Volume 307, number 3, 318–323 FEBS 11367 August 1992

Structural determination of *Saccharomyces cerevisiae rig* gene and identification of its product as ribosomal protein S21

Shin Takasawa, Akira Tohgo, Michiaki Unno, Hideto Yonekura and Hiroshi Okamoto

Department of Biochemistry, Tohoku University School of Medicine, Sendai 980, Miyagi, Japan

Received 1 June 1992

rig was originally isolated from a rat insulinoma-derived cDNA library. The 145 amino acid sequence of the *rig* protein is invariant in mammalian cDNAs. In this paper, we have isolated the cDNA and genomic clones for yeast (*Saccharomyces cerevisiae*) *rig*, determined their nucleotide sequences, and identified the gene product. The gene and the mRNA encode a basic protein of 142 amino acids which has 61.3% amino acid identity with mammalian *rig* protein. On two-dimensional gel electrophoresis, the in vitro transcription/translation product of yeast *rig* cDNA co-migrated with yeast ribosomal protein S21. These results led to the conclusion that yeast *rig* encodes ribosomal protein S21 and to the determination of the previously unknown primary structure of yeast S21 protein. Unlike most ribosomal protein genes of *S. cerevisiae*, the gene exists as a single copy in a haploid set of the yeast genome and has no intron, locating at chromosome VII or XV.

rig; Ribosomal protein S21; Chromosomal localization; Yeast

1. INTRODUCTION

rig (rat insulinoma gene) was first isolated from a cDNA library of chemically induced rat insulinoma [1]. The gene has been found to be expressed in all cell types of vertebrates examined so far [1–6] and to code for a basic protein of 145 amino acids, whose sequence has been highly conserved during vertebrate evolution [1–6]. *rig* has been found to be present as a member of a multicopy gene family in mammalian genomes and as a single copy gene in chicken and frog genome [3,5]. Human genomic *rig* is divided into four exons separated by three introns and located in a tight cluster of CpG-islands [5]. Recently, the product of *rig* has been demonstrated to be ribosomal protein S15 in mammals [7].

In this study, we have isolated yeast cDNA for *rig* and the gene, determined the complete nucleotide sequences, and identified the product as yeast ribosomal protein S21. The yeast *rig* (ribosomal protein S21) gene exists as a single copy in a haploid set of the genome and has no intron.

2. EXPERIMENTAL

2.1. *cDNA and genomic cloning of yeast homologue to* rig/*S15*

Saccharomyces cerevisiae DNA was prepared essentially as described by Struhl et al. [8]. cDNA and genomic libraries were purchased from Clontech Laboratories (Palo Alto, CA). A specific probe for the yeast homologue of *rig*/S15 was prepared using the polymerase chain reaction [9] together with yeast genomic DNA and two synthetic oligodeoxyribonucleotides (CTCGACCAGCTGCTGGACATGTCCTA and GAATTCACCCAAGTAGTGACCGATCAT) based on the protein sequence of mammalian S15/rig (amino acid residues 22–29 and 111–119) [1–7]. The resulting 300 bp fragment was purified on an agarose gel, and labeled with [α-^{32}P]dCTP by the random-priming technique [3]. The full-length cDNA and the entire gene were isolated using plaque hybridization [5] and the cDNA and the genomic inserts were subcloned in the pBS plasmid vector (Stratagene, La Jolla, CA) for DNA sequencing. The nucleotide sequences were determined by the dideoxy chain-termination method as described [3,5].

Abbreviations: UAS_{rpg}, upstream activation sequence of ribosomal protein gene(s).

Correspondence address: H. Okamoto, Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 980, Miyagi, Japan. Fax: (81) (22) 272-7273.

2.2. *In vitro expression of cloned cDNA*

The cDNA fragment containing the entire coding region of yeast homologue to *rig* was subcloned into pBS. The cDNA was transcribed in the sense orientation and translated in a rabbit reticulocyte lysate containing [^{35}S]methionine [2]. The ^{35}S-labeled product was extracted with 67% acetic acid, precipitated with 200 μg yeast ribosomal proteins in acetone and analyzed by 2D polyacrylamide gel electrophoresis [7,10] followed by fluorography.

2.3. *Southern blot analysis*

Southern blot analysis was performed essentially as described [3], using a 379 bp yeast *rig* cDNA fragment (−23–356) as a probe.

2.4. *Chromosomal location of yeast* rig/*S21*

Samples of yeast chromosome (Beckman, Palo Alto, CA) were electrophoresed on a GeneLine pulse-field gel electrophoresis system (Beckman). Hybridization was carried out as described above.

2.5. *Determination of transcriptional initiation sites*

S. cerevisiae RNA was prepared by the method of Feinberg and McLaughlin [11]. Primer extension was performed as described [5] using a synthetic 31-mer complimentary to the nucleotide residues 1–31 of the yeast homologue of *rig*/S15 mRNA.

3. RESULTS

By screening a yeast cDNA library with a specific gene fragment which was prepared by the polymerase

 Published by Elsevier Science Publishers B.V.

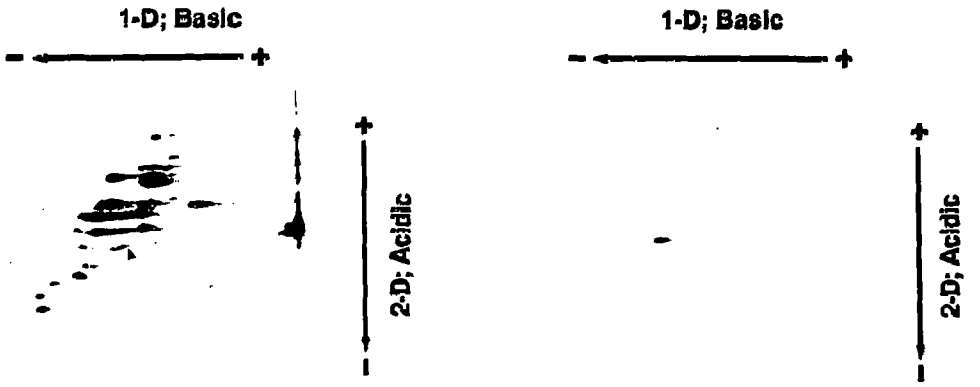

Fig. 1. Electrophoretic co-migration of in vitro translated yeast *rig* protein with yeast ribosomal protein S21. The in vitro transcription/translation product of yeast *rig* cDNA was applied to the 2D gel with 200 μg of purified yeast ribosomal proteins. Electrophoresis was from right to left in the first dimension and top to bottom in the second. (Left panel) Coomassie brilliant blue (CBB) stain. The arrow indicates the spot of S21 protein. (Right panel) Fluorography.

chain reaction using synthetic oligonucleotide based on the amino acid sequence of mammalian *rig* protein [1–7] (see section 2), a recombinant phage, λgtYRIG-5, was obtained. The insert was subcloned into plasmid pBS and the cDNA sequence was determined. The cDNA stretched for 539 nucleotides plus poly(A) and had one large open reading frame coding for a 142 amino acid protein. (The nucleotide sequence data will appear in the DDBJ, EMBL and GeneBank Nucleotide Sequence Database.)

The mRNA transcribed from the YRIG-5 cDNA in vitro was translated in rabbit reticulocyte lysate and the ^{35}S-labeled translation product was analyzed by electrophoresis in a 2D gel system [10]. The translation product co-migrated with yeast ribosomal protein S21 on the 2D gels (Fig. 1).

The cDNA insert (−23–356 nucleotide residues) was used as a probe in Southern blot (Fig. 2). In all twelve kinds of DNA digests, only one band hybridized with the probe, indicating that a single copy of the gene for the yeast homologue of *rig* (yeast ribosomal protein S21) exists in a haploid set of the genome. Using the

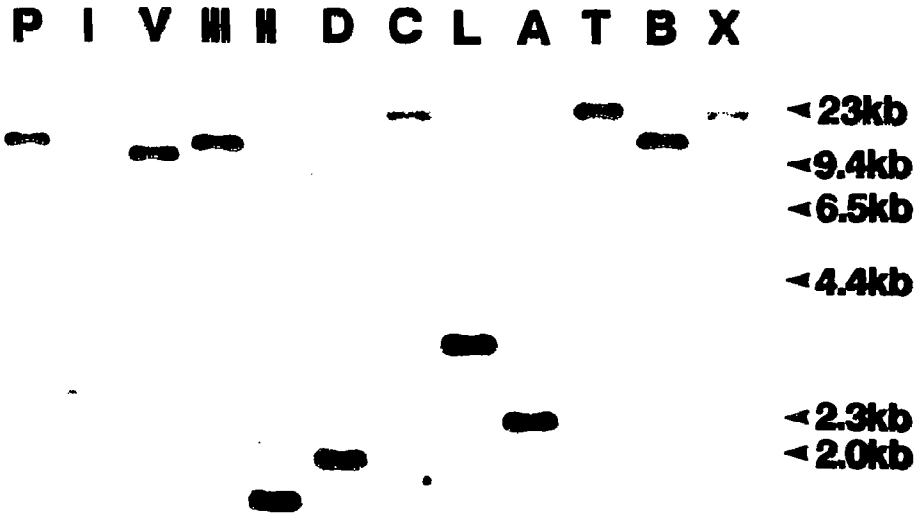

Fig. 2. Southern blotting of yeast *rig*/S21. Yeast DNA (3 μg) was digested with *Pst*I (lane P), *Eco*RI (lane I), *Eco*RV (lane V), *Hin*dIII (lane III), *Hin*cII (lane II), *Dra*I (lane D), *Sac*I (lane C), *Sal*I (lane L), *Sca*I (lane A), *Stu*I (lane T), *Xba*I (lane B), and *Xho*I (lane X), and hybridized with the cDNA fragment for yeast *rig*/S21. The numbers on the right indicate the length in kilobase pairs of *Hin*dIII-digested lambda phage DNA as standards.

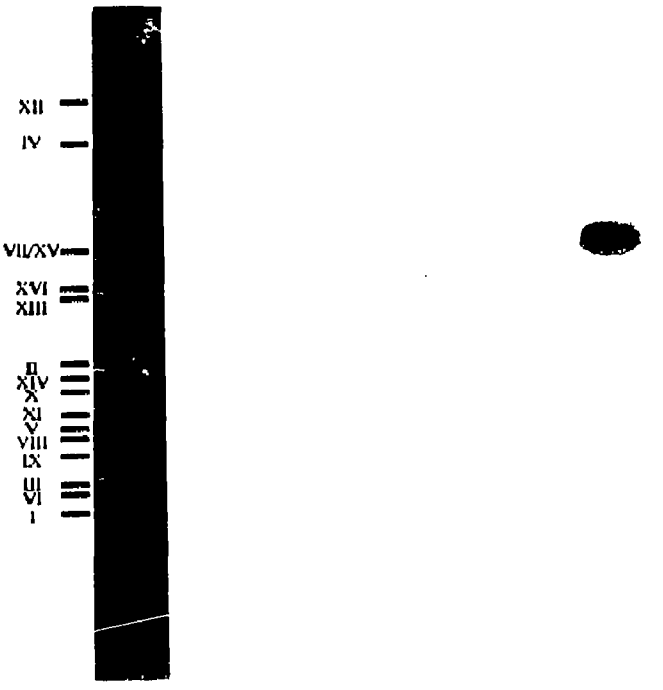

Fig. 3. Chromosomal location of yeast *rig*/S21. Yeast chromosomes were applied onto a pulse-field gel, and transferred to membranes. (Left panel) Ethidium bromide-stained gel. (Right panel) Autoradiogram of Southern blot. The chromosome numbers are indicated on the left.

same cDNA probe, a Southern blot of yeast chromosomes resolved by pulse-field gel electrophoresis was performed. Only a single band was found to hybridize with the probe (Fig. 3). Since the hybridized band is made of two chromosomes, chromosomes VII and XV [12], the yeast homologue of *rig* gene was thought to be located on one of these chromosomes.

Using the cDNA fragment (−23–356 nucleotide residues) as a probe, we screened a yeast genomic DNA library and obtained two positive clones, λgtYRIG-1 and λgtYRIG-2 (Fig. 4A). We sequenced the DNA contained in the two overlapping inserts of YRIG-1 and YRIG-2. The 1,688 nucleotide sequence was determined for both strands as shown in Fig. 4. The nucleotide sequence (residues −23–516) found in the genomic DNA (Fig. 4B) coincided with the cDNA sequence. There was a single open reading frame coding for 142 amino acids, indicating that the gene is uninterrupted by an intron. This view was confirmed by the lack of the canonical sequences [13] characteristic of the 5′ splice site (GTATGT) or of the lariat site (TACTAAC), either within or upstream of the open reading frame. In about 960 nucleotides of the 5′-flanking region there were two sequences homologous to the consensus sequences of upstream activation sequences of ribosomal protein genes (UAS_{rpg}) [14,15] (Fig. 4B). Of the two sequences, the one starting at position −434 fitted well with the HOMOL1 consensus sequence, and the other starting at −413 fitted the RPG-box [14,15]. This gene also contained pyrimidine-rich segments [16] at positions −290 to −276, −162 to −131, and −115 to −71. In the 3′-flanking region there was a polyadenylation signal starting at position 496, and the polyadenylation site was at nucleotide 516.

Initiation of transcription was determined by primer extension. A synthetic oligodeoxyribonucleotide complementary to the nucleotide residues 1–31 of the cDNA or the genomic DNA (Fig. 4B) was labeled at the 5′ end and used for hybridization with yeast RNA. After reverse transcription, three bands were found (Fig. 5); they are shown in Fig. 4B by arrowheads (residues −52, −42 and −38). Therefore, 5′-leader regions of approximately 40–50 nucleotides were estimated in the mRNA.

4. DISCUSSION

We had reported the nucleotide sequences of the yeast homologue of *rig* mRNA and of its gene and identified the gene product as yeast ribosomal protein S21. Yeast ribosomal protein S21 had been defined as the constituent of a spot on the 2D gel [10], but its primary structure had yet to be determined. In the present study, the primary structure of S21 protein was determined from the nucleotide sequences of yeast *rig* cDNA and of its gene, and yeast ribosomal protein S21 was thought to be the counterpart of mammalian *rig* protein, which has recently been identified as mammalian ribosomal protein S15 [7].

Unlike other yeast ribosomal protein genes [17,18], the gene was a single copy gene and lacked introns. Although the existence of introns upstream of the initi-

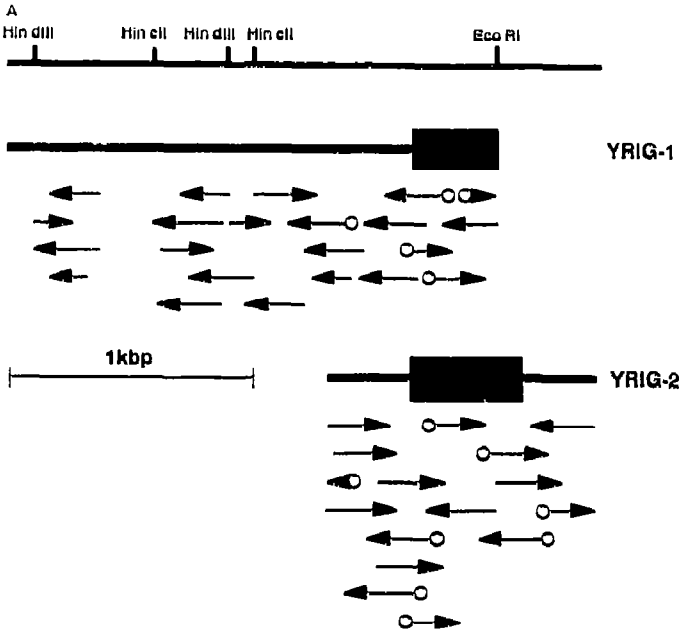

Fig. 4A. For legend see opposite.

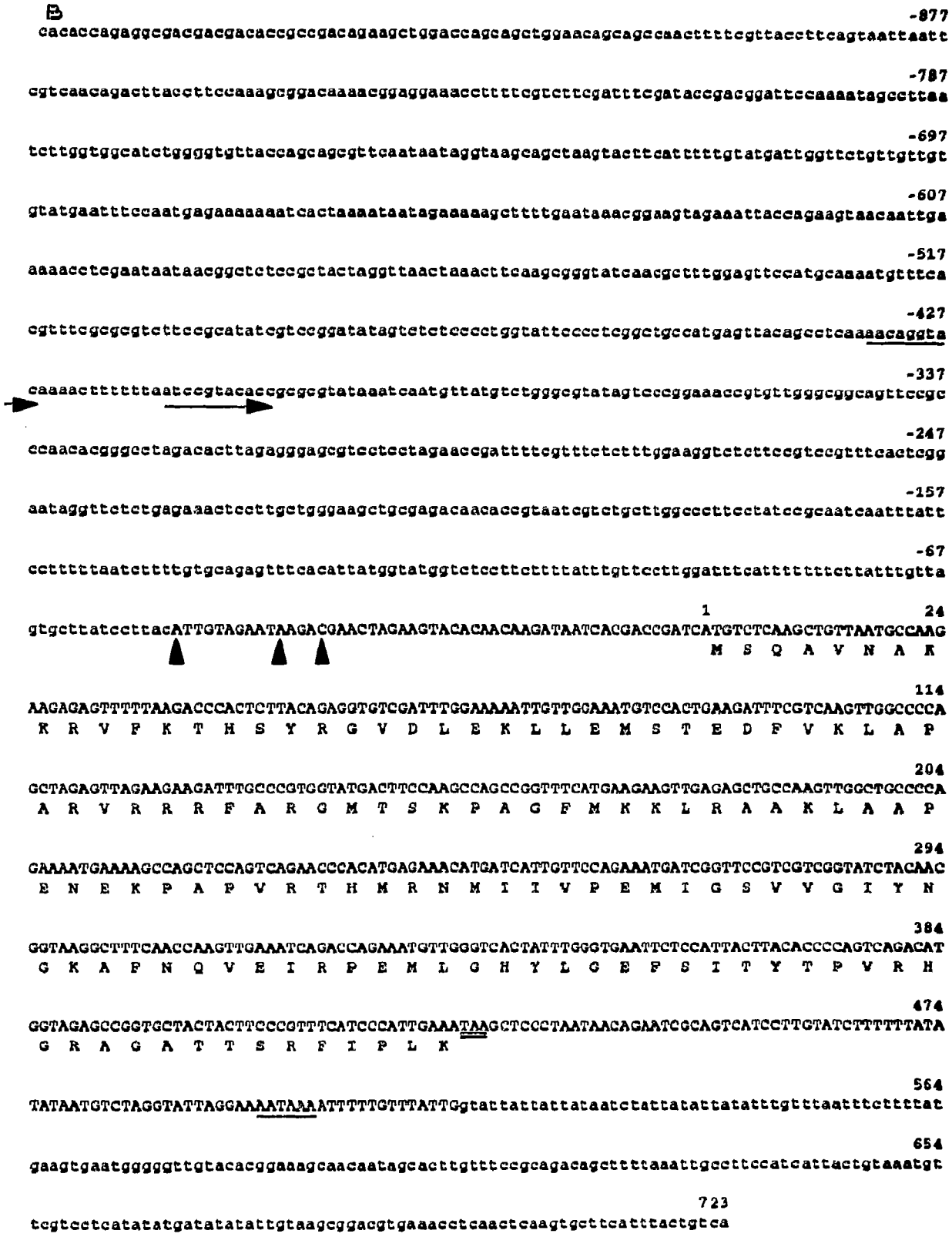

Fig. 4. Restriction map and sequencing strategy (A) and nucleotide sequence (B) for yeast *rig*/ribosomal protein S21 gene. In A, restriction sites in the yeast *rig*/S21 gene are shown at the top, the exon is indicated as a box, arrows indicate the direction and extent of sequence determination, and open circles indicate synthetic primers. In B, capital letters indicate the exon and lower case letters are used for 5′- and 3′-flanking sequences. Nucleotides are numbered starting at the first adenine in the translation initiation codon, and the termination codon is double underlined at position 427–429. The putative UAS_{rpg} boxes, starting at nucleotide −434 and −413, respectively, are marked, with their orientations indicated by arrows. The polyadenylation signal in the 3′-non-coding region is also underlined. Arrowheads indicate the transcription start sites (positions −52, −42, and −38). The nucleotide sequence data will appear in the DDBJ, EMBL and GeneBank Nucleotide Sequence Database.

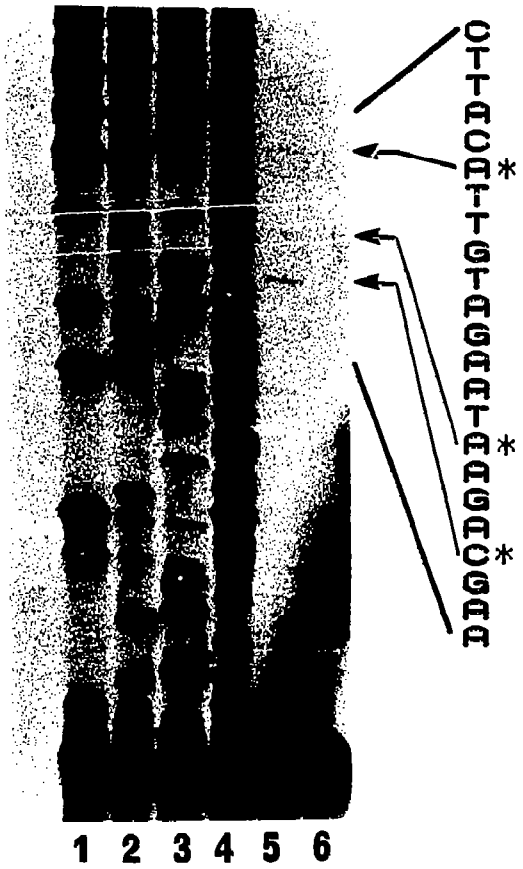

Fig. 5. Mapping of the transcription initiation site by primer extension analysis. (Lanes 1–4) A, C, G, and T of a sequencing ladder, respectively. (Lane 5) Primer extended products. Three DNA bands differing in length were detected. The nucleotide sequence of the sense strand and nucleotide numbers are indicated at the right with the initiation sites marked by asterisks. (Lane 6) Primer extension control experiment without RNA.

ation codon, as reported in two yeast ribosomal protein genes [19,20], cannot be totally excluded, this is highly improbable since the consensus splice junctions and the highly conserved TACTAAC box required for splicing in yeast [21,22] are not present in the gene copy. In the 5′-flanking region there were two nucleotide sequences homologous to the regulatory elements, named HOMOL1 and RPG-box, found in the 5′-flanking regions of the genes for most basic ribosomal proteins [14,15]. The ribosomal protein S21 gene is unique in having standard UAS_{rpg} regions (HOMOL1 and RPG-box) in the unspliced single copy gene; it has so far been reported that yeast ribosomal protein genes without introns either lack UAS_{rpg} boxes (like proteins L3 and S33) or have apparently less efficient ones (like proteins S24, L45, and L44). The T-rich region detected in the ribosomal protein genes closer to the transcription initiation site than to UAS_{rpg} [16], was present in the gene. The transcription initiation sites were located by primer extension (Fig. 5) at −52, −42, and −38 nucleotides upstream of the translation initiation site. The 3′-flanking site was also identified from the sequence of the cDNA clone and was preceded at an adequate distance by the polyadenylation signal, AATAAA.

As deduced from the nucleotide sequence of the gene and cDNA, yeast ribosomal protein S21 contains 142 amino acid residues, however, the NH_2-terminal methionine of the yeast S21 protein could be removed after translation since the residue next to the initial methionyl in S21 is serlyl, which has been reported to favor NH_2-terminal processing [23]. Thus, it is possible that the number of residues in the mature protein is 141 and that the protein is monoacetylated, as in the case of its mammalian counterpart [7]. The sequence of amino acids in S21 was searched for internal repeats. A possible duplication, 5 identities in two sequences of 7 consecutive residues, was found at positions 85–91 (IVPEMIG) and 107–113 (IRPEMLG).

The amino acid sequence in yeast ribosomal protein S21 was compared with the amino acid sequence of more than 700 other ribosomal proteins contained in a library that Dr. Tanaka of the University of the Ryukyus has compiled. Comparative analysis revealed the similarity of yeast S21/*rig* protein to mammalian S15/*rig* protein and to the proteins of the prokaryotic S19 family. We have already indicated that the prokaryotic S19 family are the homologues of mammalian S15/*rig* protein [7]. We further analyzed some consensus sequences in the ribosomal proteins. The amino acid sequence (residues 113–123, GHYLGEFSITY) in yeast S21 is completely conserved in the eukaryote proteins. There are eight identical residues within the region for *Halobacterium* S19's and 6–8 identities for eubacterial and chloroplastic S19's (see [7]). Pohl and Wittmann-Liebold [24], and Brockmöller and Kamp [25] showed that the histidine residue of this region in *Escherichia coli* and in *Bacillus stearothermophilus* S19 are in contact with S13 proteins of these species. In addition, the eukaryote proteins have three short conserved sequences (amino acid residues 8–10, 17–22 and 42–44) in their amino-terminal region. The sequence of amino acids residues 8–10 (KKR) and 42–44 (RRR) are composed of basic residues. Basic amino acid clusters located in the amino-terminal region of yeast ribosomal protein, L3 and L29, were reported to function in nuclear localization and/or assembly of the ribosomal proteins [26,27]. It is possible that the basic amino acid clusters of yeast S21/*rig* and mammalian S15/*rig* protein carry out such functions.

Acknowledgements: We are grateful to Dr. Tatsuo Tanaka, Department of Biochemistry, School of Medicine, University of the Ryukyus, for sending data from the ribosomal protein data base; to Dr. Motoo Kitagawa, Institute for Molecular and Cellular Biology, for useful suggestions; and to Hideo Kumagai and Reiko Torigoe for their skilful technical assistance. This work has been supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Takasawa, S., Yamamoto, H., Terazono, K. and Okamoto, H. (1986) Diabetes 35, 1178–1180.

[2] Inoue, C., Shiga, K., Takasawa, S., Kitagawa, M., Yamamoto, H. and Okamoto, H. (1987) Proc. Natl. Acad. Sci. USA 84, 6659–6662.

[3] Sugawara, A., Nata, K., Inoue, C., Takasawa, S., Yamamoto, H. and Okamoto, H. (1990) Biochem. Biophys. Res. Commun. 166, 1501–1507.

[4] Takasawa, S., Inoue, C., Shiga, K. and Kitagawa, M. (1990) in: Molecular Biology of the Islets of Langerhans (Okamoto, H. ed.) pp. 287–299, Cambridge University Press, Cambridge.

[5] Shiga, K., Yamamoto, H. and Okamoto, H. (1990) Proc. Natl. Acad. Sci. USA 87, 3594–3598.

[6] Taylor, K.D. and Piko', L. (1991) Mol. Reprod. Dev. 28, 319–324.

[7] Kitagawa, M., Takasawa, S., Kikuchi, N., Itoh, T., Teraoka, H., Yamamoto, H. and Okamoto, H. (1991) FEBS Lett. 283, 210–214.

[8] Struhl, K., Stinchcomb, D.T., Scherer, S. and Davis, R.W. (1979) Proc. Natl. Acad. Sci. USA 76, 1035–1039.

[9] Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 239, 487–491.

[10] Bollen, G.H.P.M., Mager, W.H. and Planta, R.J. (1981) Mol. Biol. Rep. 8, 37–44.

[11] Feinberg, B. and McLaughlin, C.S. (1988) in: Yeast: A Practical Approach (Campbell, I. and Duffus, J.H. eds.) pp. 147–161, IRL press, Oxford.

[12] Mortimer, R.K. and Schlid, D. (1985) Microbiol. Rev. 49, 181–212.

[13] Woolford, J.L. (1989) Yeast 5, 439–457.

[14] 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 Rosbach, M. (1984) Nucleic Acids Res. 12, 8295–8312.

[15] Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1985) Curr. Genet. 9, 273–277.

[16] Rotenberg, M.O. and Woolford, J.L. (1986) Mol. Cell Biol. 6, 674–687.

[17] Planta, R.J., Mager, W.H., Lee, R.J., Woudt, L.P., Raue', H.A. and El-Baradi, T.T.A.L. (1986) in: Structure, Function and Genetics of Ribosomes (Hardesty, B. and Kramer, G. eds.) pp. 699–718, Springer Verlag, New York.

[18] Warner, J.R. (1989) Microbiol. Rev. 53, 256–271.

[19] Mitr, G. and Warner, J.R. (1984) J. Biol. Chem. 259, 9218–9224.

[20] Nieuwint, R.T.M., Molenaar, C.M.T., van Bommell, J.H., van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1985) Curr. Genet. 10, 1–5.

[21] Pikielny, C.W., Teem, J.L. and Rosbach, M. (1983) Cell 34, 395–403.

[22] Langford, C.J., Klinz, F.J., Donath, C. and Gallawitz, D. (1984) Cell 36, 645–653.

[23] Flinta, C., Persson, B., Jörnvall, H. and von Heijne, G. (1986) Eur. J. Biochem. 154, 193–196.

[24] Pohl, T. and Wittmann-Liebold, B. (1988) J. Biol. Chem. 263, 4293–4301.

[25] Brockmöller, J. and Kamp, R.M. (1988) Biochemistry 27, 3372–3381.

[26] Moreland, R.B., Nam, H.G., Hereford, L.M. and Fried, H.M. (1985) Proc. Natl. Acad. Sci. USA 82, 6561–6565.

[27] Underwood, M.R. and Fried, H.M. (1990) EMBO J. 9, 91–99.