Proc. Nat. Acad. Sci. USA
Vol. 69, No. 5, pp. 1282–1286, May 1972

Location of Ribosomal Protein Binding Sites on 16S Ribosomal RNA

(ribosomal proteins/ribosomal RNA/protein-RNA interaction/RNase digestion/RNA sequence)

ROBERT A. ZIMMERMANN*, AKIRA MUTO*, PETER FELLNER†,
CHANTAL EHRESMANN†, AND CHRISTIANE BRANLANT†

* Département de Biologie Moléculaire, Université de Genève, 30, quai de l'Ecole-de-Médecine, 1211 Geneve 4, Switzerland; and
† Laboratoire de Chimie Biologique, Université Louis Pasteur de Strasbourg, Rue René Descartes, Esplanade, 67 Strasbourg, France

Communicated by J. D. Watson, March 3, 1972

ABSTRACT **The distribution of ribosomal protein binding sites on the 16S ribosomal RNA molecule has been analyzed by limited ribonuclease hydrolysis of RNA-protein complexes, as well as by the interaction of individual proteins with RNA fragments purified from partial enzymatic digests. Of the six 30S subunit proteins known to interact directly with 16S RNA, proteins S4, S8, S15, S20, and, probably, S13 bind within a fragment produced by T_1 RNase (12S RNA) that comprises some 900 nucleotides and covers almost the entire 5′-terminal half of the 16S molecule. A fragment of 500–600 nucleotides (8S RNA) that is contiguous with 12S RNA and arises from the 3′-terminal portion of the 16S molecule is believed to contain the binding site for protein S7. Protein S15 interacts specifically with a sequence of about 135 nucleotides (4S RNA) that derives from 12S RNA after more extensive hydrolysis. Protein S4, but none of the other ribosomal proteins, binds to a 500-nucleotide fragment (9S RNA), generated by pancreatic RNase, that lies at the 5′-terminus of 16S RNA and is completely overlapped by the 12S fragment. A preliminary map of the binding sites is presented.**

Reconstitution of the 30S ribosomal subunit of *Escherichia coli* proceeds by the sequential and cooperative addition of 20 or 21 different proteins to 16S RNA (1). In the initial stages of the assembly sequence, six ribosomal proteins directly interact with independent binding sites on the RNA molecule (1–4). The specific structural features that permit ribosomal RNA to recognize and bind these proteins are not known.

We have undertaken experiments to locate and characterize the protein binding sites on the 16S RNA molecule by investigating the interaction of individual ribosomal proteins with specific fragments of 16S RNA, produced by limited enzymatic hydrolysis. In particular, we have studied the binding of RNA fragments to each of the six ribosomal proteins of the *E. coli* 30S subunit that have been shown to react directly with 16S RNA: S4, S7, S8, S13, S15, and S20, according to the nomenclature of Wittmann *et al.* (5) or P4a, P5, P4b, P10a, P10b, and P14, respectively, in the nomenclature of Nomura and coworkers (see ref. 1). Proteins were labeled with [^{3}H]aminoacids and RNA with [^{14}C]uracil to facilitate identification of protein–RNA complexes. High specific activity ^{32}P-labeled RNA was used to analyze nucleotide sequences.

Large fragments of 16S RNA by limited RNase digestion

Digestion of 16S ribosomal RNA by T_1 RNase is severely curtailed under ionic conditions that favor site-specific binding of ribosomal proteins [0.05 M Tris (pH 7.6)–0.02 M $MgCl_2$–0.35 M KCl]. At an enzyme to substrate ratio of 1:20, two large RNA fragments, sedimenting at about 12 S and 8 S, are produced; together, these components account for most of the original 16S molecule. When the T_1 RNase to RNA ratio is increased to 1:5, hydrolysis is more extensive; fragments of about 10 S and 4 S are generated from the 12S RNA, as demonstrated by nucleotide sequence analysis, and the 8S RNA is degraded to small oligonucleotides. Fragments with the same sedimentation rates result when enzymatic digestion is performed after individual ribosomal proteins have been attached to the 16S RNA (Fig. 1, *dashed curves*).

The action of pancreatic ribonuclease A on 16S RNA is also limited under similar conditions, and the presence of bound protein again affords the RNA no additional protection. At an enzyme to substrate ratio of 1:20, the principal component in the digest sediments at about 11 S; an increase of the ratio to 1:5 results in fragments with apparent sedimentation coefficients of 9 S and 4 S.

The resistance of large RNA segments to RNase digestion in these experiments is attributable to the high concentration of Mg^{++} ions in the incubation buffer; reduction of the Mg^{++} concentration alone renders the 16S RNA more labile to degradation, whereas removal of monovalent cations has no effect on fragment size. Since the RNA chain possesses a high degree of secondary structure and a compact conformation at the Mg^{++} concentrations used (6), RNase attack is most likely limited by decreased accessibility of the 16S molecule to the enzyme.

RNase T_1 hydrolysis of 16S RNA–protein complexes

The distribution of ribosomal protein binding sites on RNA fragments can be directly determined by subjecting complexes between 16S RNA and individual 30S proteins to limited enzymatic cleavage. Thus, after RNase T_1 digestion at an enzyme to substrate ratio of 1:20, almost all of the [^{3}H]S4-protein bound to the 16S RNA before digestion (Fig. 1*a*) cosediments with the 12S component (Fig. 1*b*), while no protein S4 radioactivity is observed in the 8S peak. Under the same conditions, S8, S13, S15, and S20, four other 30S proteins capable of direct interaction with RNA, also cosediment with the 12S, but not the 8S, fragment. These results suggest that the binding sites for these five proteins are located on the 12S RNA. By contrast, protein S7 cosediments primarily with the smaller component (Fig. 1*d*), and it is likely that its binding site lies within the 8S RNA.

After hydrolysis at a T_1 RNase to RNA ratio of 1:5, a high proportion of the protein S4 originally bound to the 16S RNA molecule is retained by the 10S component (Fig. 1*c*). Parallel experiments showed that proteins S8, S13, and S20 also remain bound to the 10S fragment. Protein S15, however,

cosediments mainly with 4S RNA (Fig. 1e). Protein S7 is not associated with either of these RNA fragments.

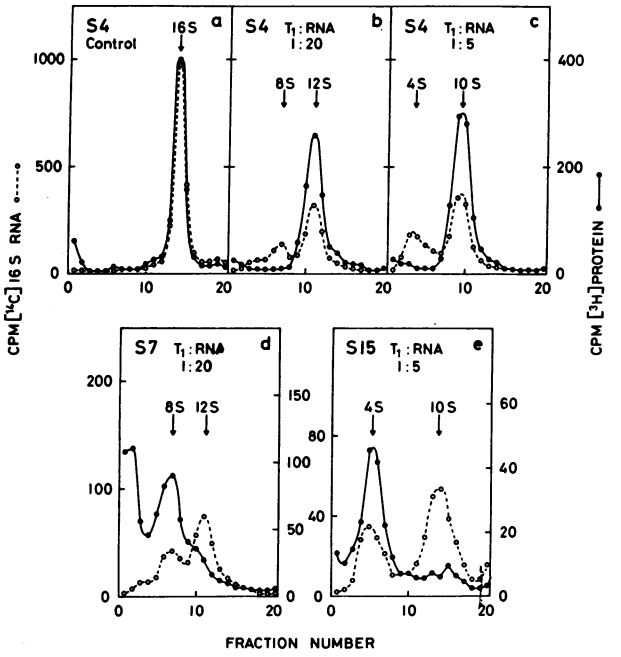

FIG. 1. Hydrolysis of 16S RNA–protein complexes with T_1 RNase. Exponentially growing *E. coli* MRE600 cells were labeled with [^{14}C]uracil or [^{3}H]aminoacid mixture, concentrated, broken open by alumina grinding, and extracted with 0.02 M Tris (pH 7.8)–0.02 M magnesium acetate–0.5 M NH_4Cl. Ribosomes were sedimented from the crude extract, resuspended, and washed by centrifugation through a cushion of 30% sucrose in the above buffer. Ribosomal subunits were separated on 5–20% sucrose gradients in 0.01 M Tris (pH 7.8)–0.3 mM magnesium acetate–0.05 M ammonium acetate. 16S [^{14}C]RNA was prepared from 30S subunits by phenol extraction. ^{32}P-labeled RNA was made in the same way. ^{3}H-labeled 30S subunit proteins were extracted and purified by chromatography on phosphocellulose at pH 6.5, according to Hardy *et al.* (12). Unresolved protein components were fractionated on carboxymethylcellulose at pH 5.6 (13). Proteins were identified by polyacrylamide gel electrophoresis; peak fractions were pooled, concentrated by ultrafiltration, and stored at −20°. Aliquots were dialyzed to remove urea before binding experiments. Unlabeled ribosomal proteins were prepared in the same manner. A complete description of the techniques used is published elsewhere (9).

Binding Assay. 20 μg of 16S [^{14}C]RNA were mixed with 1.2 molar equivalents of dialyzed proteins [^{3}H]S4 (*a*, *b*, *c*), [^{3}H]S7 (*d*), and [^{3}H]S15 (*e*) in 100 μl of TMK buffer [0.05 M Tris (pH 7.6)–0.02 M $MgCl_2$–0.35 M KCl], and incubated for 30 min at 40°. The mixture was chilled on ice, treated with the indicated amount of T_1 RNase, and centrifuged through a 12-ml linear 3–15% sucrose gradient in TMK buffer for 16 hr (*a*, *b*, *c*, *d*) or 24 hr (*e*) at 33,000 rpm. Spinco SW41 and International SB283 rotors were used interchangeably. After the gradients were fractionated, 100 μg of bovine serum albumin and 1 ml of 10% Cl_3CCOOH were added to each tube, and the precipitates were collected on glass-fiber filters. (Small oligonucleotides are not precipitable under these conditions.) Radioactivity was measured in a scintillation counter in a cocktail containing 4 g of Omnifluor per liter of toluene. ^{3}H spillover in the ^{14}C channel was negligible; that of ^{14}C in the ^{3}H channel varied between 10–12%, and was subtracted before data were plotted. Direction of sedimentation is from *left* to *right*.

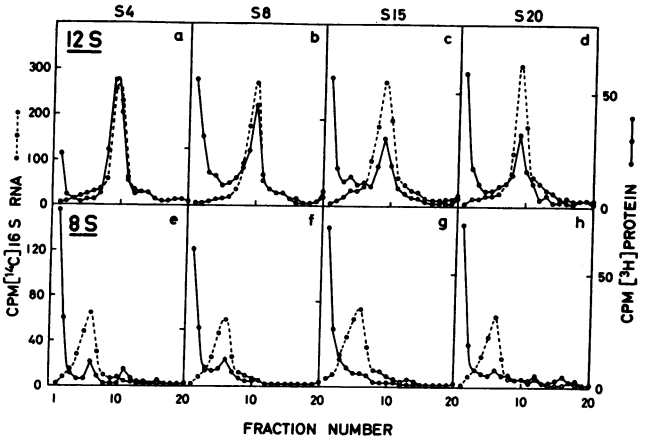

FIG. 2. Interaction of ribosomal proteins with 12S and 8S RNA fragments. ^{14}C-labeled 12S and 8S fragments were isolated by sucrose gradient centrifugation after T_1 RNase digestion of 16S RNA in TMK buffer for 15 min at 0°, at an enzyme to substrate ratio of 1:20. About 1.2 molar equivalents of ^{3}H-labeled S4, S8, S15, and S20 proteins were mixed with about 15 μg of 12S RNA (*a*, *b*, *c*, *d*, respectively) or about 10 μg of 8S RNA (*e*, *f*, *g*, *h*, respectively), and incubated 20 min at 30°. Reaction mixtures were centrifuged through sucrose gradients for 7 hr at 40,000 rpm; analysis of gradients was as described in Fig. 1.

When individual 16S RNA–protein complexes are hydrolyzed with pancreatic RNase at an enzyme to substrate ratio of 1:5, proteins S4 and S15 cosediment with the 9S and 4S components, respectively. Proteins S7, S8, S13, and S20 do not remain bound to either fragment after digestion, but are recovered at the top of the sucrose gradient.

Binding of proteins to isolated RNA fragments

To determine whether isolated 12S and 8S RNA fragments possess intact binding sites, both were prepared from uncomplexed 16S RNA and incubated individually with each of the six ribosomal proteins. Fig. 2 demonstrates that proteins S4, S8, S15, and S20 bind to 12S RNA, but not to 8S RNA. Proteins S7 and S13 interact with neither fragment, a finding that suggests that their binding sites have been modified or destroyed as a result of RNase digestion.

Three proteins that do not bind to 16S RNA, S1, S5, and S19 [P1, P4, and P13, respectively, in the nomenclature of Nomura, (refs. 1 and 5)] were tested for their ability to form complexes with 12S RNA. None of these proteins interacted with the fragment, implying that no new binding sites had been gratuitously generated through cleavage.

Purified 4S and 9S RNA fragments each bound a single ribosomal protein. The 4S component, isolated from T_1 RNase digests of uncomplexed 16S RNA, specifically interacted with protein S15 (Fig. 3*a*). The 9S RNA, produced by pancreatic RNase hydrolysis of uncomplexed ribosomal RNA, bound only protein S4 of the six proteins capable of direct interaction with intact 16S molecules (Fig. 3*b*).

Characterization of the RNA fragments

The fragments described above were isolated from ^{32}P-labeled 16S RNA, and analyzed by the fingerprint technique of Sanger and coworkers (7). Comparison of the oligonucleotides identified in the fingerprints with known sequences in the 16S RNA permitted us to deduce the size, integrity, and

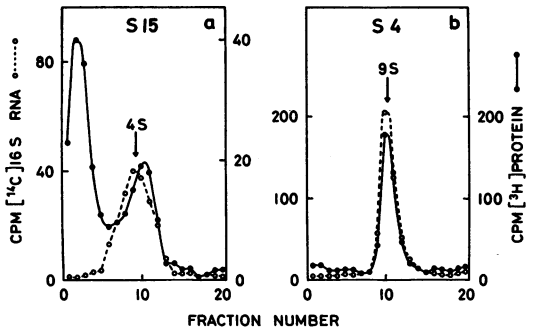

FIG. 3. Specific binding of proteins S15 and S4 to isolated RNA fragments. (*a*) 2.2 μg of 4S RNA, isolated from a digest of ^{14}C-labeled 16S RNA made in TMK buffer at a T_1 RNase to substrate ratio of 1:5, was incubated with 0.9 μg of ^{3}H-labeled protein S15 for 30 min at 0°. The reaction mixture was centrifuged for 24 hr at 40,000 rpm through a sucrose gradient, which was analyzed as in Fig. 1. Using data from three separate experiments, we estimated that between 25 and 50% of the 4S RNA binds protein S15 at a molar protein to RNA ratio of about 1.5. This finding explains the slightly faster sedimentation of the complex (^{3}H-peak) relative to the total RNA (^{14}C-peak), which consists of unbound, as well as bound, fragments. (*b*) 7.5 μg of 9S RNA, after purification from a pancreatic RNase digest of 16S RNA prepared at an enzyme to substrate ratio of 1:5, was mixed with 1.7 μg of ^{3}H-labeled protein S4 and incubated 30 min at 0°. Analysis was as in Fig. 1, except that the sucrose gradient was centrifuged for 16 hr at 33,000 rpm.

relative position of these fragments in the intact molecule.‡ Designation of segments within the 16S RNA by letter name has been made according to Fellner (ref. 8; see also Fig. 6).

The 12S and 8S components consist of about 900 and 500–600 nucleotides, respectively. The 5′-terminus of the 12S fragment lies within Section L, possibly as few as 10–15, and no more than 50, nucleotides from the 5′-terminus of the 16S RNA (see Fig. 6). The 3′-terminus of this component falls near the middle of the molecule, within the following sequence

↓or↓

of Section D, CCUUCGUGCAUCACG, and is contiguous with the 5′-terminus of the 8S fragment. The 3′-end of the 8S RNA falls within Section A, about 50 nucleotides from the 3′-terminus of the 16S RNA molecule. The 12S component could not have suffered any major excisions during preparation, since all the unique oligonucleotide sequences known to occur between its 5′- and 3′-termini are present in roughly equimolar amounts. However, at least two internal oligonucleotides (AUCUG and AAUACCG from Section K) are missing from the 8S RNA.

The 10S fragment contains most of the same oligonucleotides that are present in the 12S RNA, although the absence of several sequences known to lie in the interior of the region covered by the 12S RNA suggests that some internal segments were excised. Fingerprints of the 12S, 8S, and 10S RNAs are published elsewhere (9).

‡ A detailed analysis of 16 RNA primary structure, which provides the basis for deductions about fragment orientation and nucleotide sequence cited in this paper, is to be published in *Biochimie* by P. Fellner, C. Ehresmann & J. -P. Ebel, and by C. Ehresmann, P. Stiegler, P. Fellner & J. -P. Ebel.

The structure of the 4S RNA resulting from RNase T_1 digestion is of particular interest, since this fragment interacts specifically with protein S15. The fingerprint of the 4S RNA (Fig. 4*a*) revealed all the unique sequences of Sections C and C′, which lie near the 3′-end of the 12S RNA (see Fig. 6). The partial sequence of the fragment, which contains some 135 nucleotides, is presented in Fig. 4*b*. Although both 4S and 10S RNAs derive from the 12S fragment, the relationship between them is not clear, since the 10S component also appears to contain some of the unique sequences of Section C.

The 11S product arising from partial pancreatic RNase digestion of the 16S RNA was shown to be very similar to the 12S T_1 fragment by fingerprint analysis, although it also contained the 5′-terminal oligonucleotide of the 16S RNA (9).

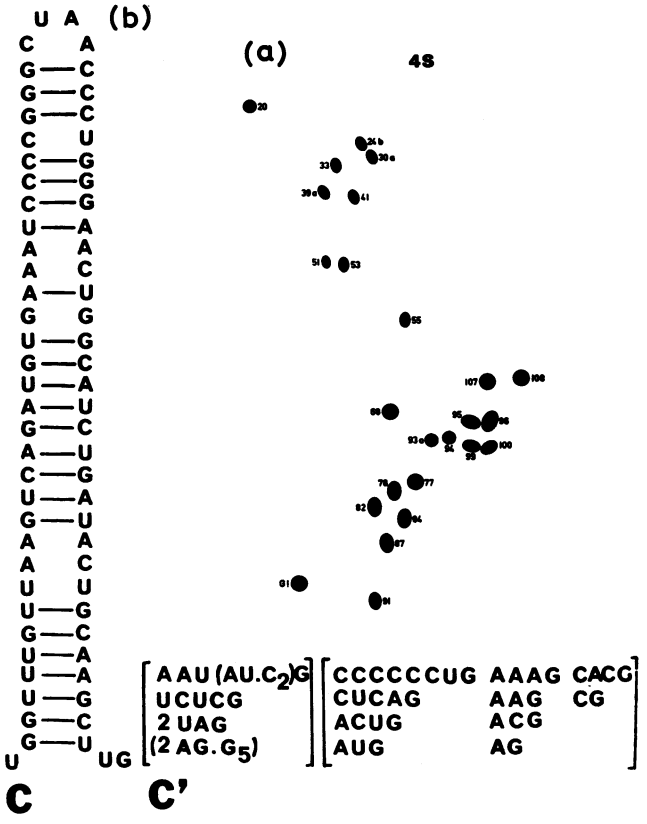

FIG. 4. Partial sequence of the binding site for protein S15. (*a*) Plan of 4S RNA fingerprint. The ^{32}P-labeled 4S fragment, prepared as in Fig. 3, was completely hydrolyzed with T_1 RNase in the presence of bacterial alkaline phosphatase (14). Resulting oligonucleotides were fractionated by two-dimensional electrophoresis on (*i*) cellulose acetate in 5% acetic acid–7 M urea, pH 3.5 (*right* to *left*) and (*ii*) DEAE-paper in 7% formic acid (*top* to *bottom*), according to Sanger *et al.* (7). It was possible to identify all of the oligonucleotides by analysis of both their position in the fingerprint and the products arising from them after digestion with pancreatic RNase; spots are numbered according to the complete catalogue of T_1 RNase oligonucleotides compiled by Fellner *et al.* (15). (*b*) Partial nucleotide sequence of 4S RNA. The major component of the 4S RNA is made up of Sections C (5′-half) and C′ (3′-half). The sequence of Section C is known, and a possible secondary structure, consisting of a loop with an extended, highly base-paired stem, is illustrated. Although the sequence of Section C′ is not yet complete, it may be able to form a similar loop, as judged from its resistance to partial hydrolysis with T_1 RNase.

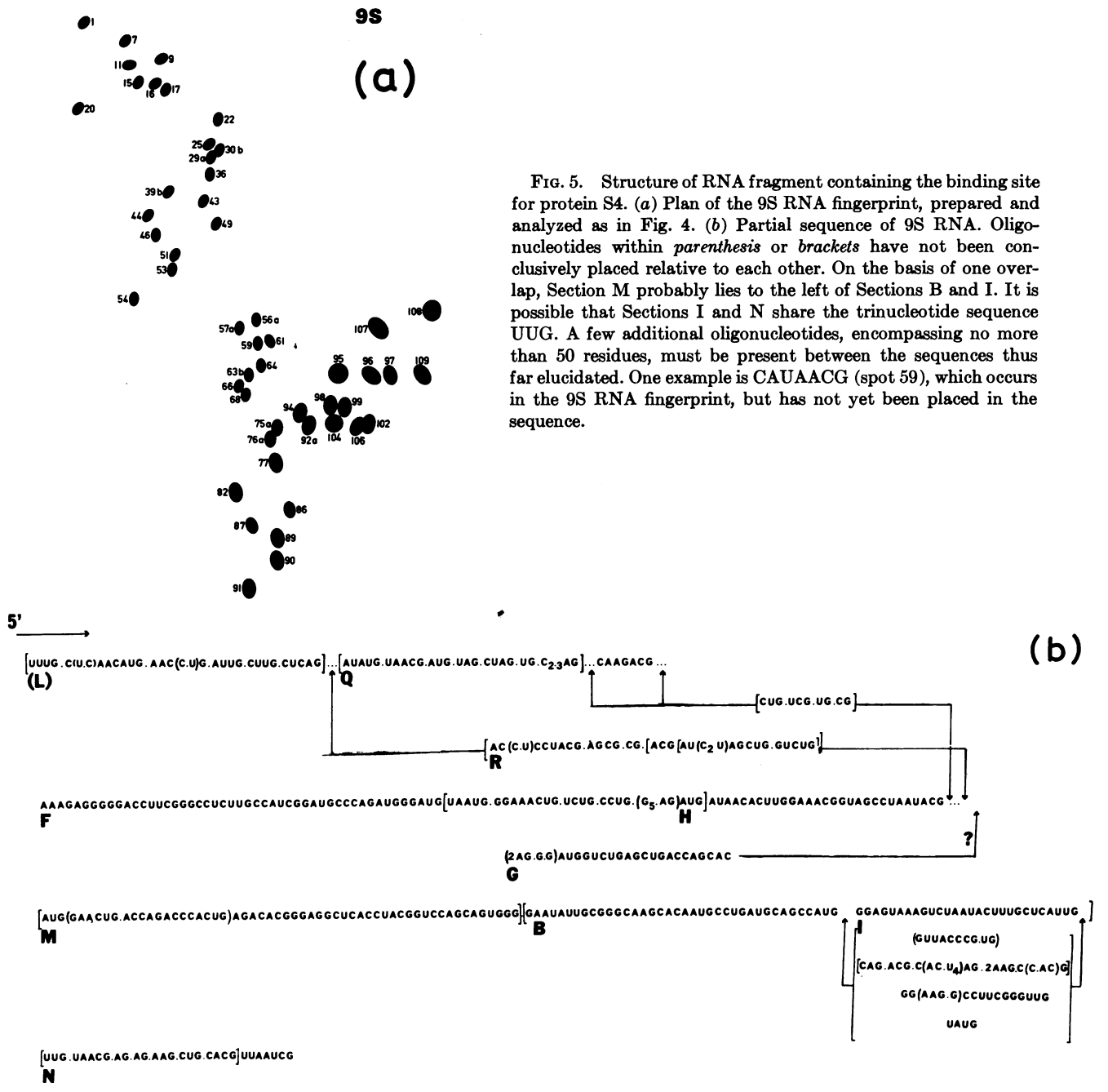

FIG. 5. Structure of RNA fragment containing the binding site for protein S4. (*a*) Plan of the 9S RNA fingerprint, prepared and analyzed as in Fig. 4. (*b*) Partial sequence of 9S RNA. Oligonucleotides within *parenthesis* or *brackets* have not been conclusively placed relative to each other. On the basis of one overlap, Section M probably lies to the left of Sections B and I. It is possible that Sections I and N share the trinucleotide sequence UUG. A few additional oligonucleotides, encompassing no more than 50 residues, must be present between the sequences thus far elucidated. One example is CAUAACG (spot 59), which occurs in the 9S RNA fingerprint, but has not yet been placed in the sequence.

Analysis of the 4S RNA, produced by digestion with the same enzyme, is not yet complete.

The 9S RNA, which specifically binds protein S4, is estimated to contain about 500 residues from an analysis of the oligonucleotides present in its fingerprint (Fig. 5*a*). This material originates within Section L at a point no more than 50 nucleotides from the 5′-terminus of the 16S RNA, and includes Sections F, H, M, B, and I, as well as intervening segments (Figs. 5*b* and 6). The 3′-terminus of the 9S RNA occurs after the sequence UUAAUCG in Section N. The absence of several internal oligonucleotides, notably CUCACCUACG and UAACG, indicates that a number of small excisions occurred during pancreatic RNase digestion. In accord with this observation, the 9S RNA was found to separate into several smaller fragments when electrophoresed in a polyacrylamide gel containing 8 M urea. An independent investigation of RNA fragments that bind to protein S4 is consistent with our findings (10).

Location and nature of the ribosomal protein binding sites

Our present knowledge about the distribution of ribosomal protein binding sites on the 16S RNA molecule is summarized in Fig. 6. Each of the proteins S4, S8, S15, and S20 can interact with the 900-nucleotide 12S fragment that covers almost the entire 5′-half of the RNA molecule. The binding site for protein S4 is located within a sequence of 500 nucleotides that lies at the 5′-end of the 12S RNA, and extends from Section L to Section N. The site for protein S15 is within Sections C and C′, near the 3′-end of the 12S fragment. Protein S13 may also bind within the 5′-half, since it sediments with

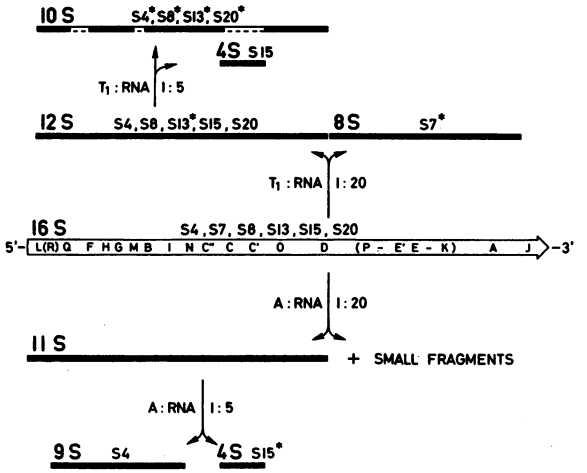

FIG. 6. Preliminary map of protein binding sites on 16S RNA. The scheme shows relative placement of RNA fragments within the 16S molecule, as well as the proteins that bind to each fragment. An *asterisk* indicates that the presence of binding site was inferred from retention of protein by a fragment after degradation of the complex between that protein and 16S RNA; in all other cases, binding was confirmed by interaction of proteins with isolated RNA fragments. Binding of protein to 11S RNA was not studied systematically, because of the similarity of this RNA to the 12S fragment. Segments within 16S RNA are denoted by upper case letters; *parentheses* indicate that relative placement of a section is uncertain. The *arrow* on 16S RNA signifies its direction of transcription. T_1 and *A* stand for T_1 and pancreatic RNases, respectively. *Dashed lines* in 10S RNA refer to probable excision of internal sequences, although secondary and tertiary bonds apparently prevent its disintegration.

the 12S RNA after degradation of its complex with 16S RNA; it does not, however, bind to the isolated 12S fragment. Only S7, of all the proteins tested, probably binds within the 3'-half of 16S RNA; it remains attached to the 8S fragment after digestion of the protein S7–16S RNA complex, but we have not yet succeeded in binding the protein to purified 8S RNA.

Since the proteins that bind directly to 16S RNA initiate 30S subunit assembly, the apparent clustering of their binding sites in the 5'-terminal portion of the RNA may be of physiological significance. The RNA molecule is transcribed in the sense 5' → 3', and it is possible that these proteins attach to nascent RNA before its synthesis is complete. Although this is not a necessary condition for the reconstitution of functional 30S particles (11), binding of certain proteins to the RNA before its release from the DNA template might be advantageous for subunit assembly in the cell.

The ability of RNA fragments to interact with ribosomal proteins shows that their binding sites are not extensively modified by partial hydrolysis of the 16S RNA molecule. Thus, the small breaks and excisions that were noted in some of the fragments do not appear to impair their ability to bind at least certain of the proteins, presumably because the overall conformation of the nucleotide chain is maintained by noncovalent bonds. Such small defects could, nonetheless, account for the failure of protein S7 to react with the 8S fragment, and of any protein but S4 to form a complex with the 9S RNA.

Binding sites for individual ribosomal proteins could arise either from the local conformation of short nucleotide sequences or from the interaction of two or more sequences maintained in a proper configuration by the three-dimensional folding of the RNA chain. Although a choice between these alternatives cannot yet be made, we can point out the salient structural features of the RNA fragments that specifically bind two of the 30S proteins. We propose that the 4S RNA, which contains the binding site for protein S15, consists of two loops with long, extensively hydrogen-bonded, stems, linked by a short nonhelical region. The postulated structure of one of the loops, comprising Section C, is presented in Fig. 4. Section C', whose sequence is only partially complete, probably exists as a similar loop. Known sequences of the 9S RNA (Fig. 5), which specifically interacts with protein S4, also appear to be capable of forming loops with more or less elongated stems (8). If they were about the same size as Section C, six or seven such loops could be accommodated within this fragment. The large amount of RNA available for binding a single protein in this case suggests the possibility that the protein interacts with more than one portion of the RNA chain. In order to further elucidate which aspects of RNA structure are important in RNA–protein interaction, we are currently testing the binding capacity of small RNA segments produced by partial fragmentation of sequences already shown to interact with individual proteins.

The interest and encouragement of Drs. A. Tissières and J.-P. Ebel, as well as the skilled technical assistance of Miss Jeanette van Dillewijn and Mrs. Barbara Seebeck, are acknowledged with appreciation. R. A. Z. thanks the Helen Hay Whitney Foundation for support. This work was done under grants from the Fonds National Suisse de la Recherche Scientifique, the Centre National de la Recherche Scientifique, and the Délégation à la Recherche Scientifique et Technique.

1. Mizushima, S. & Nomura, M. (1970) *Nature* **226**, 1214–1218.
2. Schaup, H. W., Green, M. & Kurland, C. G. (1970) *Mol. Gen. Genet.* **109**, 193–205. Schaup, H. W., Green, M. & Kurland, C. G. (1971) *Mol. Gen. Genet.* **112**, 1–8.
3. Nashimoto, H., Held, W. Kaltschmidt, E. & Nomura, M. (1971) *J. Mol. Biol.*, **62**, 121–138.
4. Garrett, R. A., Rak, K. H., Daya, L. & Stöffler, G. (1971) *Mol. Gen. Genet.* **114**, 112–124.
5. Wittmann, H. G., Stöffler, G., Hindenach, I., Kurland, C. G., Randall-Hazelbauer, L., Birge, E. A., Nomura, M., Kaltschmidt, E., Mizushima, S., Traut, R. R. & Bickle, T. A. (1971) *Mol. Gen. Genet.* **111**, 327–333.
6. Boedtker, H. (1960) *J. Mol. Biol.* **2**, 171–188.
7. Sanger, F., Brownlee, G. G. & Barrell, B. G. (1965) *J. Mol. Biol.* **13**, 373–398.
8. Fellner, P. (1971) *Biochimie* **53**, 573–583.
9. Zimmermann, R. A., Muto, A., Fellner, P. & Branlant, C. (1972) in *Functional Units of Protein Biosynthesis*, Symp. 7th FEBS Meeting Varna 1971 (Academic Press, London), in press.
10. Schaup, H. W., Sogin, M., Woese, C. & Kurland, C. G. (1971) *Mol. Gen. Genet.* **114**, 1–8.
11. Traub, P. & Nomura, M. (1968) *Proc. Nat. Acad. Sci. USA* **59**, 777–784.
12. Hardy, S. J. S., Kurland, C. G., Voynow, P. & Mora, G. (1969) *Biochemistry* **8**, 2906–2915.
13. Otaka, E., Itoh, T. & Osawa, S. (1968) *J. Mol. Biol.* **33**, 93–107.
14. Brownlee, G. G. & Sanger, F. (1967), *J. Mol. Biol.* **23**, 337–353.
15. Fellner, P., Ehresmann, C. & Ebel, J.-P. (1970) *Nature* **225**, 26–29.