Proc. Natl. Acad. Sci. USA
Vol. 78, No. 12, pp. 7331–7335, December 1981
Biochemistry

A "bulged" double helix in a RNA–protein contact site

(5S RNA/ribosomal protein L18/chemical modification/RNase digestion/comparative sequence evidence)

DEBRA A. PEATTIE*†, STEVE DOUTHWAITE‡, ROGER A. GARRETT‡, AND HARRY F. NOLLER§

*Department of Biochemistry and Molecular Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138; §Thimann Laboratories, Division of Natural Sciences, University of California, Santa Cruz, California 95064; and ‡Division of Biostructural Chemistry, Chemistry Institute, Aarhus University, DK-8000 Aarhus-C, Denmark

Communicated by Paul Doty, August 10, 1981

ABSTRACT The binding of ribosomal protein L18 affects specific nucleotides in *Escherichia coli* 5S RNA as detected by dimethyl sulfate alkylation and RNase A digestion of the 5S–L18 complex. Most of the affected nucleotides are clustered and localize a site of RNA–protein interaction in and around the defined central helix [Fox, G. E. & Woese, C. (1975) *Nature (London)* 256, 505–507] of 5S RNA. Chemical carbethoxylation of the native 5S RNA with diethyl pyrocarbonate shows that a striking feature of this region is an unstacked adenosine residue at position 66. We propose that this residue exists as a singly bulged nucleotide extending the Fox and Woese central helix by two base pairs in the *E. coli* sequence (to positions 16–23/60–68) as well as in each of 61 (prokaryotic and eukaryotic) aligned 5S RNA sequences. In each case, the singly bulged nucleotide is at the relative position of adenosine-66 in the RNA sequences. The presence of this putative bulged nucleotide appears to have been conserved in 5S RNA sequences throughout evolution, and its identity varies with major phylogenetic divisions. This residue is likely involved in specific 5S RNA–protein recognition or interaction in prokaryotic and eukaryotic ribosomes. The uridine-65 to adenosine-66 internucleotide bond is protected from RNase A digestion in the complex, and carbethoxylation of *E. coli* adenosine-66 prior to L18 binding affects formation of a stable RNA-protein complex. Thus, we identify a region of *E. coli* 5S RNA protected by the ribosomal protein L18 and propose that it contains a bulged nucleotide residue important in stable formation of this RNA–protein complex. This bulged residue appears to be evolutionarily conserved and phylogenetically defined in 5S RNA sequences in general, and consideration of other known RNA–protein binding sites shows that such a "bulged helix" may be a common feature of RNA–protein contact sites.

Little is known about the signals by which ribosomal proteins recognize their binding sites on ribosomal RNA during ribosomal assembly. Equally uncertain is our understanding of the effect of protein–RNA interactions on the conformation of either molecule, although it has often been suggested that here we may find clues to biomechanical principles underlying ribosome function (e.g., ref. 1). Any attempt to understand these processes suffers from the lack of detailed structural knowledge of the ribosomal protein binding sites. Although recent progress on the structures of the 16S and 23S RNA has shed light on aspects of this problem (e.g., ref. 2), the size and structural complexity of the large RNAs deny simple experimental approaches to these questions. For this reason, we have used the relatively simple 5S RNA system.

In bacterial ribosomes, more well defined than eukaryotic ribosomes, 5S RNA (120 nucleotides in *Escherichia coli*) specifically binds three ribosomal proteins in the large ribosomal subunit: L5, L18, and L25 in *E. coli* (3). The resulting ribonucleoprotein complexes are believed to be closely analogous to those involving the large RNAs; therefore, we expect studies of the 5S RNA system to provide insight into ribosomal RNA–protein interactions in general and possibly into nonribosomal systems as well. The exact function of the 5S RNA is, as yet, uncertain (4).

Of the three 5S RNA binding proteins, L18 binds most strongly; $K_a = 10^8\ M^{-1}$ (3). In earlier studies (5) using protein protection against modification of 5S RNA by kethoxal, we localized the L18 binding site approximately to the central helix of 5S RNA. We have now used other chemical and enzymatic probes to provide a more detailed description of the binding site—dimethyl sulfate to probe tertiary interactions at the guanosine N-7 positions and both secondary and tertiary interactions at the cytidine N-3 positions (6), diethyl pyrocarbonate to monitor the stacking of adenosines resulting from base pairing or tertiary interactions (6), and the pyrimidine-specific RNase A to monitor internucleotide bond susceptibilities. In addition, we have lightly damaged 5S RNA molecules chemically (7), refolded them, and then selected those molecules that still bind stably to ribosomal protein L18 (8). In these ways, we have characterized a major contact site in the 5S–L18 complex and have found within it a striking structural feature that is a common element of prokaryotic and eukaryotic 5S RNA sequences.

MATERIALS AND METHODS

RNA and Protein. The A conformer of 5S RNA was prepared from *E. coli* ribosomes (9, 10). Protein L18 was isolated (11) and complexed (5) with 5S RNA (with >80% efficiency) in buffer A (70 mM sodium cacodylate, pH 7.2/15 mM $MgCl_2$/300 mM KCl).

Probing of Native RNA and RNA–Protein Complex. Both 5S RNA and the RNA–protein complex were chemically modified in buffer A (8). The dimethyl sulfate reaction (1 μl of dimethyl sulfate in a 200-μl reaction volume) was 20 min at 30°C, and the diethyl pyrocarbonate reaction (20 μl of reagent grade diethyl pyrocarbonate in a 200-μl reaction volume) was 2 hr (with gentle mixing every 30 min) at 30°C. Any disruption of the RNA–protein complexes was monitored by polyacrylamide gel electrophoresis (12). After chemical modification, the RNA was radioactively labeled at its 3′ terminus with [5′-^{32}P]pCp and repurified on a polyacrylamide gel (with an extraction efficiency of >80% in 0.2 M ammonium acetate, pH 5.5) before finishing the chemical reactions and subjecting the fragments to electrophoresis on a polyacrylamide sequence analysis gel (6, 7). The 5S RNA–L18 complexes were extracted twice with neutralized phenol before the RNA was radioactively labeled.

Enzymatic Digestion of Native RNA and RNA–Protein Complex. 5S RNA and the 5S RNA–protein complex were digested at 0°C with RNase A (Sigma) in 30 mM Tris·HCl, pH 7.8/20 mM $MgCl_2$/300 mM KCl at an RNase/RNA ratio of 1:10,000–1:100 units of RNase A per μg of RNA.

The digested samples were subjected to electrophoresis on a 12% polyacrylamide gel (7 V/cm of length) at 4°C for 16 hr using a circulating buffer of 40 mM Tris·HCl, pH 8.0/10 mM $MgCl_2$. The band migrating alongside undigested 5S RNA or

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

† Present address: Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305.

the 5S RNA–L18 complex was detected by autoradiography and excised. The RNA was extracted as above (>80% efficiency) and then subjected to electrophoresis on a sequence analysis gel.

Damaged 5S RNA Selection Experiments. These were done essentially as described by Peattie and Herr (8). After renaturation, the RNA was complexed (5) with L18 in 20 μl of buffer B (30 mM Tris·HCl, pH 7.8/20 mM $MgCl_2$/300 mM KCl). Stable RNA–protein complexes were selected on nitrocellulose filters by washing three times with buffer B. After inducing strand scission within the selected molecules (8) and subjecting the products to polyacrylamide gel electrophoresis, densitometer tracings were made of individual lanes on the resultant autoradiograph by using a Joyce–Loebl instrument.

Comparative Sequence Analysis. 5S RNA sequences were chosen at random and aligned as described (13, 14).

RESULTS

Effects of ribosomal protein L18 on 5S rRNA

Chemical Probing Studies. To determine the effect of L18 on chemical methylation at the guanosine N-7 and the cytidine N-3 positions, we treated 5S RNA and the 5S RNA–L18 complex with dimethyl sulfate (6). The 5S RNA was then labeled at its 3′ terminus with ^{32}P, and strand scission was induced at the sites of modification. The products were resolved by polyacrylamide gel electrophoresis and identified by autoradiography (Fig. 1). Sites of methylation protection or enhancement could be inferred from parallel sequence lanes, and they were quantitated (data not shown) by densitometer tracings of the autoradiographs (unpublished results).

Guanosine-67 is the guanosine most susceptible to dimethyl sulfate attack in the 5S molecule, and it is strongly protected in the 5S RNA–L18 complex (Fig. 1). The N-7 positions of eight other guanosines are affected by bound L18: guanosine-16, -23, -24, -51, -54, -56, and -64 are protected from methylation, whereas guanosine-61 exhibits enhanced methylation. Guanosine-10 may also be more susceptible to methylation in the presence of complexed L18, but this result was variable and requires further study. No cytidines appear affected by L18 binding.

We also probed 5S RNA and the 5S RNA–L18 complex with diethyl pyrocarbonate to monitor adenosine stacking. We found that adenosine-66 was the most strongly modified adenosine within the native 5S molecule (see Fig. 1); however, we could not monitor any L18-induced effects on the 5S RNA because diethyl pyrocarbonate dissociated the RNA–protein complex. The damaged RNA selection experiments (see below) compensated for this inconvenience.

Enzymatic Digestion Studies. Possible L18 protection of RNase-sensitive sites was also examined. The 5S RNA was labeled at its 3′ terminus, renatured into the native A conformer, and digested in the presence and absence of L18 by the single-strand-specific RNase A. The intact 5S RNA or complex, which contained "hidden breaks," was then purified on a native polyacrylamide gel, denatured, and run on a sequence analysis gel.

L18 protects cytidine-42 and the uridine–adenosine internucleotide bonds at positions 14–15 and 65–66 from cleavage by RNase A (Figs. 2 and 3).

Damaged RNA Selection Studies. As described for tRNA molecules (8), we probed the 5S RNA to determine whether any adenosines are crucial for complex formation with L18; we were particularly interested in the adenosine-66 residue. These studies involved (*i*) modifying the purines of the RNA in a limited fashion with diethyl pyrocarbonate at 90°C, (*ii*) renaturing the RNA and complexing it with L18, and (*iii*) selecting stable 5S–L18 complexes on nitrocellulose filters. The RNA was then dissociated from its complex and fragmented at the internal sites of damage by chemical strand scission. Modification of adenosine-66, which lies within the central helix protected by L18,

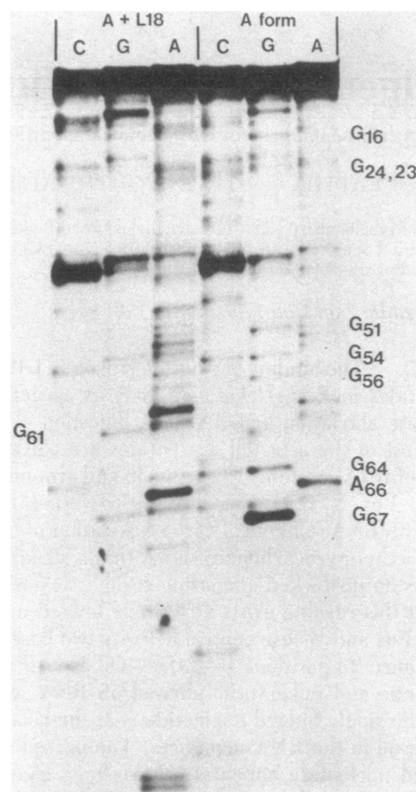

FIG. 1. Chemical modification shows that the central helix of *E. coli* 5S RNA is affected by ribosomal protein L18. L18 protects the N-7 positions of guanosine-16, -23, -24, -51, -54, -56, -64, and -67 from dimethyl sulfate attack within the molecule; guanosine-61 exhibits enhanced reactivity to dimethyl sulfate when L18 is bound. Adenosine-66 is unstacked and the most reactive adenosine to diethyl pyrocarbonate within the 5S RNA molecule. The 5S–L18 complex dissociates in the presence of diethyl pyrocarbonate, and the protein-induced effects in this A lane are variable. The 12% polyacrylamide gel was run according to Peattie (7). The numbered nucleotides correspond to residues within the dotted lines in Fig. 3.

decreases the formation of stable complexes at least 50% (Fig. 4).

Comparative sequence analysis

We compared 61 known prokaryotic and eukaryotic 5S RNA nucleotide sequences using the sequence alignment described, first by Fox and Woese (13) and later by Hori and Osawa (14). In all cases, the central helix can be extended by two base pairs (one is always a G·C pair) if a base within the helical sequence is bulged outside the helix (Table 1). The position of the bulged base is invariant within the helix but its identity varies with major phylogenetic divisions (Table 1). It is an adenosine in the 32 prokaryotic sequences, the 6 yeast sequences, and the 1 dinoflagellate (a eukaryotic protist) sequence. It is a uridine in the 7 plant sequences and the 3 RNAs from unusual niches (a cyanobacterium, a halophilic bacterium, and an anaerobic bacterium). It is a cytidine in the 11 animal sequences and in the eukaryotic slime mold sequence.

DISCUSSION

Ribosomal protein L18 affects dimethyl sulfate alkylation of nine guanosines in *E. coli* 5S RNA. Such alkylation monitors tertiary interactions at the N-7 positions of the guanosines in a RNA molecule (6), and protein-induced methylation protection or enhancement at specific guanosines can help localize a protein binding site. [Note that it is *changes* in the methylation pattern that are important because, even in the absence of protein, not all guanosines (paired or unpaired) in a native RNA molecule

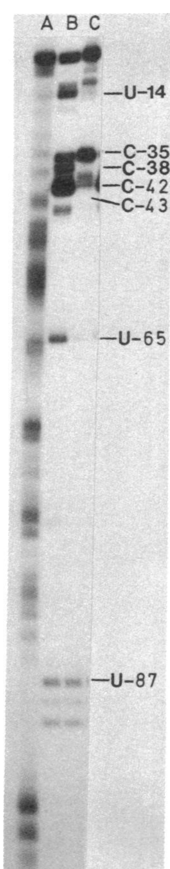

FIG. 2. Ribosomal protein L18 affects RNase A digestion of 5S RNA. Lanes: A, adenosine cleavage of 5S RNA at 90°C (7); B, 5S RNA digested with RNase A; C, 5S RNA–L18 complex digested with RNase A. The RNase A/RNA ratio was 1:500 (wt/wt) at 0°C for 10 min. The band doubling apparent on the lower half of the autoradiograph results from heterogeneity (15) at the labeled 3′ terminus of the molecule.

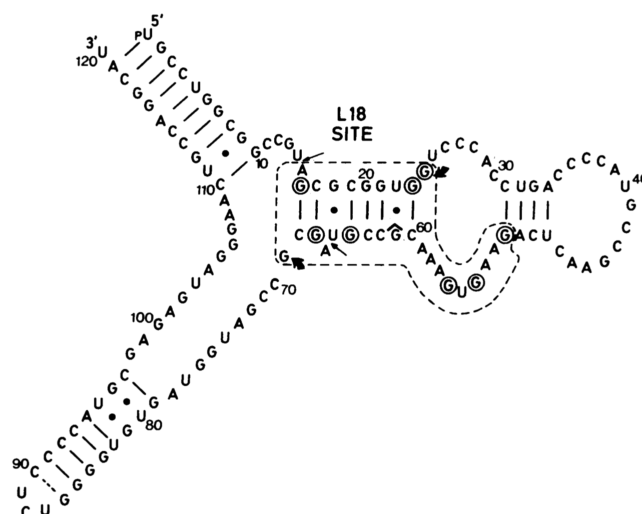

FIG. 3. The central region of *E. coli* 5S RNA (within dotted lines) is affected by ribosomal protein L18. This secondary structure is a modified version of the Fox and Woese model (13) and illustrates our proposed 2-base pair extension (guanosine-16/cytidine-68 and cytidine-17/guanosine-67) of the central helix, allowed by bulging out adenosine-66. The N-7 positions of the circled guanosines are protected from dimethyl sulfate modification in the RNA–protein complex; guanosine-67 is the most dramatically protected. Guanosine-61 (Ĝ) is more susceptible to dimethyl sulfate attack in the presence of the ribosomal protein. Curved arrows, N-1 positions of guanosine-24 and -69 protected from kethoxal modification in the RNA–protein complex. →, Internucleotide bonds (uridine–adenosine 14-15 and 65-66) protected from RNase A digestion in the complex.

react with dimethyl sulfate (6, 8).] In the 5S RNA–L18 complex, eight guanosines are protected from chemical attack, and one exhibits increased susceptibility to attack. Four of the eight protected guanosines, -16, -23, -64, and -67, are part of our extended central 5S RNA helix (see Fig. 3), and the protection of their N-7 positions suggests that L18 interacts with or lies along the major groove of the helix. The other four protected guanosines, -24, -51, -54, and -56, lie outside the central helix, indicating that the protein may also interact with extrahelical features. Our earlier chemical modification data (5) have also shown this to be the case: L18 strongly protects guanosine-24 and -69 from kethoxal modification and, conversely, kethoxal modification of these residues in the free RNA inhibits L18 binding, suggesting that these bases are directly involved in the RNA–protein interaction. Of the protected guanosines, -16, -23, -24, -54, and -69 are universal in prokaryotes, a result compatible with their being involved directly in the protein binding site. The enhanced methylation of guanosine-61 could reflect a protein-induced change in a tertiary interaction at the N-7 position; such methylation enhancement has been observed in the *E. coli* tRNA phenylalanine–ribosome interaction (8) and in DNA–protein interactions (25–27).

Adenosine-66 is both unstacked and the most reactive adenosine residue in the native 5S RNA conformer as detected by carbethoxylation with diethyl pyrocarbonate (Fig. 1). However, this chemical reagent, which inactivates some proteins (28), dissociates the 5S–L18 complex, and the chemically observed protein-induced effects vary. Yet the enzymatic RNase A digestion indicates that L18 affects the 5S central helix because the uridine-adenosine bonds at positions 14-15 and 65-66 are not hydrolyzed by RNase A in the presence of the ribosomal protein (Fig. 2). In addition, cobra venom ribonuclease (a double-strand-specific RNase) does not cleave the otherwise susceptible internucleotide bonds in the G-G-U sequence at positions 63-64-65 in this helix in the presence of L18 (unpublished results). L18 may also interact with the 5S RNA outside the central helix because it induces RNase A protection at cytidine-42 (Fig. 2). 5′-End-labeling experiments (not shown) show that loss of this cleavage site the apparent increased RNase A digestion at cytidine-35 and the apparent decreased cleavage at cytidine-43 (Fig. 2). The apparent L18 protection effect at cytidine-38 (Fig. 2) is ambiguous, since this RNase A cleavage site might be primary or secondary (unpublished results). Taken in conjunction with our earlier kethoxal data (5), the present chemical and enzymatic data indicate that the L18 binding site on *E. coli* 5S RNA is bounded at least by adenosine-14 and guanosines-24, -51, and -69 (see Fig. 3).

Fox and Woese proposed the original central helix (18–23/60–64) based on comparative sequence analysis (13). Our own analysis of 61 pro- and eukaryotic 5S RNA sequences (Table 1) supports inclusion of two additional base pairs (guanosine-16/cytidine-68 and cytidine-17/guanosine-67 in *E. coli* 5S RNA) with a single nucleotide (adenosine-66 in the *E. coli* sequence) bulged out from the helix. In support of this comparative sequence evidence, we also find (unpublished results) that the N-3 positions of cytidine-17 and -68 are blocked to chemical methylation in the native 5S RNA molecule; this further suggests that these cytidines are base paired (6). We did find three 5S sequences [trout, *Xenopus mulleri* (ovary), and *Lactobacillus brevis*] that contain single base mismatches in the proposed extended helix. However, they also contain the putative bulged base, and we cannot eliminate the possibility of an original sequence analysis error in this region of the RNA. Interestingly, certain archaebacteria possess the extended central helix but without a bulged base (refs. 29 and 30; C. Woese, personal communication). The proposed nucleotide bulge in the *E. coli* sequence is supported by the susceptibility of adenosine-66 to diethyl pyrocarbonate (Fig. 1), which monitors unstacked adenosines in native RNA molecules (6). Single noncomplementary nucleotides are known to loop out of double helices in a stable fashion (31), and single nucleotide bulges are the most common

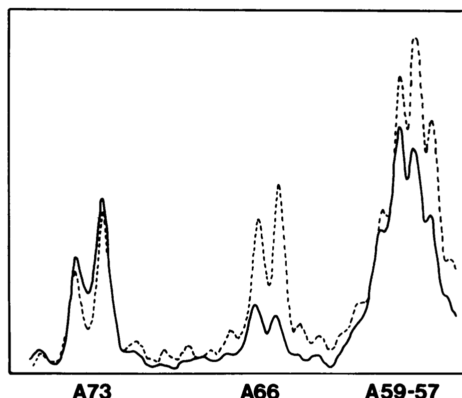

FIG. 4. Carbethoxylation of adenosine-66 prior to L18 binding decreases formation of stable 5S–L18 complexes by at least 50%. The end-labeled 5S RNA was lightly damaged chemically at adenosines (7) and then (*i*) fragmented directly (---) or (*ii*) complexed with ribosomal protein L18 and selected on nitrocellulose filters before fragmenting. The densitometer tracings were made directly from the autoradiograph of the polyacrylamide sequence analysis gel; only a portion of the tracing is shown here. Each labeled peak represents a band on the autoradiograph corresponding to a fragment produced by strand scission at a specific adenosine; the peaks are not singlets because 5S RNA isolated from cells often has less than molar amounts of the 3′-terminal base (15). The relative effects of carbethoxylation are based on the areas beneath the adenosine peaks. Adenosine-66 is the putative bulged base of the *E. coli* sequence in our proposed extension of the 5S central helix.

on the basis of random prediction (32, 33). The free energy estimation rules of Tinoco *et al.* (32, 33) predict a net increase in stability of approximately −2 kcal/mol (1 cal = 4.18 J) in the *E. coli* 5S RNA when the central helix is extended with the bulged base, and a space-filling model of this region (not shown) reveals that the bulged residue is accomodated easily with only a small deformation of the helical geometry.

The chemical damage-and-selection experiments indicate that adenosine-66 is important to a stable 5S–L18 complex. When damaged 5S molecules (damaged such that, on average, only one residue per molecule is affected) are complexed with L18, there is a selection against those containing a carbethoxylated adenosine-66 (Fig. 4). This could reflect aberrant refolding of the 5S molecule during renaturation prior to complex formation or it could directly locate a site critical to the 5S RNA contact with L18; either event suggests that this residue is important. It is not surprising that chemical modification of this residue still allows the formation of some stable 5S–L18 complexes, since our other data show that the site of contact between the RNA and the protein is fairly extensive. In addition, heterologous binding experiments (ref. 34; R. Zimmermann, personal communication) indicate that yeast 5S RNA (which has an adenosine-66 like the *E. coli* sequence) will not bind *E. coli* L18, suggesting that other features in the molecule must be crucial for stable complex formation.

What function could such a bulged nucleotide serve in a RNA–protein contact site? The conserved bulged nucleotide in the 5S RNA central helix could provide a recognition signal for the appropriate ribosomal protein. The unstacked base projecting outward from the helix could help distinguish the central helix from similar helices in cellular RNA, and protein binding to the unstacked base could in turn contribute to the stability of the RNA–protein complex. Conversely, the bulged nucleotide could have some other role in contributing to ribosome structure or function and the role of the protein might be to stabilize the bulge. Intriguingly, the identity of the putative bulged nucleotide varies among major divisions in the phylogenetic kingdom (Table 1). Among the sequences considered here, it is an adenosine for aerobic bacteria, the dinoflagellate,

Table 1. 5S RNA sequences show evolutionary conservation of a central helix containing a bulged base: The identity of the base varies among phylogenetic divisions

Central helix, proposed base pairs, and bulged base*	Organism†
G C G U U A U G **C G**$_{A}$ U A A U G C	*Phosphobacterium phosophoreum*
G C G C G G U G **C G**$_{A}$ U G C C G C	*Yersinia pestis, Serratia marcescens, Erwinia aeriodae, Proteus mirabilis, Aerobacter aerogenes, Salmonella typhimurium, E. coli*-C1, -C2, -A1, -A2, -KA, -KB, -KC, -KD
G C G G C G U G **C G**$_{A}$ C C G C G C	*Thermus aquaticus*
G C G A A G A G **C G**$_{A}$ U U U C U C	*Bacillus megaterium* -1, -2 *Bacillus pasteurii* -1, -2 *Bacillus licheniformis* -1, -2 *Bacillus subtilus*
G C G G A G R G **C G**$_{A}$ C C U C Y C	*Bacillus stearothermophilus, Bacillus brevis*
G C G A R R A G **C G**$_{A}$ C U Y Y U C	*Bacillus Q, Bacillus firmus -1, -2*
G C A U U G A G **C G**$_{A}$ U A A C U C	*Lactobacillus viridescens*
G C G U G R G G **C G**$_{A}$ Y A Y Y C C	*Streptomyces griseus* (17) *Micrococcus lysodeikticus* (18)
G C G A G A A G **C G**$_{A}$ U U C U U C	*Streptococcus faecalis*
C C G U G U C G **G G**$_{A}$ C A C G G C	*Crypthecodinium cohnii* (19) (a dinoflagellate)
G C G G U A U G **C G**$_{U}$ C C A U A C	*Anacystis nidulans* (a cyanobacterium)
G C G G U G G G **C G**$_{U}$ C C G C C C	*Halobacterium cutirubrum* (20)
A C U U A G A G **U G**$_{U}$ A A U C U C	*Clostridium pasteurianum* (an anaerobic bacterium)
C C A C C A C G **G G**$_{U}$ U G G U G C	*Chlorella pyrenoidosa* (21) (a unicellular eukaryotic alga)
Y C A G C A C **G G**$_{U}$ U C G U G	Rye, tomato, sunflower, dwarf bean, broad bean, wheat embryo
U C U A C C A G **A G**$_{A}$ A U G G U C	*Saccharomyces cerevisiae, Saccharomyces carlsbergensis, Kluyvermeyces lactis*
U C U A G C A G **A G**$_{A}$ A U C G U C	*Torulopsis utilis, Pichia membranafaciens* -1, -2
C U A G G U U G **G A**$_{C}$ U C C A G C	*Dictyostelium discoideum* (22) (a eukaryotic slime mold)
C C A C C C U G **G G**$_{C}$ U G G U A C	Human KB and HeLa cells, *Iguana iguana*, turtle, chicken, mouse (23), kangaroo rat (24), *Xenopus laevis* (somatic, ovary), *Xenopus mulleri* (somatic)

* Bases in boldface type constitute our proposed extension of the central helix (top line of helix, 5′ to 3′). The bulged base is in boldface italic type below the helical sequence. R, purine; Y, pyrimidine.

† All sequences are given in refs. 14 and 16 unless otherwise noted.

and yeast; a cytidine for animals; and a uridine for plants; it is also a uridine for the cyanobacterium, halophilic bacterium, and anaerobic bacterium. Thus, the identity of the bulged nucleotide could reflect a subtle discriminatory element for 5S RNA–protein recognition or interaction (or both) within the ribosomes of organisms in different phyla.

How general is the concept of a bulged nucleotide in a helix as a functional element of a ribonucleic acid–protein binding site? Examining the nucleic acid binding sites of other proteins, both ribosomal and nonribosomal, shows a substantial number of putative "bulged helices." For example, there are three helical regions in 16S ribosomal RNA that have single bulged nucleotides in or adjacent to protein binding sites of proteins S4, S7, and S15 (35). [However, other proteins, such as ribosomal protein S8 (35), recognize RNA structures that lack single bulged nucleotides.] In addition, translational repression of the mRNA of an *E. coli* ribosomal protein (thought to involve complex formation with a regulatory molecule or complex such as a ribosomal protein or assembly intermediate) involves a region of the message containing a 12-base pair stem with a bulged cytidine residue (36). It may be that a similar functional element also exists in other nonribosomal systems. For example, the most stable secondary structures of the coat protein binding sites of both R17 (37) and Qβ (38) RNA include a 7-base pair helix containing a singly bulged adenosine, and a possible secondary structure (39) of U-1 RNA contains two helices that have singly bulged pyrimidines. This small RNA exists as a nuclear-ribonucleoprotein complex and is suggested as having a role in processing heterogeneous nuclear RNA (40). Similarly, the initiating protein disc of tobacco mosaic virus binds to a specific site in the RNA during initiation of virus assembly, and the postulated secondary structure of this site (41) contains a hairpin loop that has two singly bulged nucleotides. In addition, the proposed secondary structure of the leader RNA of the tryptophan operon (42, 43) that is correlated with its transcription termination contains an 8-base-pair helix with a singly bulged adenosine; an alternative secondary structure (43), one that necessitates disruption of this bulged helix, is proposed for transcription of the operon.

In conclusion, *E. coli* ribosomal protein L18 binds to a region of the 5S RNA that includes its central helix. We find, in general, that the central helix proposed by Fox and Woese (13) can be extended by two base pairs in 5S RNA sequences if a single nucleotide is allowed to bulge from the helix. In the *E. coli* sequence, this bulged nucleotide (adenosine-66) is important in forming and maintaining (or both) stable 5S RNA–L18 complexes. In general, the presence of a bulged residue at this position appears to be evolutionarily conserved while its identity seems to be phylogenetically defined. Several other ribosomal and nonribosomal proteins also appear to recognize bulged helices; therefore, such helices may be a general structural element of ribonucleic acid–protein binding sites.

We thank Helga Boedtker Doty and Olke Uhlenbeck for critical reading of the manuscript and Gina Yiangou and Anni Christensen for excellent technical assistance. D.A.P. was supported by Grant GM09541-19 from the National Institutes of Health to Walter Gilbert, by a Visiting Fellowship from the Imperial Cancer Research Fund, London, and by a European Molecular Biology Organization Short-Term Fellowship. S.D. received support from Aarhus University, the Deutsche Forschungsgeneinschaft, the Max Planck Gesellschaft, and the European Molecular Biology Organization. R.A.G. was supported by grants from the Danish Research Council and the Aarhus University Fund. H.F.N. received support from National Institutes of Health Grant GM17129 and shared a North Atlantic Treaty Organization travel grant with R.A.G.

1. Fox, G. E. & Woese, C. R. (1975) *J. Mol. Evol.* **6**, 61–76.
2. Noller, H. F. & Woese, C. R. (1981) *Science* **212**, 403–411.
3. Zimmermann, R. A. (1980) in *Ribosomes: Structure, Function, and Genetics*, eds. Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. & Nomura, M. (Univ. Park Press, Baltimore, MD), pp. 135–169.
4. Noller, H. F. & Garrett, R. A. (1979) *J. Mol. Biol.* **132**, 621–636.
5. Garrett, R. A. & Noller, H. F. (1979) *J. Mol. Biol.* **132**, 637–648.
6. Peattie, D. A. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 4679–4682.
7. Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1760–1764.
8. Peattie, D. A. & Herr, W. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2273–2277.
9. Monier, R. & Feunteun, J. (1971) *Methods Enzymol.* **20**, 494–502.
10. Aubert, M., Scott, J. F., Reynier, M. & Monier, R. (1968) *Proc. Natl. Acad. Sci. USA* **61**, 292–299.
11. Hindennach, I., Kaltschmidt, E. & Wittman, H. G. (1971) *Eur. J. Biochem.* **23**, 12–16.
12. Chen-Schmeisser, U. & Garrett, R. A. (1977) *FEBS Lett.* **74**, 287–291.
13. Fox, G. E. & Woese, C. R. (1975) *Nature (London)* **256**, 505–507.
14. Hori, H. & Osawa, S. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 381–385.
15. Pace, B., Matthews, E. A., Johnson, K. D., Cantor, C. R. & Pace, N. R. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, in press.
16. Erdmann, V. A. (1981) *Nucleic Acids Res.* **9**, r25–r42.
17. Simoncsits, A. (1980) *Nucleic Acids Res.* **8**, 4111–4124.
18. Hori, H., Osawa, S., Murao, K. & Ishakura, H. (1980) *Nucleic Acids Res.* **8**, 5423–5426.
19. Hinnebusch, A. G., Klotz, L. C., Blanken, R. L. & Loeblich, A. R., III (1981) *J. Mol. Evol.*, **17**, 334–347.
20. Nazar, R. N., Matheson, A. T. & Bellemare, G. (1978) *J. Biol. Chem.* **253**, 5464–5469.
21. Luehrsen, K. R. & Fox, G. E. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2150–2154.
22. Hori, H., Osawa, S. & Iwabuchi, M. (1980) *Nucleic Acids Res.* **8**, 5535–5539.
23. Williamson, R. & Brownlee, G. G. (1969) *FEBS Lett.* **3**, 306–308.
24. Averner, M. J. & Pace, N. R. (1972) *J. Biol. Chem.* **247**, 4491–4493.
25. Gilbert, W., Maxam, A. & Mirzabekov, A. (1976) in *Control of Ribosome Synthesis*, eds. Kjeldgaard, N. C. & Maaloe, O. (Academic, New York), pp. 139–148.
26. Ogata, R. T. & Gilbert, W. (1979) *J. Mol. Biol.* **132**, 709–728.
27. Siebenlist, U., Simpson, R. B. & Gilbert W. (1980) *Cell* **20**, 269–281.
28. Ehrenberg, L., Fedorcsak, I. & Solymosy, F. (1976) *Prog. Nucleic Acid Res. Mol. Biol.* **16**, 189–262.
29. Luehrsen, K. R., Fox, G. E., Kilpatrick, M. W., Walker, R. T., Domdey, H., Krupp, G. & Gross, H. J. (1981) *Nucleic Acids Res.* **9**, 965–970.
30. Stahl, D., Luehrsen, K. R., Woese, C. R. & Pace, N. R., *Nucleic Acids Res.*, in press.
31. Fresco, J. R. & Alberts, B. M. (1960) *Proc. Natl. Acad. Sci. USA* **46**, 311–321.
32. Tinoco, I., Uhlenbeck, O. C. & Levine, M. D. (1971) *Nature (London)* **230**, 362–367.
33. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.
34. Wrede, P. & Erdmann, V. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 2706–2709.
35. Woese, C. R., Magrum, L. J., Gupta, R., Siegel, R. B., Stahl, D. A., Kop, J., Crawford, N., Brosius, J., Gutell, R., Hogan, J. J. & Noller, H. F. (1980) *Nucleic Acids Res.* **8**, 2275–2293.
36. Fiil, N. P., Friesen, J. D., Downing, W. L. & Dennis, P. P. (1980) *Cell* **19**, 837–844.
37. Gralla, J., Steitz, J. A. & Crothers, D. M. (1974) *Nature (London)* **248**, 204–208.
38. Weber, H. (1976) *Biochim. Biophys. Acta* **418**, 175–183.
39. Epstein, P., Reddy, R. & Busch, H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1562–1566.
40. Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. (1980) *Nature (London)* **283**, 220–224.
41. Zimmern, D. (1977) *Cell* **11**, 463–482.
42. Lee, F. & Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4365–4369.
43. Zurawski, G. & Yanofsky, C. (1980) *J. Mol. Biol.* **142**, 123–129.