

Sequence of the Amino-Terminal Region of Rat Liver Ribosomal Proteins S4, S6, S8, L6, L7a, L18, L27, L30, L37, L37a, and L39

B. Wittmann-Liebold, A.W. Geissler, A. Lin, and Ira G. Wool

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin 33 (Dahlem), Federal Republic of Germany (B.W-L., A.W.G.) and the Department of Biochemistry, The University of Chicago, Chicago, Illinois 60637 (A.L., I.G.W.)

The sequence of the amino-terminal region of eleven rat liver ribosomal proteins – S4, S6, S8, L6, L7a, L18, L27, L30, L37a, and L39 – was determined. The analysis confirmed the homogeneity of the proteins and suggests that they are unique, since no extensive common sequences were found. The N-terminal regions of the rat liver proteins were compared with amino acid sequences in *Saccharomyces cerevisiae* and in *Escherichia coli* ribosomal proteins. It seems likely that the proteins L37 from rat liver and Y55 from yeast ribosomes are homologous. It is possible that rat liver L7a or L37a or both are related to *S. cerevisiae* Y44, although the similar sequences are at the amino-terminus of the rat liver proteins and in an internal region of Y44. A number of similarities in the sequences of rat liver and *E. coli* ribosomal proteins have been found; however, it is not yet possible to say whether they connote a common ancestry.

Key words: rat liver ribosomal proteins, amino-terminus, yeast ribosomes

The solution of the structure of eukaryotic ribosomes requires information on the chemistry of the molecular components. Eighty-two proteins have been isolated from rat liver ribosomes and characterized (reviewed in Wool [1]). One reason for purifying appreciable amounts of protein was to have sufficient material so as to be able to do sequence studies. We have begun the exercise by determining the sequence of the amino-terminal region of eleven proteins: S4, S6, S8, L6, L7a, L18, L27, L30, L37, L37a, and L39. One purpose is to determine the relatedness of individual eukaryotic ribosomal proteins since it is possible that some arose by gene duplication [1]; a second purpose is to search for homologous ribosomal proteins by correlating sequences from rat liver and from other organisms, eg, *Escherichia coli* and *Saccharomyces cerevisiae*. In addition, we envision that the sequence data will be valuable at some later time for other purposes, for example, in determining the site of interaction of proteins and nucleic acids in comparison with similar studies made with *E. coli* ribosomes.

Received September 24, 1979; accepted October 3, 1979.

EXPERIMENTAL PROCEDURE

The isolation and characterization of rat liver ribosomal proteins S4, S6, S8, L6, L7a, L18, L27, L30, L37, L37a, and L39 have been described before [2–4].

Amino-terminal sequences were determined by automatic Edman degradation [5] in an improved Beckman sequencer [6], provided with dead volume-free delivery valves [7] and an automatic conversion device [8]. Protein samples (25–75 nmoles) were dissolved in water or dilute acetic acid and about 500 μg of dithioerythritol was added. A sample (about 1/50) of the solution was used to calculate the concentration of protein by amino acid analysis or by Folin test [9]. The sequences were determined using a program [8] that employs 0.25 M Quadrol as buffer, with two couplings and two cleavages per cycle; the program started with the second wash of solvent S3, which was changed from 1-chlorobutane to a mixture of benzene and 1,2-dichloroethane (1:1). The two solvent "S3" washes were collected and converted to PTH-amino acids in 20% trifluoroacetic acid in water containing 10 mg dithioerythritol per 100 ml; the reaction was at 54°C for 20 min. The solvents and reagents were especially purified [6]. Generally, the PTH-amino acids were identified by thin-layer chromatography [7], using in succession three solvent systems: 1) chloroform/1-propanol/2-propanol (98:1:1); 2) chloroform/methanol (1:1); 3) n-heptane/propionic acid/1,2 dichloroethane (58:17:25). Because of the small amount of protein available, in many analyses most of the sample was used for thin-layer chromatography. When possible the PTH-amino acids were identified also by high-pressure liquid chromatography (HPLC) at 254 nm. The Waters instrument was provided with a Merck Li Chrosorb RP-8 column (25 \times 0.4 cm); the analysis was at 40°C using a linear methanol and water gradient [10]. The most prominent spots were eluted for liquid chromatography. Those residues whose derivatives formed weak spots on the thin-layer sheet and could not be confirmed by HPLC analysis are put in brackets in Table I. PTH-dehydrothreonine migrates with PTH-leucine when analyzed by thin-layer chromatography, but the two were distinguished by liquid chromatography at 313 nm. In general, the PTH-derivatives of serine and cysteine were not released in amounts sufficient for unambiguous identification.

RESULTS AND DISCUSSION

The sequence of the amino-terminal region of eleven rat liver ribosomal proteins – S4, S6, S8, L7, L7a, L27, L30, L37, L37a, and L39 – was determined (Table I). The sequence studies confirm the homogeneity of the proteins, and a comparison of the sequences suggests that the individual proteins are unique.

Alanine (31%), glycine (19%), serine (17%), and valine (10%) account for most of the N-terminal amino acids of rat liver ribosomes, although proline, methionine and aspartic acid have also been identified [11]. Of the proteins whose amino-terminal sequence was determined, three had glycine (S8, L18, and L27), three had alanine (S4, L6, and L37a), and one each had serine (L39), valine (L30), methionine (S6), and proline (L7a) at the N-terminus; the N-terminal residue of L37 was not identified.

Protein S6 is the major phosphoprotein of rat liver ribosomes [12] and its phosphorylation is affected by a variety of physiologic stimuli, including hormones, cyclic AMP, viral infection, and changes of growth conditions of cells in culture (reviewed in Wool [1]). The amino-terminal 33 residues of S6 have two (positions 6 and 12) of the 15 serines in the protein; whether either is a site of phosphorylation is not yet known. Protein L6 binds to 5S and 5.8S rRNA [13–15]; L7 to elongator tRNA (N. Ulbrich and I.G. Wool, unpublished results). The functions in which the other protein participate are not yet known.

Comparison of Rat Liver Sequences

The N-terminal regions of the proteins studied contain no significant homologous sequences. Although the proteins have a large number of basic amino acids, and some (L27, L37, L37a, L39 are examples) contain only neutral and basic residues at the amino-terminus, still there are no sequence repeats to suggest (on the basis of very limited data) that the proteins are related. All *E coli* ribosomal proteins that have been sequenced appear to be unique [16].

Clusters of basic residues occur in rat liver ribosomal proteins: in S4 at positions 5–10; in S8 at positions 8–12; in L7a at positions 2–10; in L30 at positions 4–8; in L37 at positions 9–15, and at 19–25; and in L37a at positions 2–6, and at 22–27. On the other hand, S6 has an acidic region at positions 17–21. None of the proteins has an extended hydrophobic area.

Two rat liver proteins, S4 and L7a, have a N-terminal sequence rich in proline; moreover, their net charge, their distribution of charged amino acids, and their high content of alanine, valine, and leucine are similar. Still there is no evidence from inspecting the N-terminal sequences that these proteins derive from a single ancestral gene, even if one considers the replacement of amino acids caused by the exchange of one nucleotide. The possible relationship might be tested again later by a more sophisticated computer program that is under development. It should also be examined when more is known of the sequence of the two proteins.

Comparison of the Sequences of Rat Liver and Yeast Ribosomal Proteins

There is a homology between rat liver ribosomal protein L37 and the yeast *S cerevisiae* ribosomal protein SC-Y55* (the latter has been almost completely sequenced; Itoh and Wittmann-Liebold, unpublished); 16 of the 30 amino-terminal residues are identical (Fig. 1). Further, there is a relationship between rat liver L37a and yeast ribosomal protein SC-Y44 which has been sequenced [17]. However, the homology, if a real one, is more complicated. Positions 1–16 of the N-terminal region of RL-L37a resemble the region 61–79 of SC-Y44 (Fig. 2; see also Fig. 2 in ref. 17): 9 amino acids in a sequence of 16 are the same. There may also be a relation of rat liver L7a to yeast SC-Y44. A comparison of the amino-terminal region of L7a (positions 1–36) and the sequence 55–97 of SC-Y44 reveals 15 identical amino acid residues (Fig. 2). Because of the special chemical nature of ribosomal proteins (eg, their high content of basic residues) the significance of the amino acid identities in the several proteins is still not certain. Obviously more sequence data, especially for rat liver ribosomal proteins, would be helpful.

Comparison of the Sequence of Rat Liver and *E coli* Ribosomal Proteins

The sequence of 46 *E coli* ribosomal proteins has been determined [18, 19] and made part (along with the N-terminal region of the remaining proteins) of a computer program which can be used to search for similar sequences in the amino-terminal region of the rat liver ribosomal proteins and in the sequence of the two yeast proteins (SC-Y44, and SC-Y55). The results of the comparison are summarized below.

RL-S4. There are similarities in the sequence of the first 26 amino acids of rat liver ribosomal protein S4 and in the sequence of *E coli* S7 and S11 (Fig. 3); there are minimal identities with S14, S16, and S19 as well (results not shown). For the comparison the sequences were arranged so as to maximize the correspondence: Conservative replacements – for example, arginine (R) to lysine (K) or histidine (H), and valine (V) to isoleucine (I) – were scored; further, deletions were accommodated. Making those allowances, the best correlation is between RL-S4 and EC-S7; still the number of similar or identical residues is not large (12 of 26). The comparison of RL-S4 with EC-S11 reveals 10 of 26 similar or identical amino acids.

*The abbreviations used are: RL, rat liver; EC, *Escherichia coli*; SC-Y, *Saccharomyces cerevisiae*-yeast.

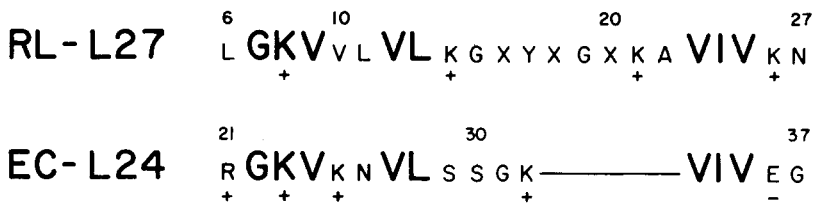


Fig. 5. Comparison of amino acid sequences from rat liver (RL-L27) and E coli (EC-L24) ribosomal proteins.



Fig. 6. Comparison of amino acid sequences from rat liver (RL-L39) and E coli (EC-L17, and EC-L18) ribosomal proteins.

CONCLUSIONS

It is difficult to judge the significance of the similarities in sequence between rat liver and E coli ribosomal proteins. Much more sequence data for the eukaryotic ribosomal proteins will be necessary to establish homologies. Moreover, there is a serious lack of criteria to assess the importance of limited similarities in sequence. Thus the present analyses serve as a guide for further studies of the sequence of rat liver proteins and, in one instance (the similarity of EC-L16 and RL-L37), of the function of a protein. It seems certain that ascertaining homologies between eukaryotic and prokaryotic ribosomal proteins by comparison of primary sequences will be difficult, since it is likely that during evolution conservative changes in amino acids, deletions, and recombination will have occurred. Some of the similarities that we have noted, if they are the vestiges of real homologies rather than merely chance phenomena, would indicate that those processes have happened (see, for example, the comparisons in Fig. 2). Immunologic studies (reviewed in refs. 1, 18, and 21) have shown that while prokaryotic and eukaryotic ribosomal proteins share determinants, cross-reaction is not common. Indeed, it has been demonstrated that antisera specific for chicken or rat liver ribosomal proteins recognize only about 20% of common determinants [22]. It is by no means surprising then that sequence homologies between rat liver and E coli ribosomal proteins are not prominent. Homologies might be more apparent if one knew and could compare the tertiary structure of these proteins. Nonetheless, intuition leads one to surmise that the ribosome arose on a single occasion and hence eukaryotic ribo-

somes are likely to be related to their prokaryotic ancestors. There is also circumstantial evidence that the fundamental structure and organization of ribosomes has been conserved (see Wool [1]), even while the primary structure of the components varies. To what extent those surmises on the evolution of ribosomes are correct is most likely to be determined from further studies of the sequence of the molecular constituents.

ACKNOWLEDGMENTS

The expenses of the research in Chicago were met by grants from the National Institutes of Health (GM-21769; CA-19265).

REFERENCES

1. Wool IG: *Ann Rev Biochem* 48:719–754, 1979.
2. Collatz E, Wool IG, Lin A, Stöffler G: *J Biol Chem* 251:4666–4672, 1976.
3. Tsurugi K, Collatz E, Wool IG, Lin A: *J Biol Chem* 251:7940–7946, 1976.
4. Tsurugi K, Collatz E, Todokoro K, Wool IG: *J Biol Chem* 252:3961–3969, 1977.
5. Edman P, Begg G: *Eur. J. Biochem.* 1:80–91, 1967.
6. Wittmann-Liebold B: *Hoppe-Seyler's Z. Physiol Chem* 354:1415–1431, 1973.
7. Wittmann-Liebold B, Geissler AW, Marzinzig E: *J Supramol Struct* 3:426–447, 1975.
8. Wittmann-Liebold B, Graffunder H, Kohls H: *Analyt Biochem* 75:621–633, 1976.
9. Lowry OH, Rosebrough NJ, Farr AL, Randall RJ: *J Biol Chem* 193:265–275, 1951.
10. Strickland M, Strickland WN, Brandt WF, von Holt C, Wittmann-Liebold B, Lehmann A: *Eur J Biochem* 89:443–452, 1978.
11. Low RB, Wool IG, Martin TE: *Biochim Biophys Acta* 194:190–202, 1969.
12. Gressner AM, Wool IG: *J Biol Chem* 249:6917–6925, 1974.
13. Ulbrich N, Wool IG: *J Biol Chem* 253:9049–9052, 1978.
14. Ulbrich N, Lin A, Wool IG: *J Biol Chem* 254:8641–8645.
15. Metspalu A, Saarma M, VILLEMS R, Ustav M, Lind A: *Eur J Biochem* 91:73–81, 1978.
16. Wittmann-Liebold B, Dzionara M: *FEBS Lett* 65:281–283, 1976.
17. Itoh T, Wittmann-Liebold B: *FEBS Lett* 96:399–402, 1978.
18. Brimacombe R, Stöffler G, Wittmann HG: *Ann Rev Biochem* 47:217–249, 1978.
19. Wittmann-Liebold B, Brauer D, Dognin J: In Previero A, Coletti-Previero MA (eds): "Solid Phase Methods in Protein Sequence Analysis." Amsterdam: Elsevier, North-Holland, pp 219–232.
20. Subramanian AR, Wittmann-Liebold B, Geissler AW, Stöffler G, Giesen M: *FEBS Lett* 99:357–360, 1979.
21. Wool IG, Stöffler G: In Nomura M, Tissières A, Lengyel P (eds): "Ribosome." Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1974, pp 417–460.
22. Fischer N, Stöffler G, Wool IG: *J Biol Chem* 253:7355–7360, 1978.