ELSEVIER

FEBS Letters 340 (1994) 133–138

FEBS 13715

FEBS LETTERS

# Structural determination and characterization of a 40 kDa protein isolated from rat 40 S ribosomal subunit

Akira Tohgo**, Shin Takasawa, Hiroshi Munakata, Hideto Yonekura, Norio Hayashi, Hiroshi Okamoto*

*Department of Biochemistry, Tohoku University School of Medicine, 2-1 Seiryo-machi, Aoba-ku, Sendai 980, Miyagi, Japan*

Received 16 December 1993; revised version received 20 January 1994

**Abstract**

We have purified a 40 kDa protein from the rat 40 S ribosomal subunit and determined its primary structure by amino acid and cDNA sequencing. The amino acid sequence of the 40 kDa protein shared 29–37% homology with prokaryotic ribosomal protein S2 of eubacteria and chloroplasts, indicating that the protein is a eukaryotic counterpart to prokaryotic S2. Moreover, the amino acid sequence shared 99% identity with those deduced from cDNAs for 68 kDa laminin binding proteins of human, murine and bovine origins. The cDNAs are capable of encoding polypeptides with predicted molecular mass of 33,000 which lacked typical signal sequences, N-linked glycosylation sites and putative transmembrane domains. These results indicate that the cDNAs for 68 kDa laminin binding proteins actually code for the 40 kDa ribosomal protein.

*Key words:* Ribosomal protein S2; 68 kDa laminin binding protein; Amino acid sequence; cDNA cloning; Rat liver

## 1. Introduction

Ribosomes play a central role in the biosynthesis of proteins and the structure of ribosomes and their constituents are highly conserved during evolution [1,2]. In this study, we have purified a 40 kDa protein from the rat 40 S ribosomal subunit and determined its primary structure by amino acid and cDNA sequencing. Comparative analysis revealed that the 40 kDa protein is a eukaryotic counterpart to prokaryotic ribosomal protein S2. The identity of this 40 kDa ribosomal protein with a protein deduced from the cDNA for a cell surface 68 kDa laminin binding protein is discussed.

## 2. Experimental

### 2.1. Protein purification and amino acid sequence determination

Ribosomes, their subunits and ribosomal proteins were prepared as previously described [2]. The 40 S ribosomal proteins were separated by SDS-PAGE and stained with Coomassie brilliant blue (CBB). A 40 kDa protein was excised from the stained gel and the protein was electroeluted as previously described [3]. Lysyl endopeptidase fragments of the 40 kDa protein were separated by reverse-phase HPLC and underwent automated Edman degradation as previously described [2]. Amino acid compositions were determined with an Applied Biosystems Model 130A amino acid analyzer [4]. SDS-PAGE and immunoblotting were performed [2], and the protein was determined by the method of Smith et al. [5].

### 2.2. cDNA cloning

The cDNA was synthesized from rat regenerating liver total RNA by reversetranscription using Superscript (Gibco-BRL, Gaithersburg, MD). A specific probe for the 40 kDa protein was prepared using a polymerase chain reaction (PCR) [6,7] together with the cDNA and two synthetic oligodeoxyribonucleotides (GAATTCAATCTTGACTTCCAGATGGA and CACGTGAAGGTTCCAGGAGTGAA) based on the amino acid sequences of the lysyl endopeptidase fragments of the purified 40 kDa protein (Fig. 3, arrows). The resulting 250-base pair (bp) fragment was purified on an agarose gel and labeled using the Megaprime DNA labeling system (Amersham, UK). The full-length cDNA was isolated from a rat liver cDNA library [8] using plaque hybridization, and the cDNA inserts were subcloned into the pBS vector (Stratagene, La Jolla, CA). Nucleotide sequences were determined as described previously [9].

## 3. Results and discussion

We found a 40 kDa protein associated with the 40 S ribosomal subunit in the course of elucidating the product of *rig* (rat insulinoma gene), i.e. ribosomal protein S15 [2,10]. As shown in Fig. 1, most of the immunoreactivity against a small subunit ribosomal protein, S15, was recovered around fraction 18 (lower panel). The 40 kDa protein coincided with the immunoreactivity and RNA peak in the 40 S ribosomal subunit region (fraction 18) on sucrose gradient centrifugation (middle panel), indicating that the 40 kDa protein is a component of the 40 S ribosomal subunit.

To determine its primary structure, we purified the 40 kDa protein from the SDS-gels (Fig. 2) (see section 2) and analyzed the purified protein on an amino acid sequencer. No phenylthiohydantoin-amino acid was liber-

*Corresponding author. Fax: (81) (22) 272 8101.

**Submitted in partial fulfillment of the degree of Doctor of Medical Science at Tohoku University.


*SSDI* 0014-5793(94)00099-H

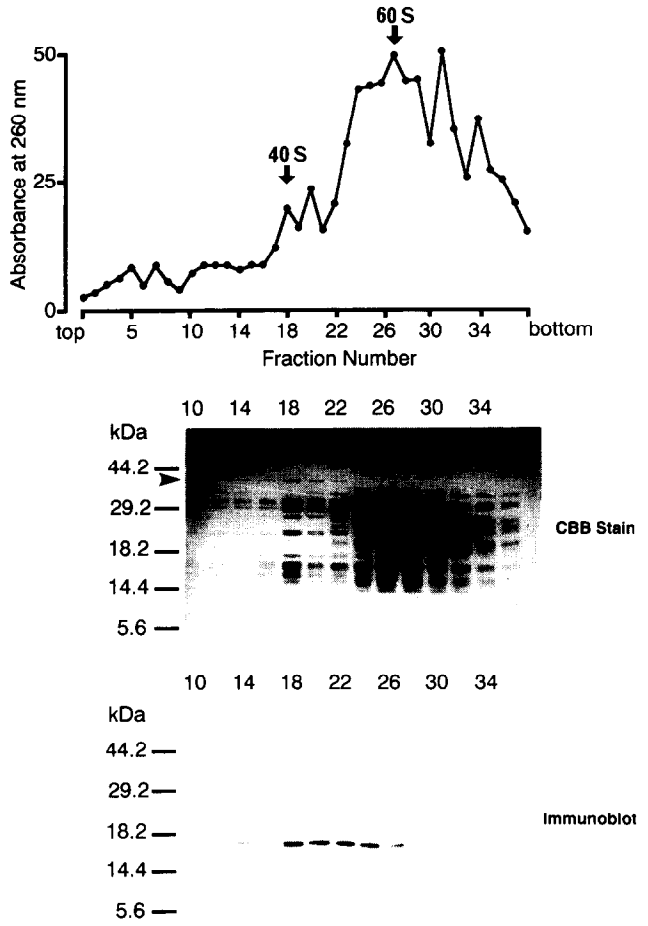


Fig. 1. Sedimentation of the 40 kDa protein in sucrose gradient. Ribosomes were prepared as described previously [2]. Ribosomes incubated with puromycin were subjected to 15–30% linear sucrose density gradient centrifugation containing 0.85 M KCl. Upper panel, absorption at 260 nm; middle panel, CBB staining of even-numbered fractions after separation by SDS-PAGE; lower panel, immunodetection of rat ribosomal protein S15 [2]. For the middle and lower panels, 10 μl of each gradient fraction was used. Positions of the size markers are presented on the left. The 40 kDa protein is indicated by arrowhead in the middle panel.

ated at the first cycle of Edman degradation, indicating that there is a blockage on the amino terminus. The purified protein was digested with lysyl endopeptidase and the digested peptides were separated on reverse phase HPLC. Seven peptides were isolated and sequenced on a gas phase sequencer. This analysis yielded a total of 109 amino acid residues from the sequences of the seven peptides (Fig. 3, boxes). Based on the amino acid sequences, we prepared a cDNA fragment specific for the 40 kDa protein by PCR (see section 2). We screened a rat liver cDNA library ($6 \times 10^5$ plaques) with the cDNA fragment and eight positive clones were obtained. The eight clones largely overlapped each other and had complete nucleotide identity in the overlapping regions. The longest clone contained a cDNA insert of approximately 1.0 kbp. A single band of about 1.3 kb was detected by northern blot analysis in all tissues examined (liver, kidney and brain) (data not shown). The 1.0-kbp cDNA insert of the longest clone was consistent with the observed mRNA size. The longest insert was subcloned into plasmid pBS and the cDNA sequence was determined. As shown in Fig. 3, the cDNA spanned 1,018 nucleotides plus poly(A) and predicted a 295-amino acid protein with a molecular mass of 32,823. The peptide sequences of the seven lysyl endopeptidase digested fragments completely matched those deduced from the cDNA sequence and were preceded by lysine, consistent with the cleavage specificity of lysyl endopeptidase (Fig. 3, boxes). Overall, 37% of the deduced amino acid sequence (109 residues) was supported by the sequential Edman degradation of the seven peptides, allowing the unequivocal identification of the cDNA clone. In addition, the amino acid composition inferred from the cDNA was very close to that obtained from a hydrolysate of the purified 40 kDa protein (Table 1).

Using the FASTA program [11], the amino acid sequence of the rat 40 kDa protein was compared with the amino acid sequences of more than 1,000 ribosomal proteins. Comparative analysis revealed that the 40 kDa protein shared homology with the prokaryotic small subunit ribosomal protein S2 of eubacteria and chloroplasts [12–15] (34.3% homology with *Escherichia coli* S2, 37.0%

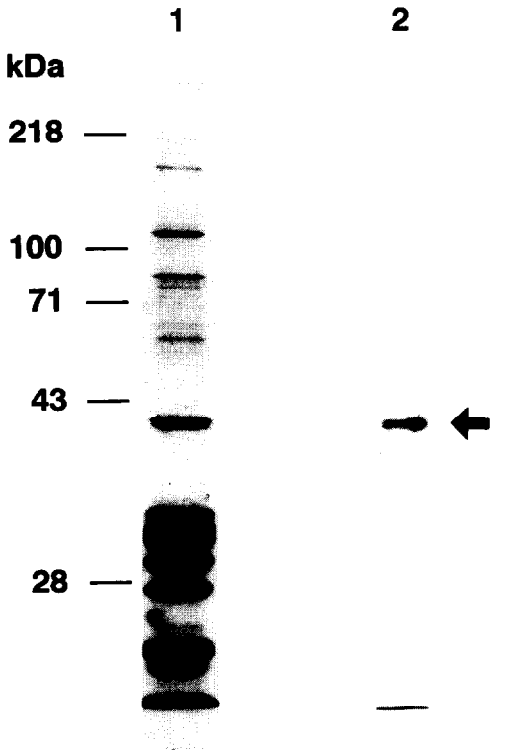


Fig. 2. Characterization of the purified 40 kDa protein in SDS-PAGE. The analysis was of 40 μg of the 40 S ribosomal proteins (lane 1) and 0.6 μg of a gel-purified 40 kDa ribosomal protein (lane 2). Proteins were separated on 10% SDS-PAGE. The gel was stained with CBB. Positions of the size markers are presented on the left. The 40 kDa ribosomal protein is indicated by the arrow on the right.

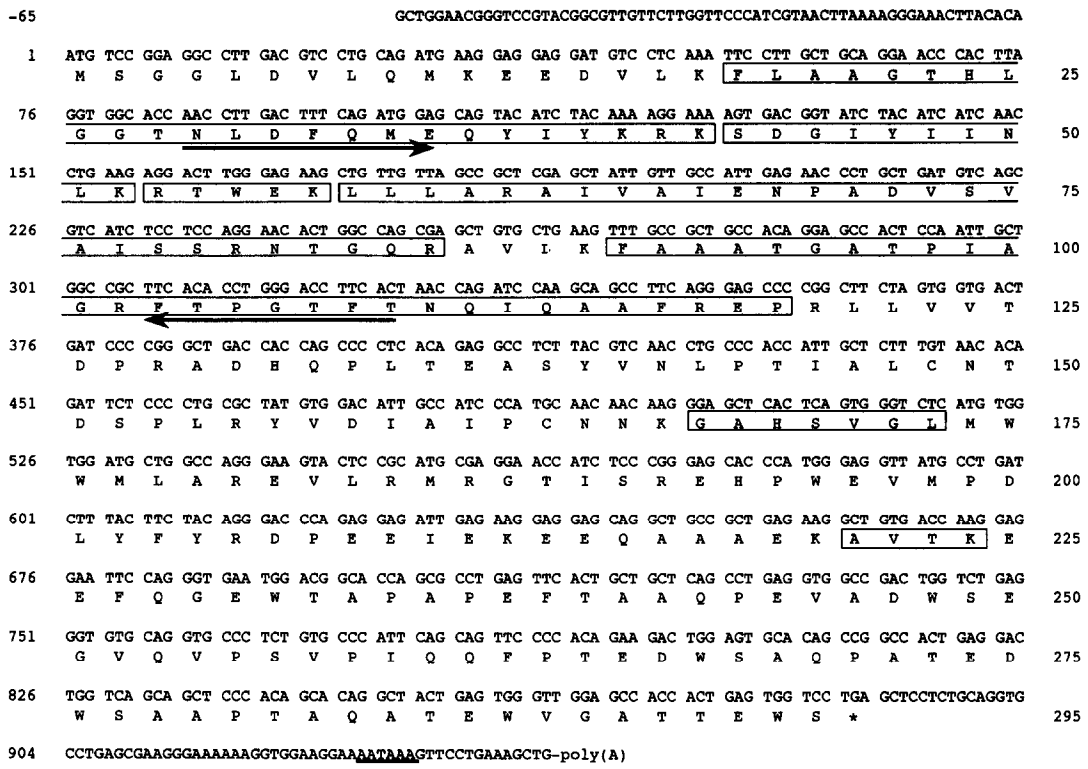


Fig. 3. Nucleotide and deduced amino acid sequences of the cDNA of rat 40 kDa ribosomal protein. Nucleotide residues are numbered in the 5′ to 3′ direction, beginning with the first residue of the ATG triplet encoding initiator methionine; nucleotides on the 5′ side of residue 1 are indicated by −65. The deduced amino acid sequence is given below the nucleotide sequence, and amino acid residues are numbered beginning with the initiator methionine. The asterisk indicates stop codon. The putative polyadenylation signal is underlined. The amino acid sequences identical with those determined from analysis of lysyl endopeptidase digested peptides are boxed. The arrows indicate the positions of the primers used for PCR. The nucleotide sequence data will appear in the DDBJ, EMBL and GeneBank Nucleotide Sequence Database under the accession number D25224.

Table 1
Amino acid composition of the rat 40 kDa ribosomal protein

| Amino acid | Amino acid composition (%) | |
|---|---|---|
| | Deduced[a] | Observed[b] |
| Asx | 7.77 | 8.6 |
| Glx | 13.85 | 13.7 |
| Ser | 4.73 | 5.7 |
| Gly | 5.41 | 8.1 |
| His | 1.35 | 1.7 |
| Arg | 5.07 | 4.3 |
| Thr | 7.77 | 7.4 |
| Ala | 12.50 | 13.0 |
| Pro | 6.76 | 8.8 |
| Tyr | 2.36 | 2.0 |
| Val | 6.42 | 5.4 |
| Ile | 5.07 | 4.1 |
| Leu | 7.09 | 6.5 |
| Phe | 3.38 | 3.8 |
| Lys | 3.72 | 3.7 |
| Met | 2.36 | –* |
| Cys | 0.68 | –* |
| Trp | 3.38 | –* |

[a]Amino acid composition deduced from cDNA sequence (Fig. 3).
[b]The amino acid composition of the protein was determined after acid hydrolysis followed by conversion of the amino acids to their phenylthiocarbamyl derivatives as previously described [4].
*Met, Cys and Trp were not determined.

homology with *Spirulina platensis* S2, 28.7% homology with *Zea mays* chloroplast S2 and 29.7% homology with *Nicotiana tabacum* chloroplast S2) (Fig. 4). The 40 kDa protein shared homology (24.7%) also with a *Saccharomyces cerevisiae* mitochondrial ribosomal protein [16]. These results suggest that the 40 kDa ribosomal protein is the mammalian counterpart to prokaryotic ribosomal protein S2.

We next compared the cDNA sequence of the rat 40 kDa ribosomal protein with sequences in the EMBL data bank (release 30). The search revealed that the rat cDNA showed 52–95% identity with the cDNAs isolated as cell surface 68 kDa laminin binding protein cDNAs of human (88%) [17,18], mouse (95%) [19], bovine (88%) [20], *Hydra vulgaris* (54%) [21] and *Drosophila melanogaster* (52%) [22]. Since most of the base differences between the species were in the wobble position, the deduced amino acid sequence of the rat cDNA shared 99% identity with those of mammalian cDNAs (human, mouse and bovine) and 57–59% identity with those of cDNAs of hydra (59%) and drosophila (57%) (Fig. 5), suggesting a possible correspondence.

In 1986, Wewer and co-workers isolated a short cDNA fragment by screening a human umbilical vein endothelial cell cDNA library with a monoclonal anti-

body against human cell surface 68 kDa laminin binding protein and reported the cDNA fragment as human 68 kDa laminin binding protein cDNA [17]. In 1988, Yow and co-workers isolated a 1.0 kbp cDNA highly expressed in human colon cancer and reported the cDNA as a full-length cDNA for the human 68 kDa laminin binding protein [18] because of its nucleotide sequence identity with the one found by Wewer and co-workers. Mouse and bovine cDNAs were isolated by nucleic acid hybridization using the human cDNA as a probe and were reported as 68 kDa laminin binding protein cDNAs for mouse [19] and bovine [20]. Hydra and Drosophila cDNAs [21,22] shared significant homology with the human cDNA and have been thought to be 68 kDa laminin binding protein cDNAs. The contradictions relate to both the physical and functional properties of the proteins encoded by the cDNAs, the most obvious of which is the large difference in size: the cDNAs encode proteins of approximately 33,000 Da, not 68,000 (Fig. 5). The cDNA-encoded proteins lack signal peptides, contain no N-linked carbohydrate attachment sites, and have no typical transmembrane domains ([18,20] and our unpublished data). The homology of the rat protein with the hydra and drosophila proteins was restricted to the 218 amino-terminal amino acids. The carboxyl-terminal part (77 amino acids), which was described as the laminin binding domain in the human protein [17], showed no sequence homology (Fig. 5, box). Why the human cDNA clone, isolated by Wewer and co-workers, reacted with an anti-68 kDa laminin binding protein antibody in the immunoscreening [17] is still unclear. However, because both the purified 68 kDa laminin binding protein from human placenta and the protein deduced from the human cDNA contained an octapeptide with the sequence Met-Leu-Ala-Arg-Glu-Val-Leu-Arg, the antibody they used may recognize the amino acid sequence containing the octapeptide.

In the present study, we have purified a 40 kDa protein in the rat 40 S ribosomal subunit and determined its primary structure, demonstrating that the protein is the eukaryotic counterpart to prokaryotic ribosomal protein S2. We also demonstrated a strong homology between the rat 40 kDa ribosomal protein and the protein deduced from the cDNA of 68 kDa laminin binding protein. In addition, the sequence of the mouse cDNA was identical to a mouse cytoplasmic protein of unknown function, p40 [23]. Recently McCaffery et al. reported that p40 coincided with a protein and RNA peak in the

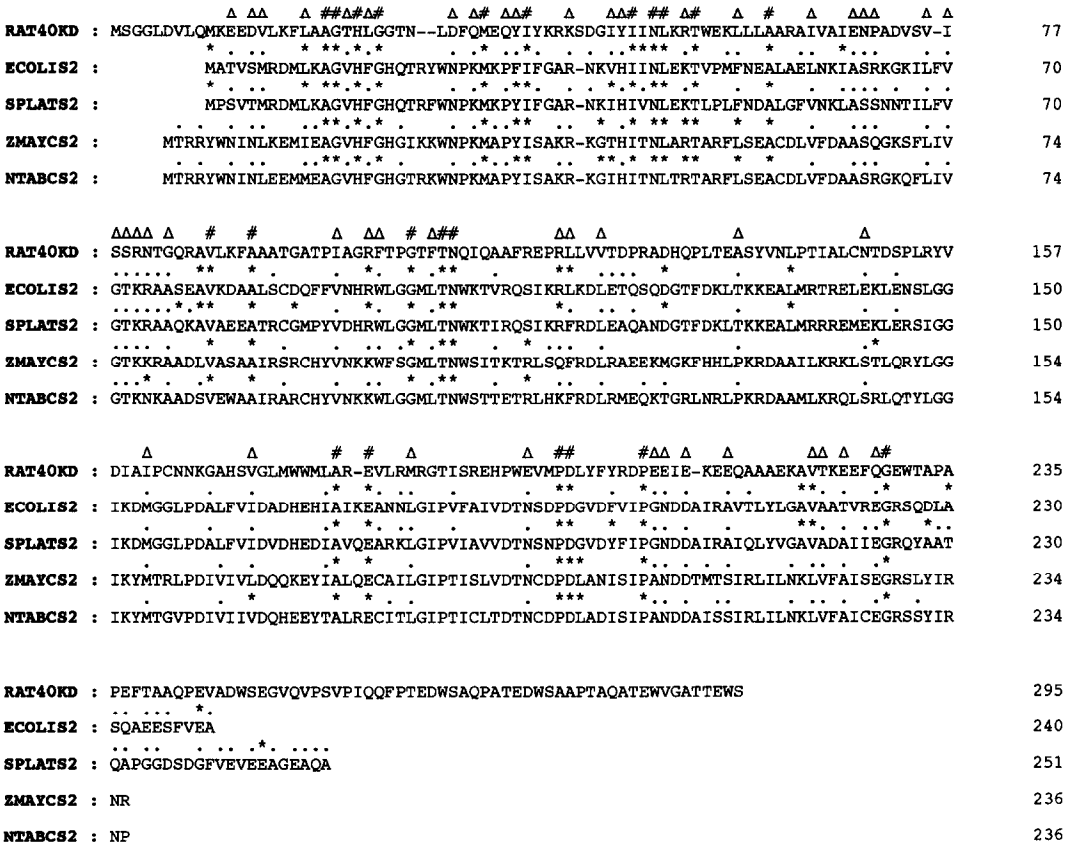


Fig. 4. Comparison of the amino acid sequence of the rat 40 kDa ribosomal protein with the sequences of ribosomal protein S2 of eubacteria and chloroplasts. Amino acid residues are numbered at the right. Residues identical to the 40 kDa ribosomal protein are designated by asterisks and those that are conservative substitutions are designated by dots over the letter symbols. Residues that are conserved among all the species are designated by # and those that are conservative substitutions among all the species are designated by Δ. Dashes denote gaps introduced to maximize homology. RAT40KD, rat 40 kDa ribosomal protein; ECOLIS2, *Escherichia coli* S2 [12]; SPLATS2, *Spirulina platensis* S2 [13]; ZMAYCS2, *Zea mays* chloroplast S2 [14]; NTABCS2, *Nicotiana tabacum* chloroplast S2 [15].

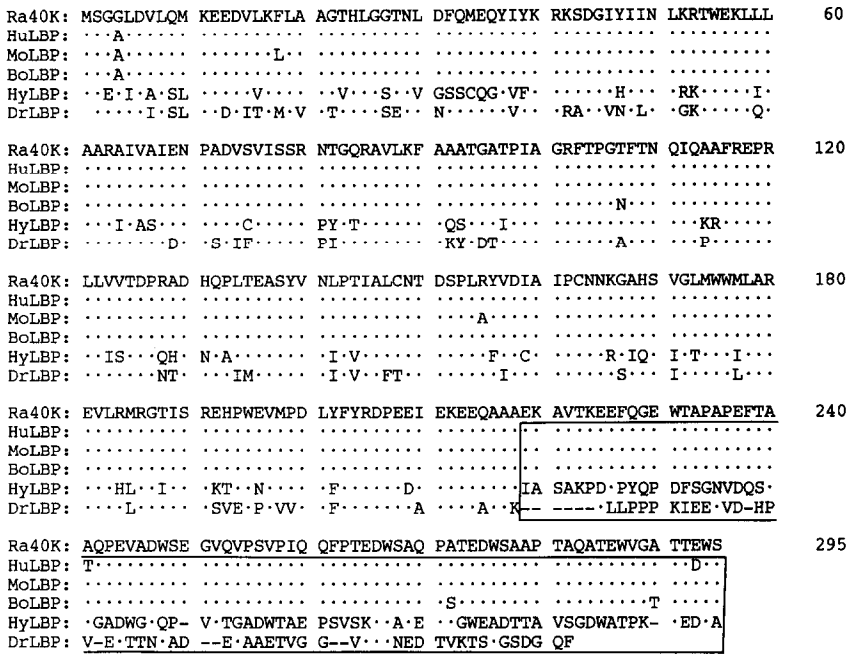


Fig. 5. Comparison of the amino acid sequences of the rat 40 kDa ribosomal protein with 68 kDa laminin binding proteins. Amino acid sequences were deduced from the cDNA sequences. Dots in place of a residue indicate identity with the rat 40 kDa ribosomal protein. Dashes denote gaps introduced to maximize homology. Ra40K, rat 40 kDa ribosomal protein; HuLBP, human 68 kDa laminin binding protein [17,18]; MoLBP, mouse 68 kDa laminin binding protein [19]; BoLBP, bovine 68 kDa laminin binding protein [20]; HyLBP, *Hydra vulgaris* homologue to mammalian laminin receptor [21]; and DrLBP, *Drosophila melanogaster* homologue to mammalian laminin receptor [22]. Amino acid residues are numbered at the right. A box indicates the corresponding region which was described as the putative laminin binding domain in human protein [18].

40 S ribosomal subunit region, suggesting that it is associated with the 40 S ribosomal subunit [24,25], and immunohistochemical study using the antibody against p40 showed a cytoplasmic granular labeling pattern, similar to that of a systemic lupus erythematosus serum specific for ribosomal proteins [24]. Our results, along with the properties of p40, strongly support the conclusion that the cDNAs which have been thought to encode the 68 kDa laminin binding protein do not encode the laminin receptor but, instead, encode the 40 kDa ribosomal protein.

*Acknowledgements:* We are grateful to Dr. Masayuki Yamamoto, Department of Biochemistry, Tohoku University School of Medicine for kindly providing a rat liver cDNA library; to Dr. Tatsuo Tanaka, Department of Biochemistry, University of the Ryukyus School of Medicine for sending data from the ribosomal protein data base; to Dr. Motoo Kitagawa, Institute for Molecular and Cellular Biology, Osaka University, for useful suggestions; to Mr. Brent Bell for reading the manuscript and to Hideo Kumagai and Reiko Torigoe for their skillful technical assistance.

## References

[1] Wittmann-Liebold, B., Köpke, A.K.E., Arndt, E., Krömer, W., Hatakeyama, T. and Wittmann, H.-G. (1990) in: The Structure, Function and Evolution of Ribosomes (Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D. and Warner, J.R., Eds.) pp. 598–616, American Society for Microbiology, Washington, DC.

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

[3] Hunkapiller, M.W., Lujan, E., Ostrander, F. and Hood, L.E. (1983) Methods Enzymol. 91, 227–236.

[4] Heinrikson, R.L. and Meredith, S.C. (1984) Anal. Biochem. 136, 65–74.

[5] Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. and Klenk, D.C. (1985) Anal. Biochem. 150, 76–85.

[6] Sommer, S.S., Sarkar, G., Koeberl, D.D., Bottema, C.D.K., Buerstedde, J.M., Schowalter, D.B. and Cassady, J.D. (1990) in: PCR Protocols (Innis, M.A.,Gelfand, D.H., Sninsky, J.J. and White, J.T., Eds.) pp. 197–205, Academic Press, San Diego, CA.

[7] Takasawa, S., Tohgo, A., Unno, M., Yonekura, H. and Okamoto, H. (1992) FEBS Lett. 307, 318–323.

[8] Yamamoto, M., Kure, S., Engel, J.D. and Hiraga, K. (1988) J. Biol. Chem. 263, 15973–15979.

[9] Takasawa, S., Tohgo, A., Unno, M., Shiga, K., Yonekura, H. and Okamoto, H. (1992) Biochim. Biophys. Acta 1132, 228–230.

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

[11] Peason, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.

[12] Wittmann-Liebold, B. and Bosserhoff, A. (1981) FEBS Lett. 129, 10–16.

[13] Sanangelantoni, A.M., Calogero, R.C., Buttarelli, F.R., Gualerzi, C.O. and Tiboni, O. (1990) FEMS Microbiol. Lett. 66, 141–146.

[14] Igroi, G.L., Döry, I. and Kössel, H. (1990) Nucleic Acids Res. 18, 663.

[15] Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K. and Sugiura, M. (1988) Plant Mol. Biol. 11, 589–600.

[16] Davis, S.C., Tzagoloff, A. and Ellis, S.R. (1992) J. Biol. Chem. 267, 5508–5514.

[17] Wewer, U.M., Liotta, L.A., Jaye, M., Ricca, G.A., Drohan, W.N., Claysmith, A.P., Rao, C.N., Wirth, P., Coligan, J.E., Albrechtsen, R., Mudryj, M. and Sobel, M.E. (1986) Proc. Natl. Acad. Sci. USA 83, 7137–7141.
[18] Yow, H., Wong, J.M., Chen, H.S., Lee, C., Steele, Jr., G.D., and Chen, L.B. (1988) Proc. Natl. Acad. Sci. USA 85, 6394–6398.
[19] Rao, C.N., Castronovo, V., Schmitt, M.C., Wewer, U.M., Claysmith, A.P., Liotta, L.A. and Sobel, M.E. (1989) Biochemistry 28, 7476–7486.
[20] Grosso, L.E., Park, P.W. and Mecham, R.P. (1991) Biochemistry 30, 3346–3350.
[21] Keppel, E. and Schaller, H.C. (1991) J. Cell Science 100, 789–797.
[22] Kim, Y.J. and Baker, B.S. EMBL Databank Accession number M77133.
[23] Makrides, S., Chitpatima, S.T., Bandyopadhyay, R. and Brawerman, G. (1988) Nucleic Acids Res. 16, 2349.
[24] McCaffery, P., Neve, R.L. and Dräger, U.C. (1990) Proc. Natl. Acad. Sci. USA 87, 8570–8574.
[25] Auth, D. and Brawerman, G. (1992) Proc. Natl. Acad. Sci. USA 89, 4368–4372.