The covalent structure of rat ribosomal protein L7 was determined in part from the sequence of nucleotides in a recombinant cDNA and in part from the sequence of amino acids in portions of the protein. The complementary analyses supplemented and confirmed each other. Ribosomal protein L7 contains 258 amino acids and has a molecular weight of 30,040. The protein has an unusual and striking structural feature near the NH₂ terminus: five tandem repeats of a sequence of 12 residues. Rat L7 appears to be related to ribosomal protein L7 from the moderate halophile Vibrio costicola and perhaps to L30 from Bacillus stearothermophilus, to L7 from the moderate halophile NRCC 41227, and to L22 from Nicotinia tobaccum chloroplast. In addition, there is a sequence of 24 amino acids in rat protein L7 that may be related to segments of the same number of residues in Escherichia coli ribosomal proteins S10, S15, L9, and L22.

It is an article of faith that a molecular account of the function of eukaryotic ribosomes will follow from knowledge of the structure and that this has as an indispensable prerequisite information on the chemistry of the constituent proteins and nucleic acids. Progress has been made in this analysis, although, a good deal remains to be done (1). The covalent structure of the four species of RNA in rat ribosomes, 5S (2), 5.8S (3), 18S (4, 5), and 28S (6, 7), has been established. In addition, 84 proteins have been isolated from the particles (8) and the sequence of amino acids in several, P2 (9), L37 (10), and L39 (11), has been determined directly. This inventory of data is not only necessary for the resolution of the structure of the organelle, but also for analyzing the interaction of the proteins with the nucleic acids. The task of determining the structure of eukaryotic ribosomal proteins is being expedited by the application of recombinant DNA technology. Thus, the structure of a number of rat ribosomal proteins has been determined from recombinant cDNAs. They include S11 (12), S26 (13), S17 and L30 (14), L35a (15), and L19 (16). In a similar way the structure of three mouse ribosomal proteins, L30 (17), L32 (18), and S16 (19) and of the Chinese hamster protein S14 (20) has been established. We report here the structure of rat ribosomal protein L7 which we have inferred for the most part from the sequence of nucleotides in a recombinant cDNA but which we have completed and confirmed by sequencing portions of the protein.

There is a second purpose that underlies the analysis of the structure of the components of the ribosome and that is to understand their evolution. It is assumed that an organelle that is at one and the same time universal, essential, and complicated, as the ribosome is, arose on a single occasion. Indeed, the evidence for homology of rRNAs from evolutionarily distant species is substantial (21), albeit the relationship is most easily seen in comparisons of their secondary structures rather than of their nucleotide sequences. In a similar manner, comparison of the sequences of amino acids in ribosomal proteins may inform us concerning the details of their evolution. In addition, identification of conserved amino acid sequences cannot but help in unraveling the function of the proteins. Finally, these comparisons may provide clues as to why the number of proteins has increased from the 52 contained in prokaryotic ribosomes to the 70-80 that are found in eukaryotes without any significant changes in the reactions that the particles catalyze.

EXPERIMENTAL PROCEDURES

RESULTS AND DISCUSSION

The Sequence of Nucleotides in the Recombinant cDNA Encoding Rat Ribosomal Protein L7—A preliminary restriction endonuclease map of the cDNA insert in pL7-2 was prepared, and a set of enzymes was selected that would generate overlapping oligonucleotides suitable for a determination of the sequence (Fig. 3). Nucleotide sequences from both strands of the DNA, and overlapping sequences for each restriction site, were obtained (Fig. 3).

The cDNA insert in pL7-2 contains 826 nucleotides and includes the 5' poly(A) and 3' poly(C) homopolymer linkers, a portion of the 5' noncoding region, and a single open reading frame (Fig. 4). In the other two reading frames the sequence is interrupted by many termination codons. The open reading frame of 756 nucleotides begins at an ATG codon at position 58 and ends with a codon (ATA) for isoleucine. There

* Portions of this paper (including "Experimental Procedures," Tables I–III, and Figs. 1–3, 6, 7, and 9) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 950 Rockville Pike, Bethesda, MD 20814. Request Document No. 87M-0778, cite the authors, and include a check or money order for $5.20 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.

1 Portions of this paper were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

This paper is dedicated to the memory of David Vazquez, an esteemed colleague.

‡ Supported by the S. A. Schonbrun Research Endowment Fund.
§ Supported by National Institutes of Health Grant GM 21769.
is no termination codon and no 3' poly(A) sequence. We presume pL7-2 lacks a 3' end, and hence that the ribosomal protein L7 cDNA insert lacked the nucleotides coding for the carboxyl-terminal amino acids of the protein L7, because of a failure to make a full copy during the synthesis of the second strand of the cDNA. This assumption was substantiated later.

The context in which the initiation codon occurs, ACCATGG, is the sequence considered optimum for initiation of translation of mRNAs by ribosomes (36, 37).

We take notice that the first nucleotide of the L7 cDNA (position 36 in Fig. 4) is opposite minus 22 if the A in the initiation codon ATG is taken as +1) is a cytosine and that it is followed by an initiation triplet of ACC. A similar run of pyrimidines has been found in the 5' untranslated region of many eukaryotic ribosomal protein mRNAs in species as diverse as mammals (16-19) and amphibians (38). The conservation of this track of pyrimidines at about the same position (i.e. near the initiation codon) certainly suggests that it plays a role in the regulation of the translation of at least some of the mRNAs for eukaryotic ribosomal proteins.

The Primary Structure of Rat Ribosomal Protein L7—The reading frame extends from nucleotide 58 to 813 and encodes a protein of 252 amino acids (Fig. 4). The polypeptide was identified as rat ribosomal protein L7, in the first instance, by positive hybridization in a translation assay. The radioactive product in this reaction migrated on one-dimensional sodium dodecyl sulfate and two-dimensional urea-polyacrylamide gels with authentic L7 (results not shown). The molecular weight, calculated from the sequence of amino acids deduced from the DNA sequence was close to that estimated from the migration of the purified protein on sodium dodecyl sulfate gels (39), and the number of individual amino acid residues obtained from the sequence of the cDNA (Table I) approximated the number derived from an analysis of a hydrolysate of L7 isolated from rat ribosomes (39). Thus, we assumed that the pL7-2 insert lacked only the codons specifying a small number of amino acids at the carboxyl terminus.

Since the recombinant cDNA lacked a termination codon it was necessary to determine the carboxyl-terminal sequence of amino acids in L7 directly. We wished also to confirm the sequence of amino acids deduced from the sequence of nucleotides, to authenticate that pL7-2 encodes ribosomal protein L7, and to establish the identity of the blocked NH2-terminal amino acid. For that purpose, L7 was cleaved with cyanogen bromide and the peptides were isolated (Fig. 2). The sequence of amino acids at the NH2 terminus of CN4 (14 residues) corresponds exactly to the sequence of amino acids deduced from the sequence of CN4; and fraction g, a mixture of CN2 and CN3 (34).

The sequence of amino acids at the NH2 terminus of CN4 was confirmed by reaction with dansyl-4'-isothiocyanate (Fig. 2). The purity and the identity of the peptides in the fractions was assessed by polyacrylamide gel electrophoresis in sodium dodecyl sulfate, by reaction with dansyl reagent, by determination of the amino acid composition, and from the sequence of amino acids at the NH2 terminus. The results of these analyses and the sequence of amino acids deduced from the sequence of nucleotides in pL7-2 were used as a guide to order the cyanogen bromide peptides.
Structure of Rat Ribosomal Protein L7

The sequencing strategy and the covalent structure of rat ribosomal protein L7. The methods used to determine the sequence of amino acids in the cyanogen bromide peptides (designated CN1-7) are given at the lower right.
richia coli L7/L12 is the prototype (32), and to rat L7.

To ascertain if L7 and the other ribosomal proteins in the library contained similar partial sequences, we examined the spectrum of count scores, calculated by RELATE (Table II) as we had done before (10, 16). For a given fragment score, n, the program determines the number of fragments with scores exceeding n for both the real and randomized comparisons. The difference is expressed as a distance and is given in standard deviation units. In the comparison of two proteins with a short conserved region one would expect only a few high scores. Unfortunately, the value at which the spectrum of count score assumes statistical significance has not been established and there is no assurance that 3.0 standard deviations, for example, is of consequence. For that reason, we have used these scores only to identify protein fragments that might show homology and then have aligned the fragments by eye. The fragments from four E. coli ribosomal proteins (S10, S15, L9, and L22) showed similarity with one segment of rat L7 (Fig. 7). Of the 24 residues in this segment (amino acids 106-129), 20 occurred at the same position as identical or related amino acids (lysine and arginine, or isoleucine, leucine, and valine, or serine and threonine) in one or more of the E. coli fragments. For S10 there were 10 identities, for S15 10, for L9 7, and for L22 11. In addition, at three of the remaining four positions there were identities amongst the E. coli proteins that did not appear in rat L7; thus, at only 1 of the 24 positions are there unrelated amino acids in the four fragments (Fig. 7). It is noteworthy that this fragment of rat L7 (residues 106-129) contains 7 basic (arginine and lysine) and 10 hydrophobic (isoleucine, leucine, and valine) amino acids and that they account for most of the identities (see also later). Finally, high scores were obtained for a comparison of separate fragments from rat L7 and B. steaothermophilus S9 and for rat L7 and N. tobacoum L2 (Table II).

These results are reminiscent of earlier findings that a segment in rat ribosomal protein L37 and in yeast YP55 might be related to sequences in three E. coli proteins, S4, L20, and L34 (10), and that a portion of rat L19 is, perhaps, related to amino acid sequences in E. coli ribosomal proteins S2, L18, and L30 (16). An analogous observation has been made for a fragment of rat ribosomal protein S10 for both the real and randomized comparisons. The raw data generated by RELATE includes a list of the segments showing the greatest similarity as well as the number of amino acids separating them, i.e. the displacement between fragments. For rat L7, multiples of 12 amino acids predominated. The fragments yielding the highest scores were at the NH₂ terminus forming five tandem repeats of 12 residues each (Fig. 8). There is within the five tandem repeats of 12 residues two longer tandem repeats of 22 amino acids. They are at positions 3-24 and 25-46 (Table III). There are 11 identities (we include arginine/lysine and leucine/valine pairs) in the two fragments. Thus, the similarity between the short repeats is more striking than for the longer ones. Finally, there is another possible tandem repeat in L7 of 20 amino acids at positions 138-157 and 158-177 (Table III), although, in this instance the number of like residues is only 5.

Although the tandem repeats are an unusual and striking feature of the structure of L7 we are unable to assign a function to them. This is at least in part because we know nothing of the function of L7 itself. Repeat sequences of amino acids in ribosomal proteins have been reported before, although they are not common. E. coli protein S1 which is required for the binding of mRNA to the ribosome during the initiation of protein synthesis (45), and which as the α-subunit of the Qβ replicase is involved in the transcription of the plus strand of the RNA phage, has multiple internal repeats (46). The exact number of the repeats and their length is not certain; either six of 87 residues or 12 of 44 residues are possible. The function of the repeats is not known, although the suggestion has been made, without evidence to support it, that they are involved in binding mRNA to the ribosome and phage RNA to the catalytic subunit of Qβ replicase. There may be internal repeats in E. coli ribosomal protein L2 also (47) but they have not been analyzed so extensively and

\begin{align*}
K_5 & K K K K V A A A L G T L_{18} \\
K_{15} & K K K K V A V P E T L_{30} \\
K_3 & R R R R F A E L K V_{42} \\
K_{43} & R L R R K F A L K T L_{54} \\
R_{50} & K A R R K L I Y E K A_{66} \\
K R & K R K R R K K L L E T L V
\end{align*}

\begin{align*}
A_{24} & R M A R K K A G N F T V_{35} \\
L_{24} & R L R Q I F N G T F V_{35} \\
N_{15} & K K R I A L T D N S L_{86} \\
G_{30} & G M K K T T H F V E_{24}
\end{align*}

**Fig. 8.** The tandem repeat of 12 amino acids near the NH₂ terminus of rat ribosomal protein L7 and possible related fragments at other sites in the protein. The subscripts to the letters designate the position of the residue in the amino acid sequence of protein L7. A consensus sequence for the tandem repeats is given; identical or related residues that occur at the same position in three or more of the fragments are designated by bold letters and those that occur in two fragments are designated by thin letters.

*J. McNally and L. G. Wool, unpublished observation.*
hence, are not as certain as those in S1.

One way to secure leads to the function of a protein, or a segment thereof, is to examine it in higher order structure. The most desirable data comes from x-ray diffraction of crystals but in the absence of this information it is sometimes useful to examine predictions of the secondary structure (48). The most desirable data comes from x-ray diffraction of crystals (alanyl, phenylalanyl, valyl, and leucyl) in the remainder (Fig. 8). The several programs we have used vary in their predictions of the secondary structure of protein L7. However, there is agreement that the region of the five tandem repeats is likely to have a good deal of a-helical structure (Fig. 9). The distribution of amino acids in the repeats and the possibility that they have an a-helical conformation led us to consider that they might be amphipathic and, hence, capable of inserting into the lipid layer of the rough endoplasmic reticulum. In this way, L7 and more specifically the tandem repeats might be responsible for positioning the large subunit of the ribosome on the membrane. As a first approximation to a test of the proposition we examined Edmundson helical wheels (48) for the region of the tandem repeats (positions 7-66) to determine whether there was segregation of hydrophilic and hydrophobic residues. The results did not lend support to the proposal, although it is still possible that the tandem repeats participate in the positioning of the ribosome on the endoplasmic reticulum as a reflection of some other aspect of their structure. The tandem repeats may also play a role in the association of L7 to one or more of the rRNAs, although there is no direct evidence to support the suggestion.

In addition to the five tandem repeats near the NH2 terminus of L7, there are variations of this 12-residue sequence of the region of the tandem repeats (positions 7-66) to determine whether there was segregation of hydrophilic and hydrophobic residues. The results did not nearly so exact as in the tandem repeats. Nonetheless, there may be as many as nine repeats of this 12-residue segment in rat ribosomal protein L7 and they are unlikely to be without functional significance.

Acknowledgment—We are grateful to Jose Olvera for valuable technical assistance and to Arlene Timosciek for help in the preparation of the manuscript.

REFERENCES
1. Wool, I. G. (1986) in Structure, Function, and Genetics of Ribosomes (Hardesty, B. and Kramer, G., eds) pp. 391-411, Springer Verlag New York Inc., New York
2. Aoyama, K., Hitaka, S., Tanaka, T., and Ishikawa, K. (1982) J. Biochem. (Tokyo) 91, 363-367
3. Nazar, R. N., Sitz, T. O., and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597
4. Torezerski, R., Blobom, A. P., and Fuke, M. (1983) Nucleic Acids Res. 11, 4879-4890
5. Chan, Y. L., Gunell, R., Noller, H. F., and Wool, I. G. (1984) J. Biol. Chem. 259, 224-230
6. Chan, Y. L., Olvera, J., and Wool, I. G. (1983) Nucleic Acids Res. 11, 7819-7831
7. Hadijolov, A. A., Georgiev, O. L., Nosikov, V. V., and Yavachev, R. S.
8. It is curious that if one takes the first of the tandem repeats (positions 7-18) as the canonical sequence then it appears that there is a tendency for the amount of variation to increase with each repeat.

L. P. (1984) Nucleic Acids Res. 12, 3677-3693
9. Wool, I. G. (1979) Annu. Rev. Biochem. 48, 719-754
10. Lin, A., Wittmann-Liebold, B., McNally, J., and Wool, I. G. (1982) J. Biol. Chem. 257, 9189-9197
11. Lin, A., McNally, J., and Wool, I. G. (1983) J. Biol. Chem. 258, 10664-10671
12. Tanaka, T., Kuwano, Y., Ishikawa, K., and Ogata, K. (1985) J. Biol. Chem. 260, 6529-6533
13. Kuwano, Y., Nakashizuka, O., Nohshima, K., Tanaka, T., and Ogata, K. (1985) J. Biochem. 97, 883-992
14. Nakashizuka, O., Oyanagi, M., Kuwano, Y., Tanaka, T., Nakayama, T., Matsui, H., Nohshima, Y., and Ogata, K. (1985) Gene (Amst.) 35, 289-296
15. Tanaka, T., Wakisaka, K., Kuwano, Y., Ishikawa, K., and Ogata, K. (1986) Eur. J. Biochem. 154, 523-527
16. Chan, Y. L., Lin, A., McNally, J., Feleg, D., Meyuhas, O., and Wool, I. G. (1987) J. Biol. Chem. 262, 1111-1115
17. Wiedemann, L. M., and Perry, R. P. (1984) Mol. Cell. Biol. 4, 2518-2528
18. Dudov, K. P., and Perry, R. P. (1984) Cell 37, 457-468
19. Wagner, M., and Perry, R. P. (1985) Mol. Cell. Biol. 5, 3560-3576
20. Rhoads, D. D., and Roufa, D. J. (1985) Mol. Cell. Biol. 5, 1655-1659
21. Noller, H. F. (1984) Annu. Rev. Biochem. 53, 119-162
22. Collatz, E., Wool, I. G., Lin, A., and Stoffler, G. (1976) J. Biol. Chem. 251, 4666-4672
23. Laslum, U. K. (1970) Nature 227, 680-685
24. Lin, A., Collatz, E., and Wool, I. G. (1976) Mol. Gen. Genet. 144, 1-9
25. Chang, J. Y., Brauer, D., and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 203-214
26. Dayhoff, M. O. (1978) in Atlas of Protein Sequence and Structure, Vol. 5, pp. 1-8, Suppl. 3, National Biomedical Research Foundation, Washington, D. C.
27. Falkenberg, P., Yaguchi, M., Rollin, C. F., Matheson, A. T., and Wydro, R. (1979) Biochim. Biophys. Acta 578, 207-215
28. Kimura, M. (1984) J. Biol. Chem. 259, 1051-1055
29. Kimura, M., and Chow, C. K. (1984) Eur. J. Biochem. 139, 225-234
30. Falkenberg, P., Yaguchi, M., Roy, C., Zunker, M., and Matheson, A. T. (1986) Biochim. Cell Biol. 64, 675-680
31. Tanaka, M., Wakisugi, T., Sugita, M., Shimozaki, K., and Suguro, M. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 6030-6034
32. Wittmann-Liebold, B. (1984) Adv. Protein Chem. 36, 56-78
33. McDonell, M. W., Simon, M. N., and Studier, F. W. (1977) J. Mol. Biol. 110, 119-146
34. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560
35. Queen, C. L., and Korn, L. J. (1980) Methods Enzymol. 65, 596-609
36. Kozak, M. (1984) Nucleic Acids Res. 12, 857-872
37. Kozak, M. (1986) Cell 44, 283-292
38. Loreni, F., Ruberti, I., Bozizzoni, I., Pierandre-Amaldi, P., and Amaldi, F. (1985) EMBO J. 4, 3483-3486
39. Tsuruji, K., Collatz, E., Wool, I. G., and Lin, A. (1976) J. Biol. Chem. 251, 7340-7496
40. Giri, L., Hill, W. E., and Wittmann, H. G. (1984) Adv. Protein Chem. 36, 1-78
41. Barker, W. C., Ketcham, L. K., and Dayhoff, M. O. (1978) J. Mol. Biol. 10, 265-281
42. Doolittle, R. F., Woodbury, N. W., and Jue, R. A. (1982) Biocoi. Rep. 2, 405-412
43. Jue, R. A., Woodbury, N. W., and Doolittle, R. F. (1980) J. Mol. Biol. 15, 129-148
44. Doolittle, R. F. (1981) Science 214, 149-159
45. Suganumayama, T., and Subramanian, A. R. (1979) J. Mol. Biol. 127, 41-54
46. Wittmann-Liebold, B., Ashkan, M., and Dzionara, M. (1983) FEBS Lett. 154, 51-51
47. Kimura, M., Mende, L., and Wittmann-Liebold, B. (1982) FEBS Lett. 149, 304-312
48. Dzionara, M., Robinson, S. M. L., and Wittmann-Liebold, B. (1977) Hoppe-Seyler's Z. Physiol. Chem. 358, 1003-1019
49. Burgess, A. W., Ponnuswamy, P. K., and Scheraga, H. A. (1974) Isr. J. Chem. 12, 239-258
THE PRIMARY STRUCTURE OF RAT RIBOSOMAL PROTEIN L7

of Five Tandem Repeats of a Sequence of Twelve Amino Acids

by

Alan Liu, Fung-Ling Chen, James McNally, David Pelling, Oded Meytal, and Bars G. Woll

EXPERIMENTAL PROCEDURES

The following were either described or cited previously (44), the source of the restriction endonucleases, of the protein-treated wheat germ extract for the translation of mRNA, and of the chemicals, radioactive substrates, and other materials; the preparation of poly(A)-mRNA and of the oligo-

transcription of the nucleic acid; the construction and identification of a rat ribosomal protein cDNA clone in the present experiments the recombinant plasmid, pL1-2, which encodes L7; the method used for the positive hybridization translation assay, and the procedures used to determine the sequence of

amino acids in the DNA.

Protein Sequence Determination - Protein L7 was isolated by high performance liquid chromatography from group fraction 240 (22) on a reverse phase column (Vydac, Hi-flm RP-318; 250 x 4.6 mm I.D.) with a linear gradient of 5 to 60% acetonitrile containing 0.1% trifluoroacetic acid at a flow rate of 0.7 ml/min for 60 min. The identity and the homogeneity of L7 was established by electrophoresis on one-dimensional polyacrylamide gels containing sodium dodecyl sulfate (21) and by two-dimensional electrophoresis (24) on polyacrylamide gels containing urea (25). The amino acid composition of L7 was determined here from the analysis of hydrolyzates. No reaction was observed when L7 was treated with dinitro-reagent (26) or when the protein was degraded with DAEAC(27) (26) and, hence, we assume that the terminal amino group is blocked. Purified L7 (50 nmol) was cleaved with cyanogen bromide (26) and the mixture of fragments of L7 was resolved (28) by high performance liquid chromatography on a reverse phase column (Spheris I-RP-300 110, 100 x 5,1 mm I.D. using the same gradient as was employed for the preparation of L7 (see above). The sequence of amino acids in the cyanogen bromide peptides was determined by the on-resin manual DABITRCPK double coupling method (26, 28), or by automated liquid-phase Edman degradation (29), or with an automated gas-phase sequence (Applied Biosystems, Model 473A). The computer program RELATE (29) was used to assess possible evolutionary relationships between ribosomal proteins. The scoring matrix was Dayhoff's NWR 95 (29). The amino acid sequences that are referred to are: Vibrio furnissi L7 (30), Bacillus stearothermophilus L3 (28) and S9 (29), the moderate halophile NERCC 4127, L7 (26), Stenotrophomonas chloroplast L7 and L2 (31), and Stenotrophomonas chloroplast L12 (32).

TABLE I

Amino Acid Composition of Rat Ribosomal Protein L7

The amino acid composition was either determined from an analysis of hydrolyzates of purified L7 (A) or inferred from the sequence of nucleotides in a recombinant cDNA (B).

| Amino Acid | A  | B |
|------------|----|---|
| Alanine    | 21 | 30 |
| Arginine   | 26 | 35 |
| Aspartic acid and Asparagine | 14 | (15 + 6) |
| Cysteine   | 16 | (18 + 6) |
| Glutamic acid and Glutamine | 19 | 16 |
| Histidine  | 17 | 10 |
| Isoleucine | 17 | 17 |
| Leucine    | 22 | 23 |
| Lysine     | 18 | 17 |
| Methionine | 4  | 7  |
| Phenylalanine | 3 | 11 |
| Proline    | 11 | 11 |
| Serine     | 6  | 7  |
| Threonine  | 5  | 5  |
| Tryptophan | 8  | 5  |
| Tyrosine   | 12 | 10 |
| Valine     | 15 | 17 |
| Cystine    | 22 | 30 |

The number of residues has been corrected for the six amino acids absent from the protein encoded in plt-L.

TABLE II

Alignment of Related Amino Acid Sequences in Ribosomal Protein L7

and in B. stearothermophilus 58 and N. tobamovirus chloroplast L5

| Rat L7 | 58 | B. stearothermophilus L3 | N. tobamovirus chloroplast L5 |
|--------|----|-------------------------|-------------------------------|
| [K]    | [K] | [K]                     | [K]                          |
| S     | R   | R                       | R                            |
| K     | K   | K                       | K                            |
| R     | L   | L                       | L                            |
| S     | K   | K                       | K                            |
| K     | R   | R                       | R                            |
| R     | L   | L                       | L                            |
| S     | K   | K                       | K                            |

Residues in pairs of fragments that are identical are designated by a vertical line (I) and those that are similar (arginine, lysine, phenylalanine, proline, or histidine) are designated by a dot (·). The subscript denotes the position of the first amino acid in the sequence.
Structure of Rat Ribosomal Protein L7

Fig. 3. A restriction endonuclease map of the cDNA insert in plasmid pL7-2 and a portion of the strategy used to determine the sequence. The upper portion depicts the restriction endonuclease sites in the cDNA insert in pL7-2 that were used to generate oligonucleotide fragments for the determination of the sequence of nucleotides. The radioactive oligonucleotides were isolated by electrophoresis on polyacrylamide gels (12%). The numbers designate the first nucleotide in the restriction enzyme recognition sequence counting from the 5' end of the insert. See figure legend for orientation and the extent of a restriction enzyme that was determined. Since the sequenced determination was always in the 5' to 3' direction the arrows also specify the strand.

Fig. 4. The time course of the release of amino acids from rat ribosomal protein L7 with carboxypeptidase A. In A, the digestion in 0.1 M Tris pH 8.0 of ribosomal protein L7 with carboxypeptidase A at an enzyme to substrate ratio of 1:100. In B, the digestion is of the same amount of L7 but with a combination of carboxypeptidase A and B, the enzyme to substrate ratio is the same.

Fig. 5. Alignment of a positive conserved sequence in rat ribosomal proteins L7 and in E. coli ribosomal proteins S12, S13, L14, and L15. The data of 24 amino acids that are compared begin at residue 186 in L7, at 18 in S12, at 19 in S13, at 79 in L14, and at 8 in L15. Residues that are in the same position in the E. coli proteins and that are identical or related are shown in green type. Residues that are identical or related but not in the E. coli proteins are highlighted.

Fig. 6. The secondary structure of rat ribosomal protein L7 predicted from the sequence of amino acids. Four computer programs were used: Pan, Panagiotopoulos, and Zavergo (1980); Chou, and Fasman (1978); Ng, Krop, and Boedtkjer (1980) designed programs where three of the four programs are in agreement in the prediction of the secondary structure. The symbols for the conformation are: helix, strand, or bend; extended; and coil.