
Location and characteristics of ribosomal protein binding sites in the 16S RNA of *Escherichia coli*

R.A.Zimmermann¹, G.A.Mackie, A.Muto², R.A.Garrett*, E.Ungewickell*, C.Ehresmann¹, P.Stiegler¹, J.-P.Ebel and P.Fellner^{1,3}

Département de Biologie Moléculaire, Université de Genève, 30, Quai de l'Ecole de Médecine, 1211 Geneve 4, Switzerland

Received 13 January 1975

ABSTRACT

Specific binding sites for five proteins of the *Escherichia coli* 30S ribosomal subunit have been located within the 16S RNA. The sites are structurally diverse and range in size from 40 to 500 nucleotides; their functional integrity appears to depend upon both the secondary structure and conformation of the RNA molecule. Evidence is presented which indicates that additional proteins interact with the RNA at later stages of subunit assembly.

INTRODUCTION

Interactions between ribosomal proteins and ribosomal RNA are essential to the proper physical assembly of the ribosome and play a critical role in maintaining the mature particle in its biologically active configuration. It is now generally agreed that of the 21 proteins present in the 30S ribosomal subunit of *Escherichia coli*, S4, S7, S8, S15, S17 and S20 independently associate with specific binding sites in the 16S RNA¹⁻⁷. Other proteins of the 30S subunit may be able to interact directly, although not independently, with the ribosomal RNA as well. Investigation of the size, distribution and nucleotide sequence of the sites in the RNA to which ribosomal proteins bind is thus of immediate relevance to the understanding of ribosome assembly and architecture.

It has been reported previously that the binding sites for proteins S4, S8, S15 and S20 are located within a segment of 900

nucleotides that encompasses the entire 5'-half of the 16S RNA⁵. Moreover, protein S4 was found to associate with a sequence of 500 nucleotides situated at the 5'-end of this segment^{5,8}, whereas protein S15 was able to bind to a sequence of 150 nucleotides that begins some 600 residues from its 5'-terminus⁵. By contrast, the attachment site for protein S7 was determined to lie within the 600 3'-terminal nucleotides of the 16S RNA molecule⁵.

As a result of recent progress in the nucleotide sequence analysis of the 16S RNA⁹ and in the isolation of specific RNA fragments, we are now able to define with greater precision the location and structural characteristics of the protein binding sites within the 16S RNA molecule. Two different experimental approaches have been used in the present studies. In the first or sedimentation method, free 16S RNA as well as protein-16S RNA complexes were subjected to partial RNase hydrolysis, and the digestion products were separated by sucrose gradient centrifugation⁵. The nucleotide sequences of the RNA fragments obtained were analyzed by fingerprinting⁶ and fragments of well-defined structure were systematically tested for their ability to interact with each of the RNA-binding proteins¹⁰. Complex formation indicated that the RNA fragment contained a binding site for the protein in question and the interaction was judged to be specific when the molar ratio of protein and RNA in the complex attained a saturation value of 1:1 or less in the presence of excess protein. In the second or electrophoretic method, polyacrylamide gels were used to isolate ribonucleoprotein particles from limited RNase digests of protein-16S RNA complexes^{11,12}. The RNA components of these particles, which had been protected from hydrolysis by the attached protein, were then dissociated in urea and SDS and further purified by

a second cycle of gel electrophoresis in the presence of urea. Individual RNA fragments were eluted and their nucleotide sequences identified by fingerprint analysis^{11,12}. In order to test the capacity of these fragments to reassociate with ribosomal proteins, the ribonucleoprotein obtained after the first electrophoretic step was extracted with phenol and the RNA was recovered by ethanol precipitation. The RNA was then mixed with a fresh sample of the protein and complexes were detected either by retention on nitrocellulose filters¹¹ or by sucrose gradient centrifugation¹². The demonstration of RNA-protein interaction under these conditions provides strong evidence that the protected RNA comprised part or all of the corresponding protein binding site in the 16S RNA molecule.

The two methods described above were used jointly to define the binding sites of each of the RNA-binding proteins of the 30S subunit. We shall now consider in detail the data which led to these assignments, progressing from the 5'-terminus to the 3'-terminus of the 16S RNA. The nomenclature of sections of the RNA molecule follows that utilized in the preceding paper⁹ (see also refs. 13 and 14).

Location of binding site for protein S4

RNA fragments specific for protein S4 have been isolated by both the sedimentation and electrophoretic techniques. Structures of the fragments are illustrated schematically in Fig.1. By means of the first method, a fragment sedimenting at 9S was isolated following mild digestion of the 16S RNA, or of the S4-16S RNA complex, with pancreatic RNase. The 9S RNA thus prepared has been shown to specifically interact with protein S4, forming a complex which contains one mole protein per mole RNA^{5,10}. This fragment encompasses approximately 500 nucleotides

Figure 1. Scheme delimiting binding sites for ribosomal proteins S4 and S20 in the 16S RNA molecule. Segments of the 16S RNA that specifically associate with proteins S4 and S20 are enclosed by dashed (----) and solid (—) lines, respectively. The nucleotide sequence and proposed secondary structure were established as described in the preceding paper⁹. Dotted lines denote portions of the RNA in which the primary structure is not entirely known, although the oligonucleotide composition and total number of residues present have been determined in all cases. Fragments specific for S4 generally contained the oligonucleotide UAAAUCG from section C'' in molar amounts; their 3'-termini could not be more accurately defined, however, since the sequence of section C'' has not yet been completed. It is of interest that these fragments were generally characterized by a low yield of oligonucleotides from sections Q through B. The arrow indicates the discontinuity between RNA fragments 12S-I and 12S-II which, after electrophoretic fractionation, can simultaneously interact with protein S4 (see Fig. 2 and text). Note that the scission occurs within the unpaired loop of section G in the labile region of the S4-specific segment. Fragments 12S-I and 12S-II both represent continuous sequences and contain no "hidden breaks". Although a limited number of scissions reproducibly occurred in the RNA chain during production of other fragments specific for both S4 and S20, their locations are not shown in the diagram for the sake of simplicity. The cuts were apparently masked by the RNA secondary structure under the conditions used for their isolation and subfragments were not released; the subfragments could, however, be separated by polyacrylamide gel electrophoresis under denaturing conditions (see Fig. 3).

and extends from the sequence CUCAG in section L to the middle of section C" (Fig. 1). A virtually identical sequence was recovered from a ribonucleoprotein particle produced by limited hydrolysis of the S4-16S RNA complex with RNase T₁ and isolated by polyacrylamide gel electrophoresis. Moreover, the RNA segment prepared under these conditions was demonstrated to reassociate specifically with protein S4 as well as with protein S20 by the membrane filter assay¹¹. Finally, the RNA moiety of a protein S4-RNA fragment complex reported earlier¹⁵ appears to be similar in size and oligonucleotide composition to the first two fragments described above. The fact that all of the S4-specific RNAs have been resolved into a number of subfragments by polyacrylamide gel electrophoresis in the presence of urea demonstrates that "hidden breaks" occur during RNase digestion.

Application of the electrophoretic method to the separation of products from a pancreatic RNase digest of the protein S4-16S RNA complex yielded a ribonucleoprotein fragment whose RNA component spanned the same portion of the molecule as those described above¹². However, all of the unique T₁ and pancreatic oligonucleotides in a 120-residue sequence extending from the 5'-end of section Q to the 3'-end of section M were absent or present only in very small amounts (see Figs. 1 and 6). Since this RNA retains the ability to interact with S4, but no other protein, it is possible that sections Q, R, G and M do not significantly contribute to the binding site for protein S4.

Limited hydrolysis of the 16S RNA with RNase T₁ has been found to yield a fragment of about 12S which, after purification, was able to specifically bind proteins S4, S8, S15 and S20^{5,10}. The 12S RNA extends from near the 5'-terminus of the 16S RNA to section O and it contains about 900 residues. Since this

fragment contained several "hidden breaks", subfragments were purified from it by gel electrophoresis under denaturing conditions (Fig. 2a) and tested for their ability to interact with each of the RNA-binding proteins. Two of the subfragments, 12S-I and 12S-11, were found to simultaneously form a complex with S4, demonstrating that this protein is capable of interaction with at least two different sequences in the 16S RNA molecule¹⁶. Oligonucleotide maps of these fragments were prepared by fingerprinting and are presented in Fig. 2. It was determined that subfragments 12S-I and 12S-II encompassed sections L through G, and sections M through I", respectively.

Location of binding site for protein S20

Pancreatic RNase digestion of a complex resulting from the incubation of 16S RNA with protein S20 in the presence of proteins S16 and S17 produced a ribonucleoprotein particle containing S20 and one or both of the other proteins¹⁶. A fingerprint of the RNA extracted from the ribonucleoprotein appeared to be identical to that of the 9S RNA described above (see Fig. 1). Moreover, the fragment was found to reassociate specifically with S20 at a molar protein:RNA ratio of 1:1 when the S16 + S17 mixture was also present in the incubation mixture. The present experiments therefore show that the binding sites for protein S20 and, most likely, for protein S17 as well, are located within the same region of the 16S RNA as the binding site for protein S4. Protein S17 has been reported to interact independently with the 16S RNA⁷, whereas the association of S16 was found to require the presence of both S4 and S20^{7,17}.

RNA sequences that interact with protein S20 were more precisely localized by the electrophoretic method. A ribonucleoprotein fragment was produced from the S20-16S RNA complex by limited digestion with RNase T₁ as shown in Fig. 3. The RNA

Figure 2. Electrophoretic fractionation and fingerprints of RNA fragments 12S-I and 12S-II. (a) ^{32}P -labelled 12S RNA from the 5'-half of the 16S RNA was prepared from a partial T_1 RNase digest of [^{32}P]-16S RNA at an enzyme:substrate ratio of 1:70 in a buffer containing 0.05 M Tris-HCl, pH 7.6., 0.02 M MgCl_2 , 0.35 M KCl, and 0.006 M β -mercaptoethanol (TMK buffer) as described by Muto *et al.*⁶. About 600 μg of the 12S RNA were applied to a 10-15% compound polyacrylamide slab gel containing 6 M urea and separated into subfragments by electrophoresis using the technique of Ehresmann *et al.*¹⁴. Migration was from top to bottom. (b) and (c). The RNA fragments in bands 12S-I and 12S-II were eluted from the gel by diffusion¹⁶, concentrated by ethanol precipitation, and aliquots fingerprinted by two-dimensional electrophoresis²⁶ after complete digestion with T_1 RNase in the presence of alkaline phosphatase²⁷. 1. First dimension: cellulose acetate in 5% acetic acid - 7 M urea, pH 3.5. 2. Second dimension: DEAE-paper in 7% formic acid. Oligonucleotides were identified by their positions and by the products obtained from them upon digestion with pancreatic RNase. Fragments 12S-I extends from the sequence UUUG in section L to the sequence UCUG in section G. Fragment 12S-II contains the sequence ACCAG at the beginning of section M and continues through section I", but does not extend into section C". In the examples shown, fragment 12S-I was 90% pure whereas fragment 12S-II was about 70% pure. Contamination of the latter fragment arose mainly from sections C and C". Asterisks denote heterogeneities. See also Fig. 1.

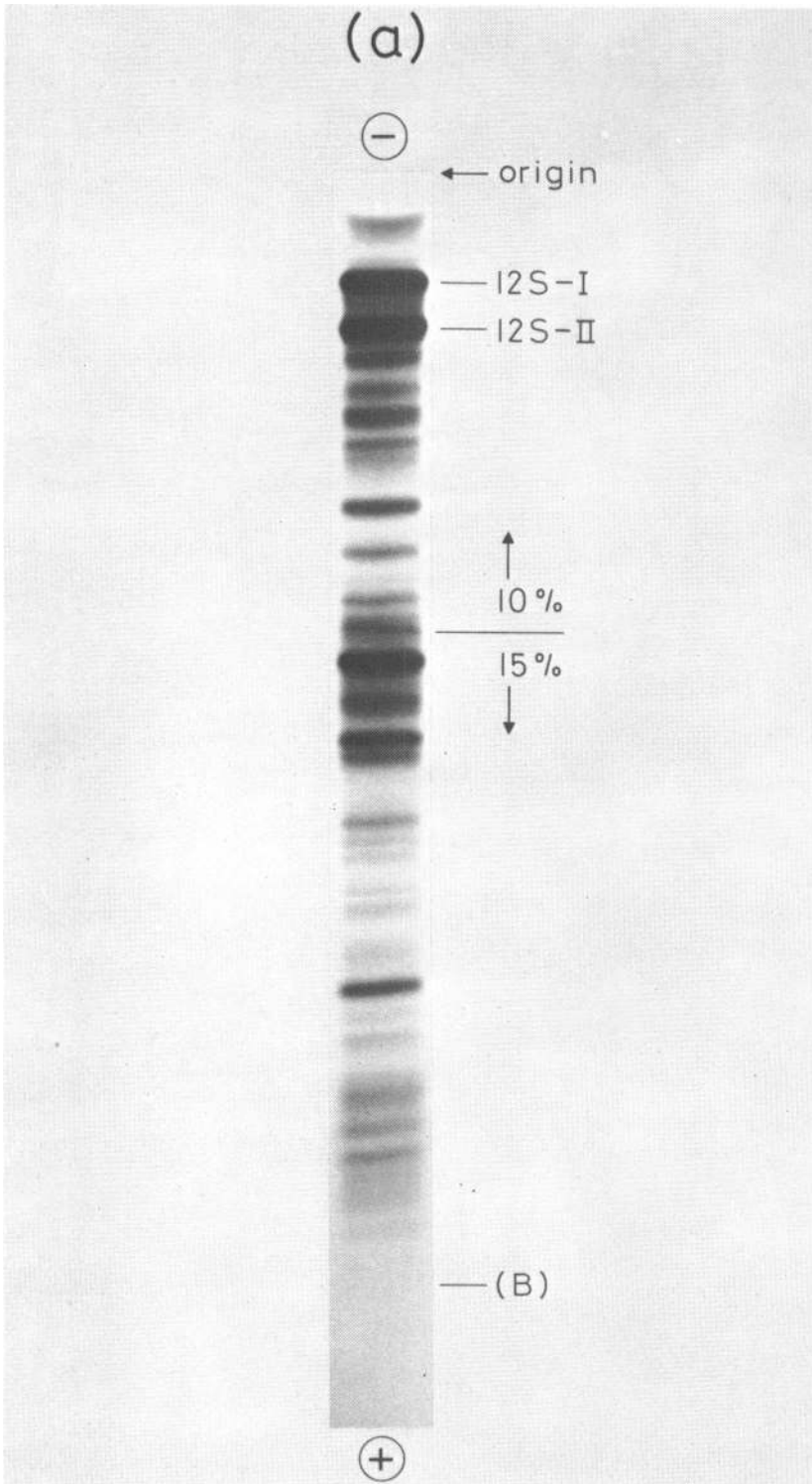


Fig. 2

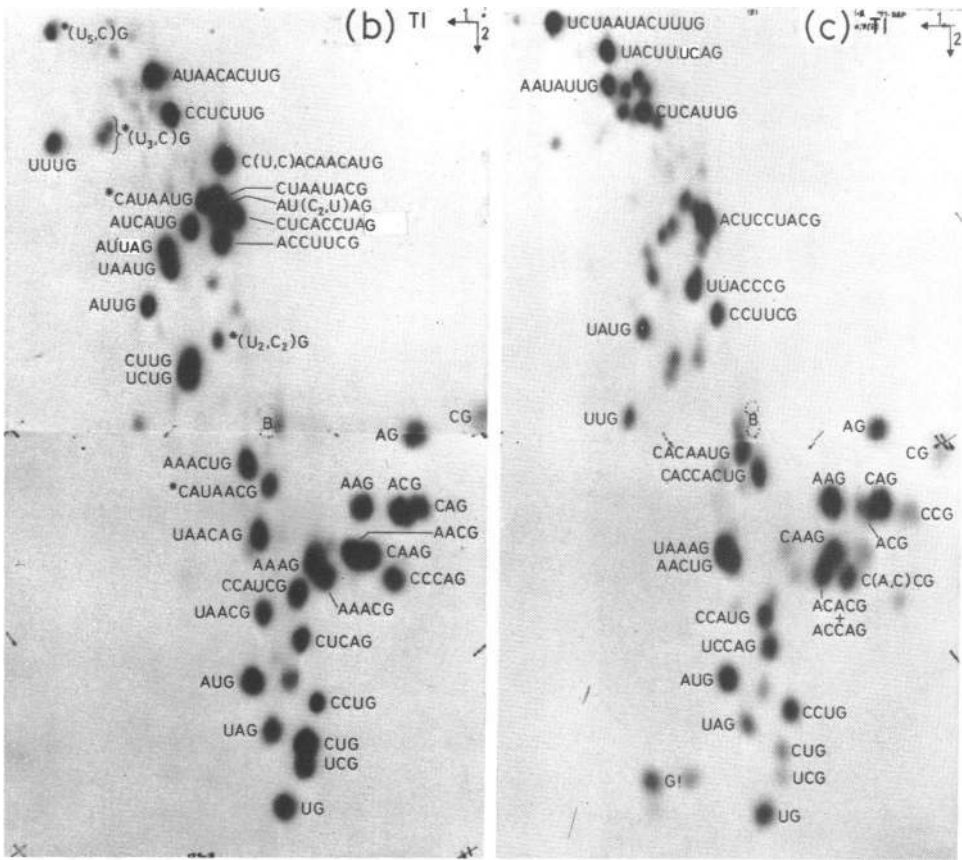


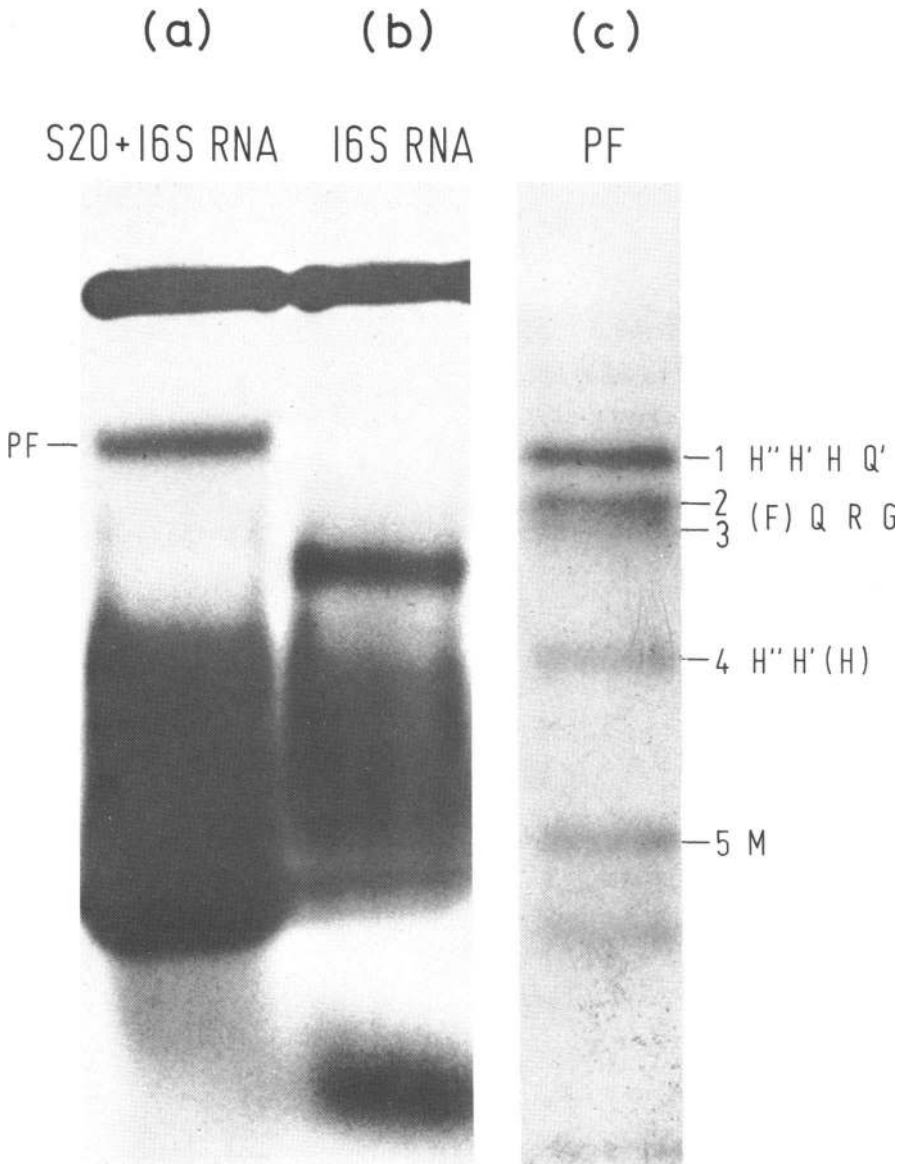
Fig. 2

associated with S20 was separated into approximately six subfragments by gel electrophoresis in the presence of urea (Fig. 3c) and each of them was subsequently analyzed by fingerprinting¹¹. The subfragments were found to comprise a colinear set from a 270-nucleotide sequence extending from section H'' to near the 3'end of section M (Fig. 1). In addition, after deproteinization of the initial ribonucleoprotein fragment, the resulting RNA was specifically retained by protein S20 on nitrocellulose membrane filters. The RNA segment protected by S20 represents the middle portion of the S4-specific region and is shorter than it by about 60 residues at the 5'-end and 170 residues at the 3'-end.

Location of binding sites for proteins S15 and S8

The binding sites for proteins S15 and S8 in the 16S RNA are intimately related to one another as the scheme in Fig. 4 shows and they will therefore be discussed together. The position of the binding site for protein S15 was initially inferred from its ability to interact with a 4S RNA fragment containing sections C, C₁', and C₂' which was isolated by the sedimentation technique from T₁ RNase digests of 16S RNA⁵. A similar, partially-protected RNA fragment of 5S arises upon hydrolysis of the S15-16S RNA complex with pancreatic RNase and, if S8 is present as well in the original ribonucleoprotein, it is recovered in the 5S fragment complex together with S15 after RNase digestion¹⁸. The 5S RNA, which encompasses 150 nucleotides, has been freed of bound protein by phenol extraction and purified on sucrose gradients. In this state, it can reassociate specifically with freshly-added protein S15 with a binding stoichiometry of 0.75 moles protein per mole RNA¹⁸. Moreover, purified 5S RNA can select and bind proteins S8 and S15 from an unfractionated mixture of 30S subunit proteins¹⁸. It is thus

Figure 3. Isolation of RNA fragments protected by protein S20. (a). The complex of S20 with ^{32}P -labelled 16S RNA was digested with T_1 RNase at an enzyme:substrate ratio of 1:20 at 0°C in TMK buffer (see Fig. 1). The products were separated by electrophoresis on an 8% polyacrylamide gel in 0.005 M Tris-acetate, pH 8.0, and 0.005 M Mg acetate¹¹. Migration was from top to bottom. RNA bands were located by autoradiography. PF marks the band containing protein S20 and associated RNA fragments. (b) Same as (a) except that the protein was omitted. Note that there is no band at the position of PF. (c) Band PF was excised from the gel shown in (a), soaked in 7 M urea to dissociate the RNA-protein complex, and polymerized into a second gel consisting of 12% polyacrylamide and 7M urea according to Branlant *et al.*²². The layer surrounding the sample contained in addition 0.1% SDS to facilitate separation of protein and RNA. The RNA fragments were then fractionated by electrophoresis into several bands, and the nucleotide sequences in each were identified by fingerprinting¹¹. Letters to the right of the five largest bands indicate the sections of the 16S RNA that were found to be present (see Fig. 1). In one experiment, the protection of 28 nucleotides from section L, at the 5'-end of the 16S RNA, was also noted.



apparent that the binding sites for both of these proteins lie within section C, C₁' and C₂' (Fig. 4). Section C consists of a long hairpin loop with an extensively hydrogen-bonded stem; sections C₁' and C₂' are believed to comprise the base-paired stem of an adjacent loop⁹.

The electrophoretic method has been used to provide further information on the location of the binding sites for S15 and S8. A ribonucleoprotein particle produced by pancreatic RNase digestion of the S15-16S RNA complex has been isolated on polyacrylamide gels¹¹. The RNA sequences associated with S15 arise from sections C, C₁' and C₂' as in the previous experiment, although only the lower portions of the two putative hairpin loops are protected in this case (Fig. 4). After phenol extraction, the protected RNA fragments were shown to interact with protein S15 by the membrane filter assay¹¹.

Electrophoretic fractionation of products arising from limited pancreatic RNase digestion of the S8-16S RNA complex yielded a small ribonucleoprotein particle¹¹. The RNA component of the particle migrated as a single, relatively sharp band in a polyacrylamide gel containing urea. A fingerprint of the purified fragment, presented in Fig. 5a, revealed the presence of oligonucleotides corresponding to about 40 residues of section C. Since the sequences are located in the lower part of the long and largely base-paired hairpin loop of that section (Fig. 4), it was concluded that the protected RNA consisted of two complementary fragments of identical size which migrated together in the urea gel. Reassociation of these fragments with protein S8 has been demonstrated¹¹. Evidence for the protection of a similar portion of section C by protein S8 has also been obtained by an independent method¹⁹.

Figure 4. Primary and secondary structures of nucleotide sequences within the 16S RNA protected from nuclease attack by proteins S8, S15, S6 and S18. RNA sequences enclosed within the various boxes were wholly or partially protected from pancreatic RNase digestion by proteins S8 (—), S15 (-----), S8 + S15 (.....) or S8 + S15 + S6 + S18 (— - —). Fingerprints of two of these fragments are presented in Fig. 5 and their structures are discussed in detail elsewhere^{11,18,20}. Whereas internal scissions within the two complementary sequences protected by S8 were minimal (see Fig. 5), a small number of "hidden breaks" were detected in each of the other fragments. For clarity, their positions have not been shown except for a major cut in section D' (arrow). Although not indicated, protein S15 confers weak protection upon the sequence from the left side of section C which is strongly protected by protein S8. The base-pairing among sections C₁, K' and C₂' must remain hypothetical since it has been predicated upon incomplete sequence data; however, it is consistent with all analytical information presently available. Dotted lines designate portions of the nucleotide sequence in which the oligonucleotides present are known, but not conclusively ordered. In addition, the overlaps between sections C₁ and K', and between sections C₂' and D', have not yet been formally established. Since the RNA fragments delineated in the scheme were found to reassociate specifically with the proteins which protect them from RNase attack, it is believed that they comprise part or all of the respective protein binding sites.

_____ **S8**

----- **S15**

..... **S8, S15**

----- **S8, S15, S6, S18**

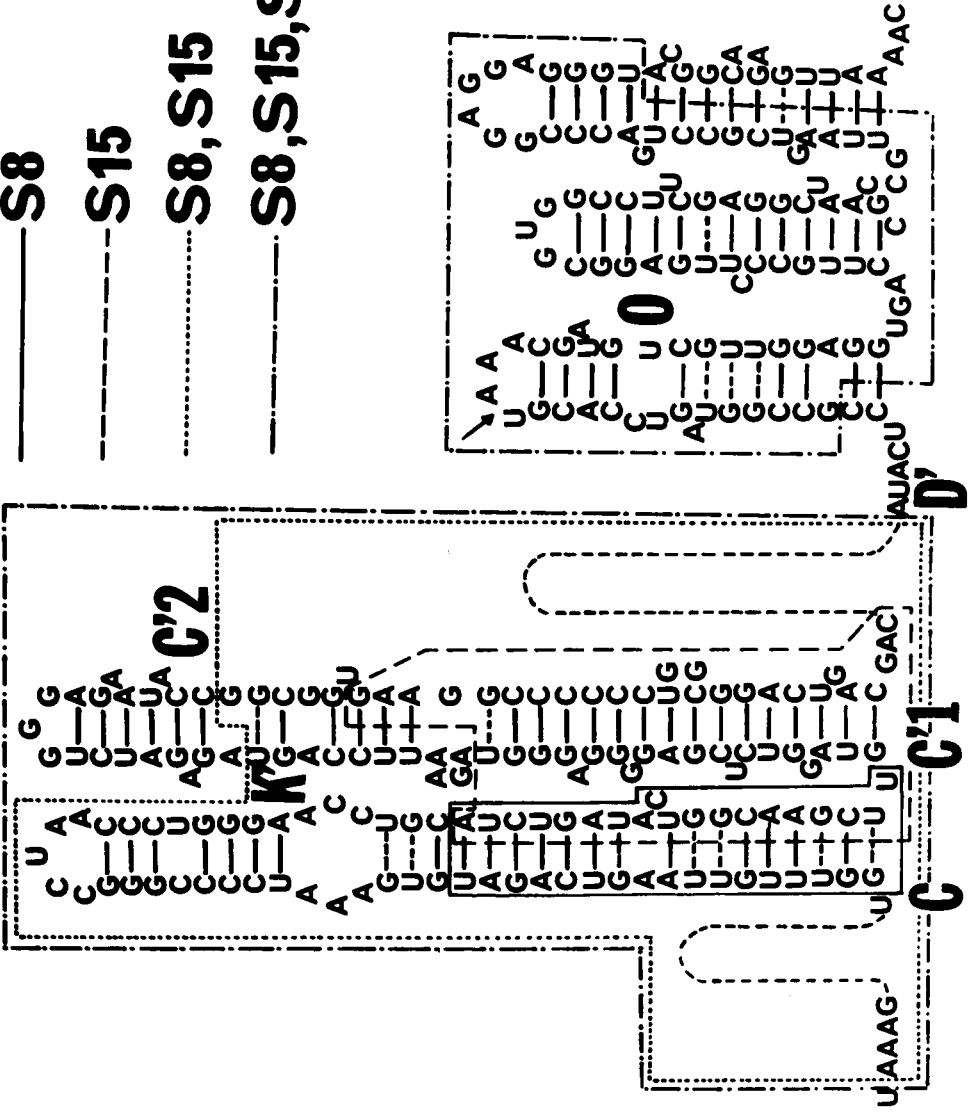
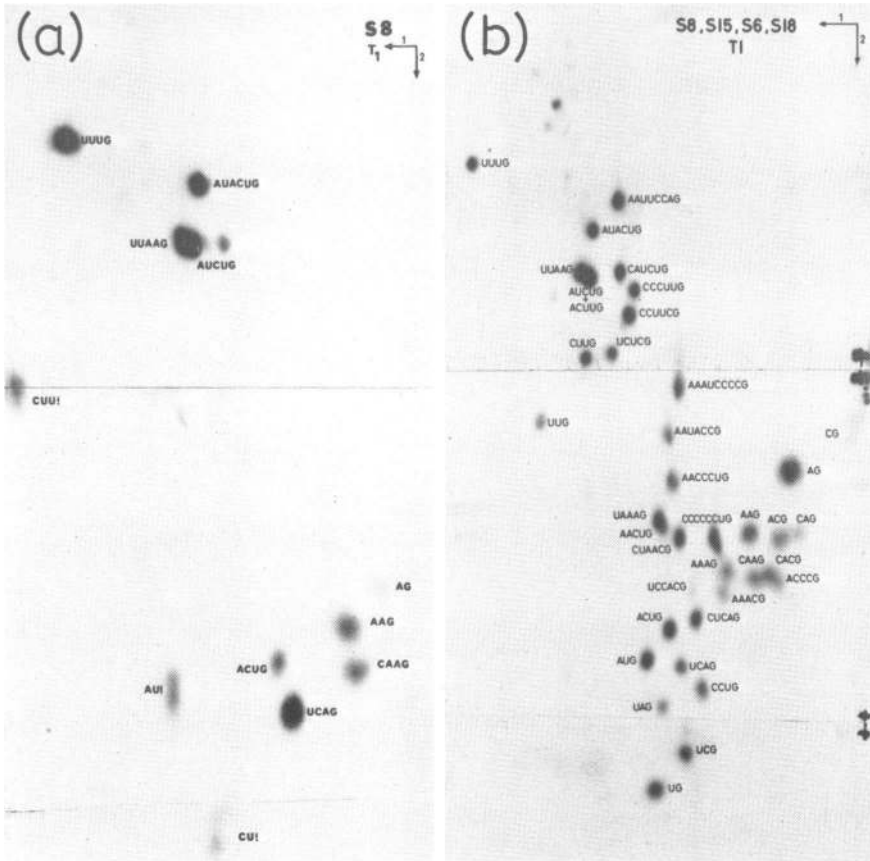


Figure 5. Fingerprints of RNA fragments specific to protein S8 and to proteins S8, S15, S6 and S18. (a) Fragment recovered following pancreatic RNase digestion of the complex between S8 and ^{32}P -labelled 16S RNA at an enzyme:substrate ratio of 1:30 in TMK buffer at 0°C . The protected RNA sequences were isolated by the two-step electrophoretic technique described in the legend to Fig. 3 except that the first electrophoresis buffer contained 0.03 M Tris-acetate, pH 7.8, and 0.0005 M Mg acetate¹¹. After complete T_1 RNase digestion, the fragment was fingerprinted by the method of Sanger and co-workers^{26,27} as outlined in the legend to Fig. 2. Two of the T_1 oligonucleotides, AUACUG and CAAG, were partially but reproducibly cut by pancreatic RNase during fragment preparation to yield the products AU + ACUG and C + AAG, respectively, which can be seen in the fingerprint. Although such cuts in a fraction of the RNA fragments might be expected to liberate sequences which run faster than the large protected segments during electrophoresis on the polyacrylamide-urea gel, their presence in the fingerprint suggests that they remain associated with the main fragments because of an unusually stable secondary structure in this region of the 16S RNA molecule. (b) Fragment produced by RNase digestion of a complex containing 16S RNA and proteins S8, S15, S6 and S18. Twenty-five μg [^{32}P] 16S RNA were mixed with two molar equivalents of each protein in TMK buffer and incubated at 0°C for 30 min. Reaction mixtures were treated with 5 μg pancreatic RNase A for 5 min at 30°C , chilled, and fractionated by sucrose gradient centrifugation for 16 hr at 32,000 rev/min in an IEC SB283 rotor²⁰. The protein-RNA fragment complex, which sedimented at 7S, was deproteinized by phenol extraction, precipitated with ethanol and fingerprinted as described in Fig. 2. Note that all the T_1 oligonucleotides present in (a) can also be detected in (b). The occurrence of the oligonucleotide AAACG results from a scission at the top of the loop in section D' (see Fig. 4). The sequence UCCACG from section D' is present in very low yield.

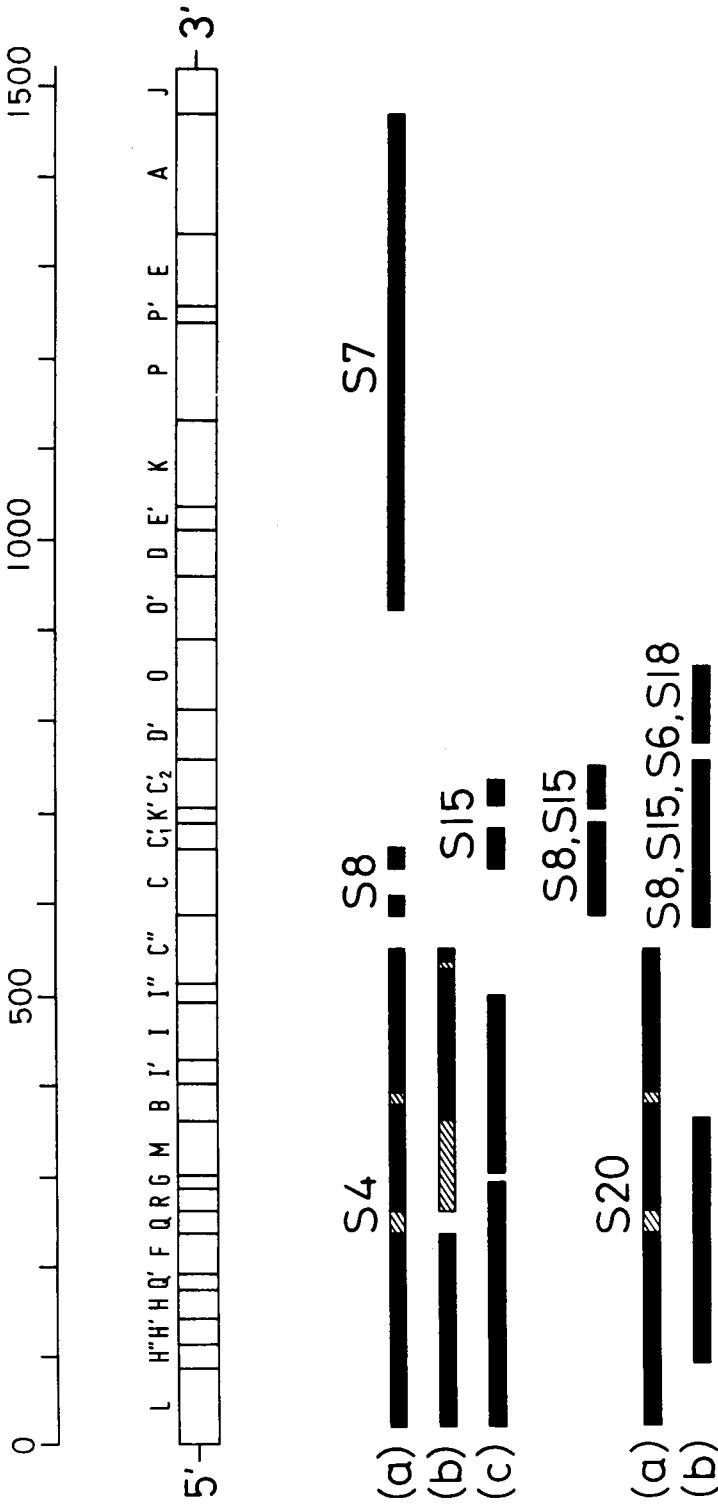


A ribonucleoprotein arising from pancreatic RNase digestion of a complex between 16S RNA and proteins S8, S15, S6 and S18 was isolated by the sedimentation technique²⁰. In addition to the four proteins, the particle contained an RNA fragment of approximately 7S that extended from section C to the middle of section 0 (Fig. 4). Moreover, the fingerprint of this fragment, presented in Fig. 5b, revealed the presence of section K', a partially single-stranded sequence that lies at the top of the hypothetical hairpin loop of which sections C₁' and C₂' comprise the stem. The implications of this finding will be discussed below. After deproteinization by phenol, the purified 7S RNA was able to bind S8, S15, S6 and S18, but no other proteins, from a mixture of unfractionated 30S subunit proteins²⁰.

Location of binding site for protein S7

Protein S7 has been shown by the sedimentation method to remain associated with a fragment of 3S following T₁ RNase digestion of its complex with 16S RNA⁵. The 8S RNA contains approximately 600 nucleotides from the 3'-terminal portion of the 16S RNA, and extends from section 0' through section A (see Fig. 6). The purified 8S fragment did not rebind protein S7, however, nor was a protected ribonucleoprotein fragment recovered from RNase digests of the S7-16S RNA complex by the electrophoretic method. It thus appears that complexes between RNA fragments and protein S7 alone were unstable under the conditions employed, and further localization of the S7 binding site was not possible. An RNA fragment with essentially the same nucleotide sequence as the 8S RNA has been isolated in association with proteins S7, S9, S10, S13 and S19 following mild RNase digestion of unfolded 30S subunits²¹ and it is possible that the additional proteins promote a more stable interaction between S7 and its binding site.

Figure 6. Location of protein binding sites in the 16S RNA. The 16S molecule is depicted schematically at the top of the diagram. The molecule is divided into sections, designated by upper case letters, that have been defined by the primary structure analysis of Ehresmann *et al*⁹. The scale above the 16S RNA denotes distances in nucleotide residues from the 5'-terminus of the molecule. The bars in the lower part of the diagram indicate the relative positions and lengths of the RNA fragments that specifically bind proteins S4, S7, S8, and S20 of the 30S ribosomal subunit. Fragments capable of interacting simultaneously with proteins S8 + S15 and proteins S8 + S15 + S6 + S18 are also located in the scheme. Cross-hatched regions within the bars indicate oligonucleotides recovered in reduced yield. S4. (a) 9S fragment produced by pancreatic RNase digestion of 16S RNA or of protein-16S RNA complexes⁵. (b) Protected fragment isolated from pancreatic RNase digest of S4-16S RNA complex by gel electrophoresis¹². (c) Two fragments isolated by polyacrylamide-urea gel fractionation of the 12S RNA fragment¹⁶, left, 12S-I, right, 12S-II (see Fig. 2). S20 (a) same as for S4. (b) Protected RNA recovered from T₁ RNase digest of S20-16S RNA complex¹¹ (see Fig. 3). S8. Protected fragment produced by pancreatic RNase hydrolysis of S8-16S RNA complex¹¹ (see Fig. 5). S15. Protected fragment isolated from pancreatic RNase digest of S15-16S RNA complex¹¹. S8, S15 5S RNA resulting from pancreatic RNase digestion of S8, S15-16S RNA complex¹⁸. S8, S15, S6, S18. RNA segments protected from pancreatic RNase hydrolysis in the complex of 16S RNA with proteins S8, S15, S6 and S18²⁰ (see Fig. 5). S7. 8S fragment produced from 16S RNA by T₁ RNase digestion⁵.



Characteristics of ribosomal protein binding sites

The isolation and nucleotide sequence analysis of a variety of specific fragments of the 16S RNA has enabled us to locate with some precision the binding sites for several proteins of the 30S ribosomal subunit. A scheme summarizing our results is presented in Fig. 6. It is evident that there is marked overlapping of the RNA segments that contain the binding sites both for proteins S4 and S20 and for proteins S8 and S15. In particular, proteins S4 and S20 interact with sequences within the 500 5'-terminal residues of the 16S RNA and it is likely that protein S17 can attach within this portion of the molecule as well. The binding site for 50S subunit protein L24 has been located within a segment of similar length at the 5'-end of the 23S RNA²² and it will be of interest to see whether other proteins of the large subunit bind within the same region.

Two criteria were used to establish that a given RNA fragment contained binding sites for one or more specific proteins. First, it was demonstrated that the isolated fragment could re-associate with the protein or proteins in question, but not others. The specificity of the interaction was further substantiated in many instances by measurement of binding stoichiometry. Second, RNA sequences protected from nuclease attack in the RNA-protein complex, but labile in the free RNA molecule, were judged to be at or near the protein binding site.

Where the apparent protection extended to sequences of several hundred nucleotides, as for S4 and S20, it seemed unlikely that the RNA could be shielded by direct contact with the protein molecule. It has been suggested that in such cases, the protein masks a small number of normally sensitive single-stranded segments which connect a series of base-paired loops²².

Double-stranded portions of the 16S RNA have been shown to be much less sensitive to partial hydrolysis with T_1 and pancreatic RNases than unpaired sequences¹⁴. Moreover, the compact conformation assumed by the RNA chain at the high Mg^{++} concentrations used in these experiments further reduces its susceptibility to RNase⁶. Accordingly, large RNA fragments containing intact protein binding sites can frequently be recovered from enzymatic digests of free 16S RNA⁵. The stability of these fragments must therefore be attributable to the substantial intrinsic resistance of the RNA molecule arising from its own particular configuration.

Since our conclusions about binding site size and structure have been based upon the analysis of a finite number of specific RNA fragments, it is not yet possible to ascertain whether the actual sites of protein attachment are smaller, or perhaps even larger, than those delineated in the present study. However, all of the fragments examined so far have been able to reassociate specifically with individual ribosomal proteins, suggesting that they encompass the sequences that directly contact the protein molecule. Although the presence of extraneous portions of the RNA cannot be excluded, sequences not in direct contact with the protein may be essential to ensure that the proper secondary and tertiary of the binding sites are maintained.

The importance of RNA configuration in the formation of specific ribonucleoproteins complexes is in fact suggested by several features of the S4-16S RNA interaction. The RNA segment that contains the binding site for protein S4 possesses an intricate secondary structure consisting of 15-16 hairpin loops connected by short single-stranded sequences (see Fig. 1). Despite its complexity, this segment must apparently assume a

highly specific configuration in order to associate with the protein since small structural alterations induced by variations in temperature and ionic environment exert a strong influence upon the S4-16S RNA interaction^{23,24}. When the Mg^{++} ion concentration is reduced below 0.01 M, for example, reversible conformational transitions in the 16S RNA are manifested by alterations in sedimentation rate and RNase susceptibility. At the same time, the ability of the RNA to interact with protein S4 sharply decreases^{6,23,24}. The sensitivity of the S4 binding site to conformational alterations may reflect the fact that this protein attaches to two or more distinct sequences at the 5'-end of the 16S RNA (see Fig. 6). The need for a molecular framework to maintain these sequences in their correct three-dimensional relationship might explain in part why the binding site for protein S4 spans such a large portion of the 16S RNA molecule.

Changes in RNA conformation have also been found to occur as a consequence of the protein S4-16S RNA interaction. First, the accessibility to RNase of certain bonds within the S4 binding site changes following attachment of the protein, suggesting an overall shift in RNA conformation¹⁶. Second, S4 binding provokes dissociation of ethidium bromide from the 16S RNA, indicating an alteration in the secondary structure of the nucleic acid molecule²⁵. Although the functional significance of these observations cannot be definitively assessed at present, it is possible that small rearrangements of RNA structure upon RNA-protein interaction might permit a better fit between the two components.

The nature of RNA-protein interactions in the ribosome

The evidence presented into this report suggests that specific RNA-protein recognition and interaction in the ribosome can

occur by several different mechanisms. In particular, the regions of the 16S RNA which interact with individual ribosomal proteins and which are protected from RNase digestion by them vary markedly in size and structure. The binding site for protein S8, for instance, consists of two complementary sequences of 20 nucleotides derived from the stem of a single hairpin loop and is thus believed to be largely double-stranded (Fig. 4). By contrast, protein S4 interacts with sequences extending over 400-500 nucleotide residues. In addition to many base-paired loops of varying size, this region contains a substantial number of single-stranded sequences, some of which are protected from RNase attack in the S4-16S RNA complex. Furthermore, we have shown here that protein S4 can bind simultaneously to two distinct and nonoverlapping RNA fragments. This protein must therefore make at least two, and perhaps more, contacts with the 16S RNA molecule. Similar conclusions were reached by Schaup and Kurland⁸, although the extent to which the fragments examined by them overlapped one another was not determined. No evidence of multiple interactions has been obtained so far for any of the other RNA-binding proteins. Finally, the stability of different protein binding sites varies widely as a function of solution conditions. Thus, interactions of S4, S7 and S20 with the 16S RNA are much more sensitive to alterations in the Mg^{++} ion concentration than are those of S8 and S15²³. Furthermore, the dependence of complex formation and stability upon temperature differs materially for proteins S4 and S8²⁴.

Although the chemical nature of the specificity exhibited by RNA-protein interactions is not yet understood, structural regularities in the RNA molecule might provide the bases for recognition⁹. First, palindromes are found throughout the

nucleic acid chain. Several of these occur in regions such as sections G, I, I' and D' that are thought to participate in protein binding. Second, twofold axes of symmetry are present in a number of base-paired hairpin loops. Symmetrical segments with only slight imperfections can be noted in sections F and O, as well as in section C, which contains the binding site for protein S8.

In addition to providing sites of attachment for the six proteins discussed above, the 16S RNA may well participate in other RNA-protein interactions during ribosomal subunit assembly. In particular, it is quite possible that proteins unable to bind independently to the 16S RNA may nonetheless be able to interact directly with specific sites in the nucleic acid molecule once other proteins are attached. Evidence that proteins S6 and S18 take part in such interactions is provided by experiments described in the present report. Although neither S6 nor S18 can bind independently to the 16S RNA, both will form a stable complex with the RNA molecule if S8 and S15 are also bound. When a complex containing all four proteins is treated with RNase, twice as much RNA is shielded from digestion as when S8 and S15 alone are present^{18,20}. Although other explanations can be imagined, it is likely that the increased protection arises from direct interactions of S6 and S18 with the additional nucleotide sequences recovered, because the presence of these sequences is necessary for reassociation of S6 and S18 with the RNA fragment²⁰. Since the additional protected sequences lie adjacent to those containing the binding sites for proteins S8 and S15, it is possible that the independent attachment of these two proteins imposes specific conformational constraints on nearby segments of the RNA chain that lead to the organization of binding sites accessible to proteins S6 and S18. The cooperative mechanism

envisioned here could play a major role in correctly positioning proteins within the subunit structure during the assembly process.

ACKNOWLEDGEMENTS

We wish to express our appreciation to Drs. A. Tissieres and H. G. Wittmann for encouragement and counsel, and to Miss A-M. Piret, Mrs. Z. Zabielski, Mrs. C. Bran and Mrs. O. Marconato for expert technical assistance. G.A.M. was a Fellow of the Medical Research Council of Canada, R.A.G. the recipient of a short-term fellowship from the European Molecular Biology Organization. This work was supported by grants from the Fonds National Suisse de la Recherche Scientifique, the Centre National de la Recherche Scientifique, the Delegation Generale a la Recherche Scientifique et Technique and the Commissariat a l'Energie Atomique.

* Max-Planck Institut für Molekulare Genetik, Inhestrasse 63-73,
1 Berlin 33, Germany

† Institute de Biologie Moléculaire et Cellulaire du CNRS, 15, rue
René Descartes, 67000 Strasbourg, France

Present
addresses:

¹ Department of Biochemistry, University of Massachusetts, Amherst,
Massachusetts 01002, USA

² Research Institute for Nuclear Medicine and Biology, Hiroshima
University, Hiroshima, Japan

³ Searle Research Laboratories, Lane End Road, High Wycombe,
Bucks., England

REFERENCES

1. Mizushima, S., and Nomura, M., Nature, 226, 1214-1218 (1970).
2. Schaup, H. W., Green, M., and Kurland, C.G., Molec. Gen. Genetics, 109, 193-205 (1970).
3. Schaup, H. W., Green, M., and Kurland, C. G., Molec. Gen. Genetics, 112, 1-8 (1971).
4. Garrett, R. A., Rak, K. H., Daya, L., and Stoffler, G., Molec. Gen. Genetics, 114, 112-124 (1971).

5. Zimmermann, R. A., Muto, A., Fellner, P., Ehresmann, C., and Branlant, C., Proc. Natn. Acad. Sci. U.S.A., **69**, 1282-1286 (1972).
6. Muto, A., Ehresmann, C., Fellner, P., and Zimmermann, R. A., J. Molec. Biol. **86**, 411-132 (1974).
7. Held, W. A., Ballou, B., Mizushima, S., and Nomura, M., J. Biol. Chem. **249**, 3103-3111 (1974).
8. Schaup, H. W., and Kurland, C. G., Molec. Gen. Genetics, **114**, 350-357 (1972).
9. Ehresmann, C., Stiegler, P., Mackie, G. A., Zimmermann, R. A., Fellner, P., and Ebel, J. -P., Nucleic Acids Research, **2**, 265-278 (1975)
10. Zimmermann, R. A., Muto, A., and Mackie, G. A., J. Molec. Biol. **86**, 433-450 (1974)
11. Ungewickell, E., Garrett, R. A., Ehresmann, C., Stiegler, P., and Fellner, P., Eur. J. Biochem, (in the press).
12. Mackie, G. A., and Zimmermann, R. A., J. Biol. Chem, (in the press).
13. Fellner, P., Ehresmann, C., and Ebel, J. -P., Biochimie, **54**, 853-900 (1972).
14. Ehresmann, C., Stiegler, P., Fellner, P., and Ebel, J. -P., Biochimie, **54**, 901-967 (1972).
15. Schaup, H. W., Sogin, M., Woese, C., and Kurland, C. G., Molec. Gen. Genetics, **114**, 1-8 (1971).
16. Mackie, G. A., and Zimmermann, R. A., J. Molec. Biol, (in the press).
17. Green, M., and Kurland, C. G., Molec. Biol. Reports, **1**, 105-111 (1973).
18. Zimmermann, R. A., and Bergmann, K., J. Molec. Biol., (in the press).
19. Schaup, H. W., Sogin, M. L., Kurland, C. G., and Woese, C. R., J. Bacteriol, **115**, 82-87 (1973).
20. Zimmermann, R. A., Proc. Natn. Acad. Sci. U.S.A., (in the press).
21. Szekely, M., Brimacombe, R., and Morgan, J., Eur. J. Biochem, **35**, 574-581 (1973).
22. Branlant, C., Krol, A., Sriwidada, J., Fellner, P., and Crichton, R., FEBS Lett., **35**, 265-272 (1973).
23. Muto, A., and Zimmermann, R. A., J. Molec. Biol., (in the press).
24. Schulte, C., Morrison, C. A., and Garrett, R. A., Biochemistry, **13**, 1032-1037 (1974).
25. Bollen, A., Herzog, A., Favre, A., Thibault, J., and Gros, F., FEBS Lett., **11**, 49-54 (1970).
26. Sanger, F., Brownlee, G. G., and Barrell, B. G., J. Molec. Biol., **13**, 373-398 (1965).
27. Brownlee, G. G., and Sanger, F., J. Molec. Biol., **23**, 337-353 (1967).
28. Noller, H. F., Biochemistry, **13**(23), 4694-4703 (1974).