THE PRIMARY STRUCTURE OF PROTEIN 44 FROM THE LARGE SUBUNIT OF YEAST RIBOSOMES

T. ITOH* and B. WITTMANN-LIEBOLD
Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, D-1000 Berlin-Dahlem, Germany

Received 30 October 1978

1. Introduction

Ribosomes are complex organelles consisting of three RNAs and 50–60 proteins in prokaryotes [1] and of four RNAs and 70–80 proteins in eukaryotes [2]. The best known ribosome is that of the bacterium *Escherichia coli;* for instance the primary structures of 45 *E. coli* ribosomal proteins have been determined in this laboratory or in collaboration with other groups [1,3].

Work on the primary structure of eukaryotic ribosomal proteins has recently started. The N-terminal sequences of an acidic ribosomal protein from the brine shrimp *Artemia salina* and from the yeast *Saccharomyces cerevisiae* have been elucidated, and they show a high degree of homology [4]. In addition, the N-terminal regions of 12 basic proteins [5] and of one acidic protein [6] from rat liver ribosomes have been determined. Finally, the amino acid sequences of one acidic ribosomal protein from *Artemia salina* [7] and of one basic protein from rat liver ribosomes [8] have almost been completed.

Determination of the primary structures of ribosomal proteins from various species gives direct information on changes in the structure of ribosomes during evolution. We report here the amino acid sequence of the protein 44 from *Saccharomyces cerevisiae* ribosomes, the first eukaryotic ribosomal protein whose primary structure has fully been determined.

* Present address: Department of Biochemistry and Biophysics, Research Institute for Nuclear Medicine and Biology, Hiroshima University, Hiroshima, Japan

2. Materials and methods

Proteins from 60 S subunits of *Saccharomyces cerevisiae* ribosomes were isolated as in [9]. Protein 44 was purified by CM-cellulose column chromatography in the presence of 6 M urea and 0.5 mM DTT with a linear gradient from 0.05–0.65 M sodium acetate buffer (pH 5.5) followed by Sephadex G-100 gel filtration. Desalting was done with Bio-Gel P10 in 15% acetic acid. The identity and purity of the protein were checked by two-dimensional SDS–polyacrylamide gel electrophoresis [9]. Enzymatic digestions were performed with trypsin, treated with 1-chloro-4-phenyl-3-tosylaminobutan-2-one for 4 h at 37°C and with chymotrypsin at pH 8.1 for 1 h at 37°C. Thermolysin digestion was at pH 8.0 for 2 h at 55°C, and digestion with *Staphylococcus aureus* protease was on performic acid oxidized protein at pH 4.0 for 20 h at 37°C.

Peptides were isolated by the fingerprint technique [10] on thin-layer plates (Cel 300 from Macherey and Nagel, Düren) or by chromatography on a micro-column (0.3 × 10 cm) of Dowex M71 at 55°C [11] developed with the following pyridine-formate gradients: 0.1 M (pH 2.7) to 1.0 M (pH 4.0) followed by 1.0 M (pH 4.0) to 2.0 M (pH 6.0).

Amino acid analyses were performed with a Durrum D500 analyser. For determination of cysteine the protein or the peptides were oxidized with performic acid before hydrolysis.

Determination of the amino acid sequence of peptides was either by the manual dansyl-Edman procedure [12,13] or by automatic Edman degradation in a solid-phase sequenator [14,15]. The N-terminal

region of the intact protein was sequenced by the automatic Edman degradation procedure [16] in an improved Beckman sequenator [17]. Aspartic acid, glutamic acid and their amides in the manually sequenced peptides were determined by converting the released 2-anilino-5-thiazolinones to the phenyl-thiohydantoins and identifying them by chromatography [17]. For quantitative determination, and where the identification after dansylation was uncertain, the free amino acids were liberated from their 2-anilino-5-thiazolinones or phenyl-thiohydantoins by hydrolysis and identified with a Durrum amino acid analyser. The determination of cysteine was achieved by oxidation of the released thiazolinone derivative with performic acid followed by amino acid analysis.

3. Results and discussion

The 30 N-terminal amino acid residues of yeast protein 44 were established with a liquid-phase sequenator (fig.1). Treatment of the protein with trypsin gave 25 peptides which, with the exception

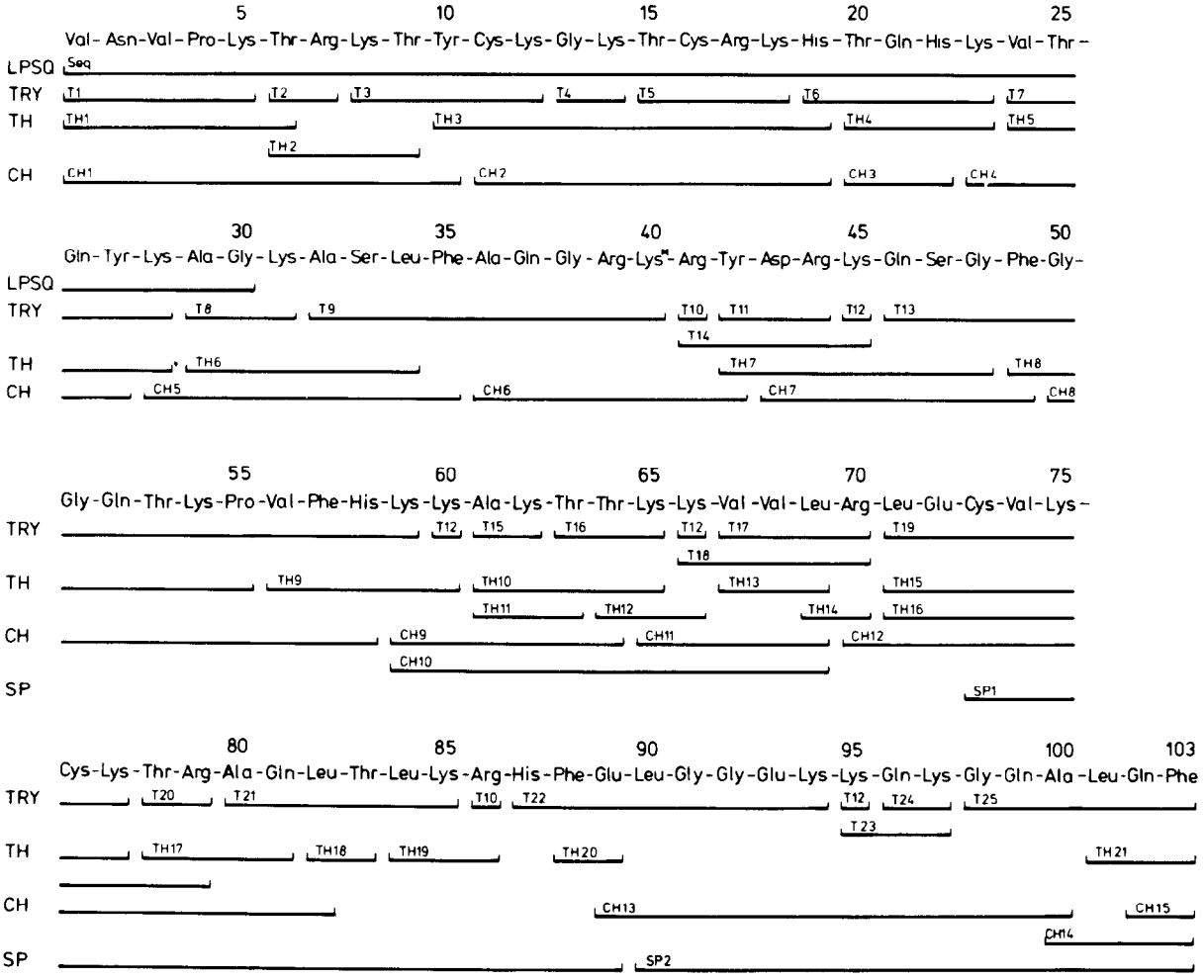

Fig.1. Primary structure of protein no. 44 from *Saccaromyces cerevisiae. Abbreviations:* TRY, tryptic peptides; TH, thermolytic peptides; CH, chymotryptic peptides; SP, peptides isolated after digestion with *Staphylococcus aureus* protease; LPSQ, degradation of the intact protein with a liquid phase sequenator.

of T13, were sequenced manually or in a solid-phase sequenator. Some tryptic peptides, e.g., T14, T18, T23, were due to incomplete cleavage.

Digestion with thermolysin gave 21, and with chymotrypsin 15 peptides. Determination of the sequence of peptides TH1, TH2, TH4, TH5, TH6, CH1, CH2, CH3, CH4 and CH5 established the alignment of the tryptic peptides T1-T2-T3-T4-T5-T6-T7-T8 (fig.1). The alignment of these 8 tryptic peptides was confirmed by the analysis with the liquid-phase sequenator.

Peptides CH5, CH6, CH7, CH8, CH10, CH12, TH7, TH8, TH9 and TH10 were sequenced manually or automatically, and the results established the alignment of T8-T9-Arg-T11-Lys-T13-Lys-T15-T16-Lys-T17-T19 in the middle part (positions 32–77) of the protein.

Digestion of the oxidized protein with *Staphylococcus aureus* protease gave two peptides, SP1 and SP2, and they were sequenced manually, resulting in the alignment of T19-T20-T21-Arg-T22-Lys-T24-T25 (positions 71–103). This alignment was confirmed by the sequence analyses of thermolytic and chymotryptic peptides (TH16, TH17, TH19, CH12, CH13). Treatment of the protein with carboxypeptidase A and B released the C-terminal residues phenylalanine and glutamine.

The combination of the described results gave the alignment of all peptides and the complete amino acid sequence as shown in fig.1.

An unusual amino acid was observed in positions 40 and 54; it eluted from the amino acid analyser at the same position as monomethyl-lysine. However, it migrated slightly slower on thin-layer plates in pyridine–acetone–3 M NH_4OH (50:30:25) [18], and trypsin split the protein chain after this unusual amino acid. These results indicate that the amino acid derivative is related to monomethyl-lysine. Further analyses are in progress to identify it unambiguously.

The amino acid composition of the yeast protein 44 derived from its sequence is: Asp_1, Asn_1, Thr_{10}, Ser_2, Glu_3, Gln_9, Pro_2, Gly_9, Ala_6, Val_7, Leu_7, Tyr_3, Phe_5, His_4, Lys_{20}, Lys^{*}_{2}, Arg_8, Cys_4. This is in excellent agreement with the data obtained from the total hydrolysis of the intact protein.

The mol. wt 11 849 as calculated from the protein sequence is close to the value of 10 800 determined by SDS–gel electrophoresis.

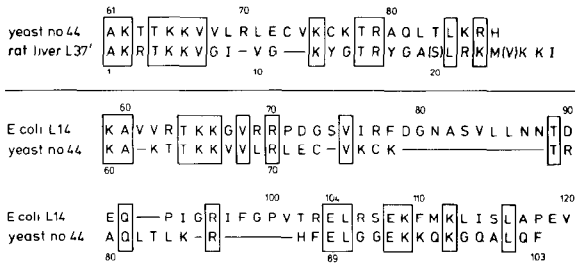

Fig.2. Comparison of regions from the yeast protein 44 on the one hand and from rat liver protein L37′ (above) and *E. coli* protein L14 (below) on the other hand. Abbreviations for amino acids are according to the one-letter code.

A computer analysis has been made to search for homologous structures among ribosomal proteins from different species. Yeast protein 44 was compared with 45 *E. coli* ribosomal proteins of known primary structures and to several eukaryotic ribosomal proteins whose sequences are partly or completely known. Possible homologous regions are shown in fig.2; however, it remains to be seen whether the observed structural similarities are significant.

Acknowledgements

We thank Dr K. Higo and Mr S. Tani for kindly providing ribosomal subunits, Miss B. Greuer for excellent technical assistance and Dr I. G. Wool for reading the manuscript.

References

[1] Brimacombe, R., Stöffler, G. and Wittmann, H. G. (1978) Ann. Rev. Biochem. 47, 217–249.
[2] Wool, I. G. (1979) Ann. Rev. Biochem. in press.
[3] Wittmann-Liebold, B., Brauer, D. and Dognin, J. (1977) in: Solid Phase Methods in Protein Sequence Analysis (Previero, A. and Previero-Coletti, A. M. eds) pp. 219–232, Elsevier/North-Holland Biomedical Press, Amsterdam, New York.
[4] Amons, R., Van Agthoven, A., Pluijms, W., Möller, W., Higo, K., Itoh, T. and Osawa, S. (1977) FEBS Lett. 81, 308–310.
[5] Wittmann-Liebold, B., Lin, A. and Wool, I. G. (1979) in preparation.

[6] Amons, R., Van Agthoven, A., Pluijms, W. and Möller, W. (1978) FEBS Lett. 86, 282–284.
[7] Amons, R. and Möller, W. (1978) personal communication.
[8] Lin, A., Wittmann-Liebold, B. and Wool, I. G. (1978) unpublished results.
[9] Otaka, E. and Kobata, K. (1978) Mol. Gen. Genet. 162, 259–268.
[10] Heiland, I., Brauer, D. and Wittmann-Liebold, B. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 1751–1770.
[11] Chen, R. and Wittmann-Liebold, B. (1975) FEBS Lett. 52, 139–140.
[12] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 379–380.
[13] Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873–886.
[14] Previero, A., Derancourt, J., Colletti-Previero, M. A. and Laursen, R. A. (1975) FEBS Lett. 33, 135–138.
[15] Wittmann-Liebold, B. and Lehmann, A. (1975) in: Solid Phase Methods in Protein Sequence Analysis (Laursen, R. A. ed) pp. 81–91, Pierce Chemical Co., Rockford, IL.
[16] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[17] Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z. Physiol. Chem. 354, 1415–1431.
[18] Kakimoto, Y. and Akazawa, S. (1970) J. Biochem. 245, 5751–5758.