein genes in *Saccharomyces* is is the T-rich region (44) usually found between the RPG and the transcriptional start. Analogous sequences were not found in this *Mucor* gene.

A comparison of the promoters of the *Mucor* 'S19 analogue protein' and the EF-1α gene family reveals similarities based primarily on the *RPG* and *Homol1* sequences (Table 1). The promoter of the TEF-1 gene of *Mucor* contains duplicate, over-lapping copies of the sequence CATCCATCCAT , homologous with the *RPG* box, 30 nucleotides 5′ of the translational start site. The TEF-3 gene has two copies of this same sequence located 40 and 76 nucleotides 5′ of the translational start site. In addition, the TEF-3 gene contains the sequence CATCCATCAAAC, which is identical to the *Homol1* consensus of the *Mucor* S19 gene. These similarities suggest a common mechanism may control transcription of these components of the translational apparatus. Certainly the presence of these sequences warrants further study .

Mucor was previously shown to contain three copies of the gene encoding EF-1α(3). Similarly, the ribosomal protein gene reported here was shown to hybridize with three genomic DNA fragments (Fig.7). One of the hybridization bands was unique to the cloned gene. The other two bands may represent a single additional copy of the gene with a restriction site polymorphism in the chromosomal homologs of this diploid organism or may represent two addition copies of the gene. We do not yet know if either of these other copies is expressed or if any of the hybridization bands represent pseudogenes. Isolation of these additional genes will permit these questions to be addressed.

ACKNOWLEDGEMENTS

We wish to acknowledge the excellent assistance during various portions of this work, of Carol Katayama, Terry Woodworth and Lucy Lira. This work was supported by a grant (GM23999) from the National Institute of General Medical Science. Luis Sosa was supported in part by a fellowship from the Mexico/Chicano Program of the University of California.

*To whom correspondence should be addressed

+Present address: Universidad de Guanajuato, Instituto de Investigacion en Biologia Experimental, Apartado Postal No. 187, Guanajuato, Morales, Mexico

REFERENCES

1. Orlowski, M. and Sypherd, P.S. (1978) *Biochemistry* **17**, 569–575.
2. Fonzi, W.A., Katayama, T., Leathers, T. and Sypherd, P.S. (1985) *Mol. Cell. Biol.* **5**, 1100–1103.
3. Linz, J.E., Katayama, C. and Sypherd, P.S. (1986) *Mol. Cell. Biol.* **6**, 593–600.
4. Sundstrom, P., Lira, L.M., Choi, D., Linz, J.E. and Sypherd, P.S. (1987) *Nucleic Acids Research* **15**, 9,997–10,006.
5. Larsen, A.L. and Sypherd, P.S. (1980) *J. Bacteriol.* **141**, 20–25.
6. Maniatis, T., Fritch, E.F. and Sambrook E.F. (1982) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
7. Parnes, J.R., Velan, B., Felsenfeld, A., Ramanathan, L., Ferrini, U., Appella, E. and Seidman, J.G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2253–2257.
8. Woolford, J.L, Hereford, L.M. and Rosbash, M. (1979) *Cell* **18**, 1247–1259.
9. Wessel, D. and Flugge, U.I. (1984) *Anal. Biochem.* **138**, 141–143.
10. Larsen, A.L. and Sypherd, P.S. (1979) *Mol. Gen. Genet.* **175**, 99–109.
11. Otake, E. and Osawa S. (1981) *Mol. Gen Genet.* **181**, 176–182.
12. Hames, B.D. (1981) In Hames, B.D. and Rickwood, D. (eds), *Gel Electrophoresis of Proteins. A Practical Approach*, IRL Press, Oxford. pp. 1–91.
13. Cihlar, R.L. and Sypherd, P.S. (1980) *Nucleic Acids Research* **8**, 793–804.
14. Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517.
15. Thomas, P. (1982) *Methods Enzymol.* **100**, 255–260.
16. Feinberg, A.P. and Vogelstein, B. (1984) *Analytical Biochemistry* **137**, 266–267.
17. Sanger, F. and Courlson, A.R. (1975) *J. Mol. Biol.* 94, 441–448.
18. Hu, M.C. and Davidson, N. (1986) *Gene* **42**, 21–29.
19. Calzone, F.J., Britten, R.J., & Davidson, E.H. (1987) Meth. Enzymol.**152**:611–632.
20. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K (eds.) (1989) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience, New York.
21. Wagner, M. and Perry, R.P. (1985) *Mol. Cell. Biol.* **5**, 3560–3576.
22. Kozak, M. (1986) Cell **44,** 283–292.
23. Hamilton, R., Watanabe, C.K., and deBoer, H.A..1987. Nucl. Acids Res. **15**, 3581–3593.
24. Lewin, B. (1983) Genes, John Wiley & Sons, New York.
25. Amaldi, F., Beccari, E., Bozzoni, I., Luo, Z. and Pireandrei-Amaldi, P. (1982) *Gene* **17**, 311–316.
26. Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, W.H., Planta, R.J., Schultz, L., Friesen, J.D., Fried, H. and Rosbash, M. (1984) *Nucleic Acids Research* **12**, 8295–8312.
27. Leer, R.J., van Raamsdonk-Duin, M.C.C., Mager, W.H. and Planta, R.J. (1985) *Current Genetics* **9**, 273–277.
28. Wardt, L.P., Mager, W.H., Beek, J.G., Wassenaar, G.M., and Planta, R.J. 1987. Current Genet. **12**, 193–198.
29. Nistch, R., Gross, T., Gatermann, K., Swida, U. and Kauter, N. 1987. Nucl. Acids Res. **15**, 1477–1491.
30. Huet, J., Cottrelle, P., Cool, M., Vignaes, M., Thiele, D., Marck, C., Buhler, J., Sentenac, A. and Fromageot, P. (1985) *EMBO J.* **4**, 3539–3547.
31. Rotenberg, M.O. and Woolford, J.L. (1986) *Mol. Cell. Biol.* **6**, 674–687.
32. Woudt, L.P., Mager, W.H., Nieuwint, T.M., Wassenaar, G.M., van der Kuyl, A.C., Hoekman, M.F.M., Brockhoff, P.G.M. and Planta, R. J. (1987) *Nucleic Acids Research* **15**, 6037–6048.
33. Bollen, G., Molenaar, C.M.T., Cohen, L.H., van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1982) *Gene* **18**, 29–38.
34. Nieuwint, R.T.M., Molenaar, C.M.T., van Bommel, J.H., van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1985) *Current Genetics* **10**, 1–5.
35. Teem, J.L.and Rosbash, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4403–4407.
36. Kimura, J., Arndt, E. and Kimura, M. (1987) *FEBS Letters* **224**, 65–70.
37. Chen, I.T., Dixit, A., Rhoads, D. and Roufa, D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6907–6911.
38. Tanaka, T., Kuwano, Y., Ishikawa, K. and Ogata, K. (1985) *J. Biol. Chem.* **260**, 6329–6333.
39. Bennetzen, J. and Hall, B. (1982) *Science* **230**, 32–36.

40. Nagata, S., Nagashima, K., Tsunetsugu-Yokota, Y, Fujimura, K., Miyazaki, M. and Kaziro, Y. (1983) *EMBO J.* 3, 1825–1830.
41. Pearson, N.J., Fried, H.M. and Warner, J.R. (1982) *Cell* **29**, 347–355.
42. Woudt, L.P., Smit, A.B., Mager, W.H. and Planta, R.J. (1986) *EMBO J.* **5**, 1037–1040.
43. Huet, J. and Sentenac, A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3648–3652.
44. Rosbash, M., Harris, P.K.W., Woolford, J.L. and Teem, J.L. (1981) *Cell* **24**, 679–686.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.