Electron Microscopy of Secondary Structure in Partially Denatured *Escherichia coli* 16S rRNA and 30S Subunits†

Barbara K. Klein, Penny Forman, Yukio Shiomi, and David Schlessinger*

ABSTRACT: Loops observed in partially denatured 16S rRNA lie within three domains, each about 500 nucleotides long. The loops observed in the 5′ and central domains agree well with features of the model proposed by Woese et al. [Woese, C. R., Gutell, R., Gupta, R., & Noller, H. F. (1983) *Microbiol. Rev. 47*, 621–669]. The structure in the 3′ domain is more complex and variable but is still consistent with the model. Published psoralen cross-linking studies have reported one of the observed loops but have also identified loops other than those observed here or predicted by any secondary structure model. These loops are stabilized by increasing concentrations of Mg^{2+} ions and by bound ribosomal proteins. For example, protein S4 in LiCl core particles stabilizes a loop of 370 nucleotides which forms part of its putative binding site on rRNA. The loop structures are characteristic enough to permit an overall comparison of the most stable of the predicted and observed loops in 16S and 23S rRNAs. Both rRNAs show a stable 5′-terminal loop and a set of subterminal nested loops near the 3′ end.

Detailed models for the secondary structure of 16S and 23S rRNA[1] have been proposed (Woese et al., 1983; Stiegler et al., 1981; Zweib et al., 1981; Noller et al., 1981; Glotz et al., 1981). Base-paired stems in the models are inferred by comparison of rDNA sequences from different bacteria and by the relative protection of base-paired sequences from chemical or enzymatic attack. Each smaller stem and loop fall into one of three domains in 16S rRNA; each domain, about 500 nucleotides long, is delimited by a base-paired stem.

However, such an overall structure has not been directly demonstrated. In particular, the model is static, but reconstitution assays have long indicated that rRNA undergoes extensive conformational changes during ribosome formation (Traub & Nomura, 1969). Many discussions have also suggested conformational changes in rRNA during the cycle of ribosome function. Furthermore, the powerful experimental technique of psoralen cross-linking has detected a number of the small hairpins and stems predicted in the model, but some of the predicted long-range interactions have not been seen, and other long-range base pairings have been inferred that are not found in the model (Wollenzein et al., 1979; Wollenzein & Cantor, 1982; Thompson & Hearst, 1983a). Some cross-links even directly suggest alternate conformations in the rRNA (Zweib et al., 1981; Thompson & Hearst, 1983b).

It would therefore be useful to determine directly the degree to which major features of the model actually exist and are stable. Electron microscopy can easily see interactions that form loops or hairpins of length at least 100 nucleotides, and we have looked for stable features in that size range in partially denatured rRNA. Jacobson (1976) has inferred a pattern of loops in RNA from phage MS2, and we have already reported a set of loops observed in 23S rRNA that agrees with proposed secondary structure models (Klein et al., 1983). Here we extend the analysis to 16S rRNA and also examine LiCl core particles and 30S ribosomes prepared ("spread") for electron microscopy. The results again strongly support critical features of the secondary structure models and also suggest features of the relative roles of base-paired stems, Mg^{2+} ions, and r-proteins in stabilizing ordered rRNA.

Materials and Methods

Sample Preparation and Characterization. 16S rRNA. Escherichia coli 30S and 50S ribosomal subunits were prepared, and the 16S rRNA in the 30S particles was isolated, characterized, and stored as described previously for the 23S rRNA of the 50S ribosomes (Klein et al., 1983).

30S Subunits. Vacant couples were isolated by first incubating mid-log-phase cells for 30 min at 15 °C and then isolating subunits as described by Noll et al. (1972).

LiCl Core Particles. Vacant couples isolated as above were incubated at 4 °C in a buffer containing 10 mM Tris, pH 7.8, 10 mM $MgCl_2$, 0.1 mM EDTA, and 3 M LiCl (Homann & Nierhaus, 1971). This removed all but the most tightly bound r proteins (see text and Figure 7). The particles were then collected by centrifugation, resuspended in a buffer containing 10 mM Tris, pH 7.8, and 10 mM $MgCl_2$, and stored frozen in small portions.

Protein Content Determination. To determine which r-proteins remain bound in the LiCl core particles, one-dimensional SDS–polyacrylamide gels were run (Laemmli & Favre, 1973) and stained with silver (Wray et al., 1981). The fraction of proteins that remain bound to 30S subunits in the partial denaturing conditions used for electron microscopy was also estimated, by sedimenting 30S subunits through a buffer containing 50% formamide, 10 mM Tris (pH 8.0), 50 mM NaCl, and 0.3 mM $MgCl_2$. The pellet was resuspended in 10 mM Tris (pH 8.0), electrophoresed in 15% SDS–polyacrylamide gels, and again stained with silver. This determination was necessary because ribosomes partially denatured with formamide have not previously been examined (see Results and Discussion).

DNA Hybridization Probe. A DNA probe with a single-stranded end complementary to the 3′ end of 16S rRNA was constructed and hybridized to 16S rRNA by using a procedure similar to that used previously for 23S rRNA (Klein et al., 1983). The DNA probe was constructed from the plasmid pZZ101, a subclone of pKK3535 which contains approximately 90 nucleotides at the 3′ end of 16S (King & Schlessinger, 1983). The hybridization buffer contained 30% form-

† From the Department of Microbiology and Immunology, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110. *Received November 15, 1983.* Supported by National Science Foundation Grant PCM-80 17402 and NIH Postdoctoral Fellowship 5-F32-GM 08012 to B.K.K.

[1] Abbreviations: EDTA, (ethylenedinitrilo)tetraacetic acid; rRNA, ribosomal ribonucleic acid; Tris, tris(hydroxymethyl)aminomethane; SDS, sodium dodecyl sulfate; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

0006-2960/84/0423-3927$01.50/0

amide (v/v), 0.4 M NaCl, 0.1 M HEPES (pH 7.8), 1 mM EDTA, 16S rRNA (20 μg/mL), and DNA probe (50 μg/mL). After incubation for 1 h at 25 °C to permit hybrid formation, a portion of the reaction mixture was diluted 40-fold and brought to a final buffer concentration of 50% formamide, 10 mM NaCl, 10 mM Tris (pH 8), and 1.0 mM $MgCl_2$ for electron microscopy (see below).

Electron Microscopy. The sample preparation for electron microscopy was similar to that described earlier (Klein et al., 1983); the hyperphase contained 50% (v/v) formamide, 10 mM Tris (pH 8.0), 70 μg of cytochrome *c*/mL (Sigma), and the indicated concentrations of NaCl and $MgCl_2$. In some experiments the rRNA was spread on a hypophase of 1 mM Tris (pH 8.0). The concentration of Tris in the hyperphase was 20 mM for experiments with LiCl core particles and 30S subunits.

Minor modifications were also made in the procedures for data acquisition. Features of individual rRNA molecules were traced at a magnification of 250000×. All full-length molecules with a clear contour were traced. At levels of Mg^{2+} higher than 1 mM, some molecules, particularly those with more than four loops, were too complex; thus, the data at high Mg^{2+} might underestimate the average number of loops per molecule. Linear contours of the molecules were digitized by using a GTCO digitizing tablet interfaced directly to the VAX 11/780 computer.

Data Analysis. The procedures used for determining loop location domains and for identifying discrete loop sizes within these domains were the same as described earlier. Briefly, loop sizes and midpoint locations were identified from molecules that were oriented and scaled to a constant length. Histograms of loop midpoint locations were constructed and domain boundaries estimated (see Figure 5). Loop size histograms were then constructed for each domain, and the data were fit with a set of Gaussian peaks (see Figure 6A,B) to estimate the size of significant composite loops (Klein et al., 1983; Fraser & Suzuki, 1970).

In this work 552 molecules of 16S rRNA were analyzed, with a mean length of 0.53 ± 0.06 μm. From the known length of 16S (1542 nucleotides), the projection length observed is 3.5 Å/nucleotide. This projection length agrees with the value obtained earlier for 23S rRNA (3.4 Å/nucleotide) and with the value of 3.5 Å/nucleotide reported for 16S rRNA by Wollenzien et al. (1979). The apparent length in these cases could be an underestimate if there are hairpins or loops too small to be seen (<50 nucleotides). Some evidence for shortening of the rRNA was seen in our data for spread LiCl core particles (173 molecules) and 30S subunits (450 molecules) compared to 16S rRNA. The bound proteins led to a length slightly but significantly shorter by about 100 nucleotides, yielding a projection length of 3.3 Å/nucleotide rather than 3.5 Å/nucleotide.

While the molecular lengths obtained were reproducible between different sample preparations, the average number of loops per molecule obtained differed somewhat. For example, preparations and analyses 18 months apart gave as much as one more loop per molecule in the later experiments. No comparable variations were observed with 23S rRNA. Possible explanations for these differences are that the secondary structure in 16S rRNA may change during storage or may be particularly sensitive to the spreading conditions for electron microscopy. In either case, the number of loops per molecule would vary between experiments, especially in regions of the molecule where the structure appears to be more variable (as in domain III; see below). The structure and relative stability of different loops is generally clear, however (see Results).

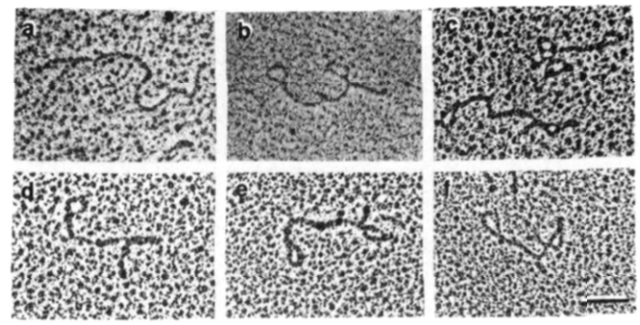

FIGURE 1: Loop patterns in rRNA. 16S rRNA in the following: 50 mM NaCl, 0.3 mM $MgCl_2$ (a); 50 mM NaCl, 1 mM $MgCl_2$ (b); 10 mM NaCl, 1 mM $MgCl_2$ (c). LiCl core particles spread in 50 mM NaCl, 0.3 mM $MgCl_2$ (d). 30S subunits spread in the following: 50 mM NaCl, 0.3 mM $MgCl_2$ (e); 10 mM NaCl, 0.5 mM $MgCl_2$ (f). The bar at the lower right is 0.1 μm long.

To compare the stability of features observed in the electron microscope with those predicted by the secondary structure models, the free energy of individual base-paired regions was calculated. The constraints of free energy values are the same as used earlier (Klein et al., 1983). The sources for the secondary structures were the following: Woese et al. (1983), Figure 1; Zweib et al. (1981), Figures 1a, 2 (top), and 3; Steigler et al. (1981), Figure 1.

Results

Partially denatured 16S rRNA shows a characteristic pattern of loops whose locations fall into three domains. In the following sections a detailed analysis of this loop structure is presented. In this analysis, the structural features of individual molecules are compared, and a single composite structure is constructed by incorporating these individual features. To complement the work on the rRNA, the loop patterns observed in spread partially denatured 30S ribosomal subunits and LiCl core particles were also studied. These samples provide a way to investigate the effect of r proteins on the stability of individual structural features in the rRNA.

Orientation of Observed Molecules. To analyze loop patterns of a large number of individual molecules, the molecules must first be oriented with respect to one another by means of recognizable structural features. Figure 1 shows a sample of electron micrographs of partially denatured 16S rRNA Figure 1a–c), spread LiCl core particles (Figure 1d), and spread 30S subunits (Figure 1e,f). Panels b, c, and f of Figure 1 exemplify how the characteristic loops in 16S rRNA tend to occur in three regions ("domains") along the contour of the molecule. One occurs near each end and a third in the middle. In general, elements of the same basic pattern were seen for all the species investigated, and over a range of ionic conditions. A small field of molecules is shown in Figure 2, indicating that many of the molecules have these characteristic features.

Loops tend to occur most frequently in the two external domains. A terminal loop (about 400 nucleotides long) is clearly seen in Figure 1 (panels b–f). At the opposite end of the molecule a set of loops 300–500 nucleotides long is often observed (again see panels b–f). Because these two loop locations occur frequently and are characteristic both in location and size, they were used to orient the molecules relative to one another.

To determine the absolute orientation of this loop pattern with respect to the known primary sequence of rRNA, a partially single-stranded DNA probe complementary to the

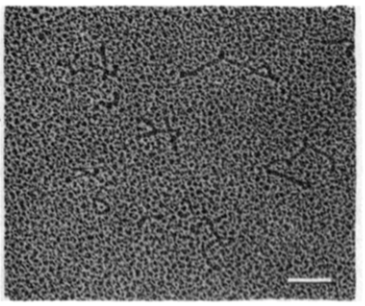

FIGURE 2: 16S rRNA molecules spread in 1 mM $MgCl_2$. A random portion of a large field is shown; the molecules show characteristic loop structures. The marker bar is long.

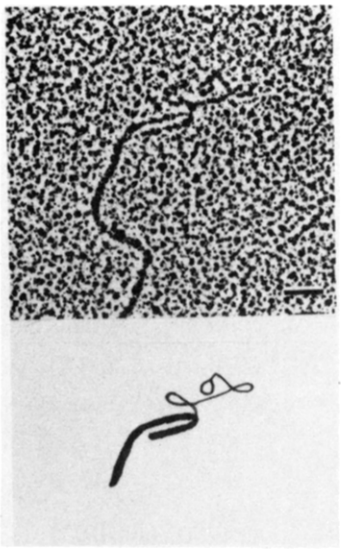

FIGURE 3: 16S rRNA hybridized to a DNA probe at its 3′ end. The thicker DNA portion and a loop pattern in the RNA showing a 5′-terminal loop and two other loops are shown. The bar is 0.1 μm long.

3′ end of 16S rRNA was hybridized to 16S rRNA. The hybrids were spread for electron microscopy by using hyperphase conditions which gave a characterisitc pattern of about three loops per molecule including a prominent terminal loop (10 mM NaCl, 1.0 mM $MgCl_2$). Figure 3 shows an electron micrograph of one such hybrid; note the terminal loop at the free end of the rRNA. The loop patterns of 25 hybrids were also analyzed in detail. The same three domains seen in isolated 16S molecules were evident, including a frequent terminal loop of about 400 nucleotides at the free (5′) end of the hybrid. This implies that the terminal loop occurs at the 5′ end of 16S rRNA, consistent with the inferred models of 16S rRNA structure.

Loop Frequencies, Sizes, and Locations in 16S rRNA. Linear representations of randomly chosen oriented molecules are shown in Figure 4 to indicate the information from which consensus loop structures were determined. In this figure a simple loop is represented by a bar; a complex loop (a loop or hairpin occurring within a larger loop) is represented by a thicker bar. Each individual loop identified in this analysis is characterized by two parameters, the size of the loop in nucleotides and the position of the loop midpoint along the contour of the RNA strand. Figure 5 graphs the relative frequency of loops with a given midpoint location for rRNA and ribosomes spread in different conditions. Because loops tend to occur in three domains along the contour of the rRNA strand (see above; Figure 1), loop midpoints also occur in the same three domains. Three of the individual data sets in

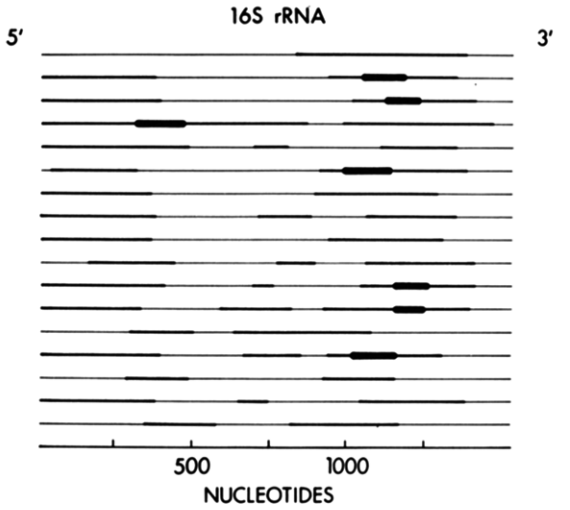

FIGURE 4: Oriented representation of 17 16S rRNA molecules spread in 1 mM $MgCl_2$. Thin bar: simple loop. Thick bar: complex loop, as defined in the text.

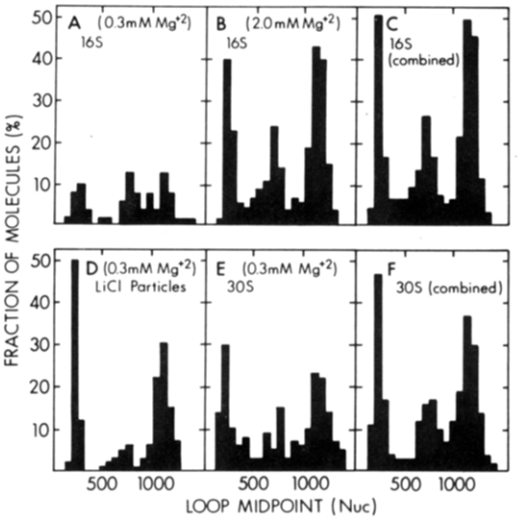

FIGURE 5: Loop frequency as a function of position in 16S rRNA. Panels A and B show data for 16S rRNA spread in 50 mM NaCl and 0.3 or 2 mM $MgCl_2$, respectively; panel C shows a composite for 1 and 2 mM Mg^{2+} data sets (listed in Table II). Panels D and E show spread LiCl core particles (D) or spread 30S ribosomes (E) in the same buffer as panel A. Panel F shows a composite for all data sets for spread 30S ribosomes (listed in Table II).

Figure 5 show a distinct three-domain structure (panel B, 16S rRNA in 2 mM $MgCl_2$; panel D, LiCl particles; panel E, 30S subunits) that is also evident in composite histograms combining the 1 and 2 mM Mg^{2+} data sets for 16S rRNA (panel C) or all data sets for 30S ribosomes (panel F). The 16S rRNA molecules spread in 50 mM NaCl–0.3 mM $MgCl_2$ (panel A) have less than 1 loop each on average, so that the three domain structures are blurred or absent.

The 16S rRNA at high Mg^{2+} concentrations (1 to 2 mM) was chosen for more detailed analysis, since it demonstrated a clear domain structure with multiple loops per molecule. The top panels of Figure 6 show histograms of the relative frequency of loops of different sizes in each of the three domains. For comparison, arrows at the top of each panel indicate the loop sizes predicted by the model of Woese et al. (1983).

Loops occur more frequently in domains I and III. Furthermore, the distribution of loop sizes differs significantly between domains. Domain I is dominated by a single 5′-terminal loop (panel A), whereas a large variety of loop sizes are observed in domain III (panel C). In domains I and II

Table I: Observed Loop Size and Locations

domain	16S rRNA[a] size	16S rRNA[a] location	LiCl core particles[b] size	LiCl core particles[b] location	30S subunits[b] size	30S subunits[b] location
I	110 ± 90	300 ± 120			160 ± 40	190 ± 70
I	370 ± 60	220 ● 90	370 ● 60	200 ± 40	360 ● 60	200 ± 40
II	110 ± 30	710 ± 90			120 ● 30	720 ● 90
II	190 ± 10	740 ● 70	180 ± 50	680 ● 50	200 ± 10	730 ● 90
II	280 ± 60	700 ± 60			270 ± 60	750 ± 80

[a] Loops observed in molecules spread in 1 or 2 mM $MgCl_2$. [b] Loops observed in molecules spread in 0.3 or 0.5 mM $MgCl_2$.

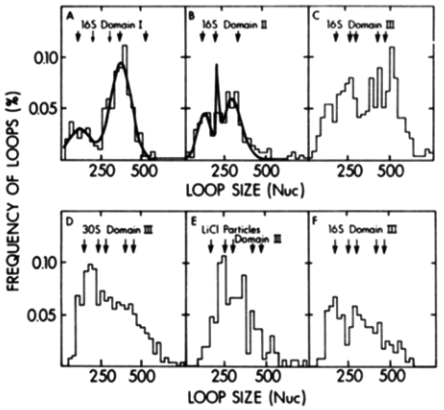

FIGURE 6: Loop frequency as a function of loop size in individual domains. Panels A–C, the three domains of 16S rRNA from molecules spread in 1 or 2 mM $MgCl_2$. In panels A and B, Gaussian fit curves are superimposed on the histograms. Panels D–F, domain III of spread 30S ribosomes, LiCl core particles, and 16S rRNA spread in 0.3 or 0.5 mM $MgCl_2$. The arrows in each panel represent the loop sizes predicted by Woese et al. (1983), as listed in Table III. Smaller arrowheads are used for loops predicted to be significantly less stable.

the pattern of loop sizes is readily evaluated by using the multiple Gaussian fit procedure described earlier (Fraser & Suzuki, 1979; Klein et al., 1983). Superimposed on the loop size histograms in Figure 6 (panels A and B) are examples of these Gaussian data fits that were used to infer the sizes of loops within the individual domains. The consensus loop pattern is given in Table I.

Domain I includes two loops, the more frequent of which is a terminal loop approximately 370 nucleotides long. The less frequent loop is smaller, about 150 nucleotides long, and is less well-defined. Domain II contains three loops of different sizes (110, 190, and 280 nucleotides). In general, the smallest and largest loops (110 and 280 nucleotides) tend to dominate.

The loop size pattern for domain III cannot be as readily interpreted by using this Gaussian fit program. As seen in Figure 6 (panel C) the distribution of loop sizes for the high Mg^{2+} data set does not appear to be discrete. Loop sizes appear to vary from 100 to 500 nucleotides, and the relative proportions of the different loop sizes differ substantially in different spreading conditions. As a result, when partial data sets or slightly shifted boundaries are used, the pattern of fitted loops changed significantly. A few loops are more prominent in particular conditions; for example, at high Mg^{2+} concentrations (Figure 6, panel C), a large loop of about 500 nucleotides is prominent, whereas at low Mg^{2+} concentrations (0.3 and 0.5 mM) (Figure 6, panel F), much smaller loops predominate. The loop pattern varies even more widely when r proteins are bound (see below). Because, in contrast to domains I and II, the Gaussian fits are not consistent for all conditions, no loops from domain III have been included in Table I.

As was observed for 23S rRNA (Klein et al., 1983), 16S rRNA shows a trend to more loops per molecule with increasing Mg^{2+} concentration. These results are shown in Table II for the 522 16S rRNA molecules used in this analysis. The levels of Mg^{2+} ions used were higher than in our earlier studies with 23S rRNA (Klein et al., 1983), because the structural elements in 16S rRNA appeared less stable.

Table II: Loops per Molecule in Molecules Spread in Different Conditions

$[Na^+]$ (mM)	$[Mg^{2+}]$ (mM)	16S N[a]	16S loops per molecule[b]	LiCl N[a]	LiCl loops per molecule[b]	30S N[a]	30S loops per molecule[b]
10	0.3	45	1.6 ± 1.0			101	3.0 ± 0.9
10	0.5	71	2.3 ● 1.0	68	2.3 ● 0.7	122	3.1 ● 0.8
10	1.0	59	2.7 ± 0.9				
10	2.0	71	3.4 ● 1.0				
50	0.3	52	1.0 ± 0.9	105	1.7 ● 0.9	115	2.0 ± 1.2
50	0.5	72	1.7 ± 1.0			120	2.5 ● 1.2
50	1.0	46	3.1 ± 0.9				
50	2.0	127	2.8 ± 1.2				

[a] Number of molecules measured. [b] The numbers are the average values ● one standard deviation.

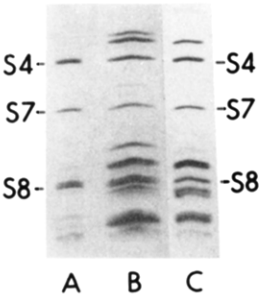

FIGURE 7: Gel electrophoretic analysis of ribosomal proteins bound in LiCl cores or in 30S ribosomes treated with spreading buffer. Lane A, LiCl core proteins from 30S ribosomes; lane B, proteins from 30S ribosomes; lane C, proteins from particles centrifuged through buffer containing 50% formamide (see Materials and Methods). Positions of proteins S4, S7, and S8 are indicated.

Loops in Spread LiCl Cores and 30S Ribosomes. Analyses of LiCl cores and 30S ribosomes suggest that r proteins remain bound to rRNA in the conditions of partial denaturation employed here. In preliminary experiments (data not shown), 30S ribosomes were doubly labeled with [^{14}C]uridine and [^{3}H]leucine to follow both their RNA and protein components. They were then mixed with 10 mM Tris, pH 7.4, and 1 mM Mg^{2+} or the same buffer containing 50% formamide and sedimented through replicate 10–30% sucrose gradients with or without 50% formamide. Most of the labeled protein continued to sediment with ribosomes even in buffers containing formamide and cytochrome *c* (as in the conditions used for electron microscopy). A small amount of protein was released from the ribosomes and sedimented much more slowly.

The retention of r proteins in spreading buffers was confirmed by gel electrophoresis. The 30S ribosomes were centrifuged out of spreading buffer containing 50% formamide and resuspended in 10 mM Tris, pH 7.4, and 5 mM $MgCl_2$. Figure 7 shows the gel electrophoretic pattern of proteins from these particles compared to those of proteins from LiCl core

particles and untreated 30S ribosomes. Because the hypochromicity of rRNA and ribosomes is the same (Schlessinger, 1960), the number of particles applied to the gel could be accurately assessed by their absorbane at 260 nm. The lanes of Figure 7 received proteins from LiCl particles containing 10 μg of rRNA (lane A) or proteins from 30S ribosomes (lane B) or formamide-treated ribosomes (lane C) containing 5 μg of rRNA. When the pattern of bands obtained was compared with published standards (Dijk & Litlechild, 1979), three r-proteins were found to remain in LiCl core particles: S4, S7, and S8 (cf. Robakis & Boublik, 1981). Densitometry confirmed that the formamide-treated ribosomes had lost some proteins but retained most of them with the 1:1 stoichiometry found in intact ribosomes. The LiCl core particles were more heterogeneous, containing about half the content of S4, S7, and S8 found in 30S ribosomes.

Ribosomes and LiCl core particles were then prepared for electron microscopy in formamide-containing buffers and the loop structures analyzed. The photographs in Figure 1d–f show sample molecules. For convenience, such molecules are referred to below as "spread 30S ribosomes" and "spread LiCl core particles", but we do not know whether the rRNA chains visualized still have r proteins bound or whether the proteins have been displaced during the preparation of grids for microscopy (see Discussion). In either case, r proteins bound in the particles at the time of spreading distinctly stabilized features of secondary structure in 16S rRNA. For example, in 50 mM NaCl and 0.3 mM $MgCl_2$ (Table II), where 16S rRNA showed 1.0 loop per molecule, spread LiCl core particles showed 1.7 loops and spread 30S ribosomes, 2.0 (or compare the frequencies of loops in Figure 5; the ribonucleoprotein particles spread in 0.3 mM Mg^{2+} have the loop frequencies seen in spread rRNA only at 2 mM Mg^{2+}).

The loop sizes and locations listed in Table I were determined, as described in detail for 16S rRNA above (see Figure 6, panels A–C). Spread 30S subunits, even at low Mg^{2+} levels (0.3–0.5 mM), showed all five loops seen in 16S rRNA, whereas in those conditions, spread LiCl core particles showed the selective stabilization of only the 370-nucleotide 5′-terminal loop (domain I). The loop size pattern in domain III was again more difficult to evaluate. The presence of proteins in the spread samples tended to favor some loops; spread 30S subunits showed a major loop of about 150 nucleotides (Figure 6, panel D), whereas 16S rRNA in high Mg^{2+} showed a prominent 500-nucleotide loop (see above). The spread LiCl core particles showed an even more disparate and narrower distribution, with the most frequent loops about 300 nucleotides long (see Discussion).

Discussion

The Most Stable Loops Are Consistent with Secondary Structure Models. Of eight loops larger than 90 nucleotides predicted in domains I and II by the model of Woese et al. (1983) (also, H. Noller, personal communication), five are seen in the electron microscope (Table III and Figure 6). The secondary structure model also predicts five loops in domain III ranging in size from 130 to 460 nucleotides, in excellent agreement with the range of loop sizes observed in the experimental data (100–500 nucleotides).

From this consistency, and the fact that loop patterns directly observed in 23S rRNA also agreed well with predictions (Klein et al., 1983), we assume that the structure observed in the electron microscope corresponds to that predicted by the secondary structure models. The agreement extends to the relative stabilities of the base-paired regions that form these loops in the model of Woese et al. (1983). Structures with the lowest free energies are located in domains I and III (Table III), consistent with the prominence of these domains (see high Mg^{2+} data set, figure 5, panel C). Furthermore, the most stable loop in each domain is the most frequently observed structure in that domain in the presence of high Mg^{2+}: the 360-nucleotide terminal loop in domain I and the 460-nucleotide loop in domain III (Figure 6, panels A and C).

Table III: Loops in the 16S rRNA Model of Woese et al. (1983)

domain	loop size[a] (nucleotides)	loop location (nucleotides)	$\Delta G°$ (kcal)[b]
I	110	180	−15.9
I	200	210	−4.2
I	300	210	−9.9
I	360	220	−18.5
I	520	292	−16.8
II	90	700	−12.8
II	180	670	−11.5
II	320	730	−15.0
III	130	1130	−17.1
III	230	1100	−18.0
III	280	1090	−18.4
III	400	1140	−11.0
III	460	1160	−21.7
I + II	900	467	−6.0

[a] All loops larger than 85 nucleotides are listed. The precise base-paired nucleotides are from Woese et al (1983). Two other recent models for secondary structure of 16S rRNA (Stiegler et al., 1981; Glotz et al., 1981) are very similar. [b] Values were estimated as under Materials and Methods.

Comparison with Previous Microscopy and Cross-Linking Studies. In a previous study, with no divalent ions added, Edlind & Bassel (1980) found four small hairpin loops, three of them similar in location to the smallest loops observed in each of the domains in the present work. The 5′-terminal loops were also seen, though one was somewhat shorter (270 nucleotides) and the other considerably longer (1160 nucleotides) than the ones observed here. In addition, a large open loop in the center of the molecule, not seen in the present work, was identified.

Another major source of information on higher order structure is psoralen cross-linking. Studying inactive 30S ribosomal subunits and 16S rRNA, Wollenzien (Wollenzien et al., 1979; Wollengien & Cantor, 1982) and Thompson (Thompson & Hearst, 1983a,b) have identified about 25 cross-links. Many of the psoralen cross-links detected correspond to the predicted secondary structure; two of these interactions are also reported here.

However, the nature of all the interactions preserved in cross-links has not been clear. Another class of psoralen cross-links involves long-range interactions of sequences from different domains, in disagreement with present models for secondary structures. These interactions are not seen regularly in the present work. The "unexpected" cross-links could arise from base-paired stems that only exist in alternate conformations of rRNA (Thompson & Hearst, 1983b). Alternatively, tertiary interactions may be cross-linked. The accessibility of proximal regions to psoralen may also affect the frequency of cross-links.

Mg^{2+} Ions Stabilize Loops in rRNA. Increasing the concentration of Mg^{2+} ions increases the average number of loops. Figure 5 illustrates the increase between 0.3 (panel A) and 2.0 mM Mg^{2+} (panel B). The 5′ and 3′ domains are more stable than the interior domain (see the comparison of the thermodynamic stabilities above), and loops frequently occur there at Mg^{2+} concentrations of 0.5–1.0 mM.

The trends with added Mg^{2+} ions and the consensus loop patterns in domains I and II were the same for all data sets.

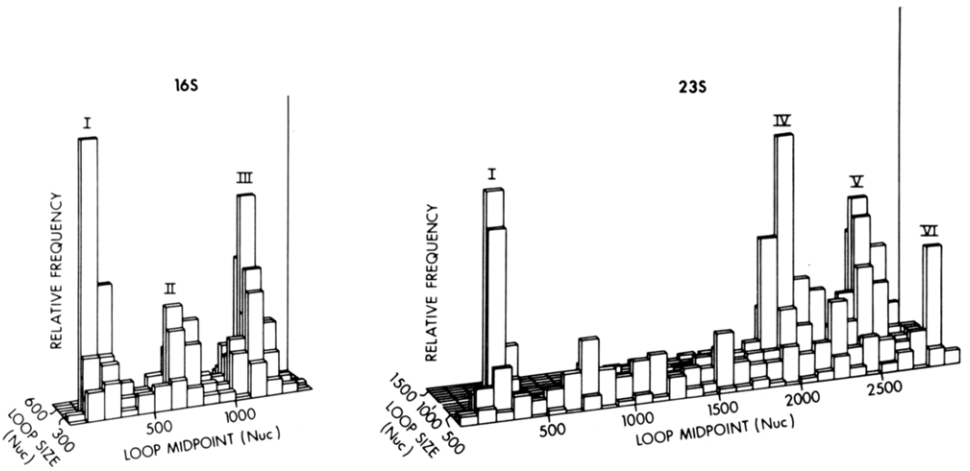

FIGURE 8: Relative frequency of loops plotted as a function of both size and midpoint location along 16S rRNA (left) or 23S rRNA (right). The total data from Tables I and II here and in Klein et al. (1983) are represented. Bin size 100 nucleotides.

However, the uniqueness of the structure is much less clear for 16S rRNA than that obtained previously for 23S rRNA. The loop pattern in domain III, in contrast to that of other domains, showed qualitative as well as quantitative differences in different preparations and at different Mg^{2+} concentrations (compare Figure 6, panels C and F, for the differential effect of Mg^{2+} concentrations on loops of different sizes; see Results). One loop appeared unique to the samples in high Mg^{2+}: the large loop (~500 nucleotides) which is prominent in domain III. This loop is also a prominent feature of domain III when the antibiotic neomycin is added to samples (work in progress). However, in other conditions, alternative sets of structural features were observed (see Figure 6 and below). Proteins and other ligands probably contribute to determining which features are more prominent.

Ribosomal Proteins Stabilize Particular Loops Preexisting in rRNA. More loops per molecule are seen in spread LiCl particles or spread 30S ribosomes compared to 16S rRNA (Table II). We do not know as yet whether these "spread particles" have retained r proteins or have become naked rRNA during grid preparation. Most proteins remain bound when ribosomes are treated with spreading buffers and formamide (see Results), but the proteins are too small to verify their continued presence in the molecules spread on grids (Cole et al., 1978). In any case, the proteins are at least bound long enough to give additional stability to features of secondary structures in the RNA. The greater number of loops per molecule observed when 30S ribosomes or LiCl particles are spread relative to free 16S rRNA is shown in Table II. In the same spreading conditions protein-free 16S rRNA shows few loops; spread LiCl core particles have two loops per molecule, and spread 30S subunits have even more loops per molecule.

Figure 5 demonstrates the striking effect of the LiCl core proteins on 16S rRNA stability, particularly in domains I and III (panels A and D). With spread 30S subunits (panel E), the loop frequency increases somewhat in domain II as well.

Ribosomal proteins thus seem to stabilize structural features of domains I and II that can also exist to some degree in 16S rRNA in their absence. However, in domain III, r proteins stabilize a small 150-nucleotide loop (Figure 6, panel D) more than the large 500-nucleotide loop favored by higher Mg^{2+} concentrations (panel C). In some experiments the large loop that predominates at high Mg^{2+} concentrations (Figure 6, panel C) was replaced by a pair of loops: one about 300 nucleotides long centered in domain III; the other, a small 150-nucleotide loop near the 3′ terminus. Therefore, in domain III of ribonucleoprotein particles, several alternative conformations could have comparable stability.

The specific effects of particular r proteins can be assessed more precisely by reference to their putative binding sites on 16S rRNA. Protein S4 is known to bind primarily to the 5′-terminal domain (Zimmerman, 1980). Two loops, one terminal of 371 ± 45 nucleotides and another of 464 ± 40 nucleotides, were observed by electron microscopy of 16S rRNA associated with protein S4 (Cole et al., 1978). From our data, the 360-nucleotide loop of domain I may be the primary site stabilized by S4 binding.

Earlier nuclease protection studies have shown that the 5′ domain of isolated rRNA is very resistant to digestion, producing a pattern of fragments similar to that obtained from rRNA complexed with S4 (Ehresmann et al., 1980). The results presented here, the same structure being stabilized by Mg^{2+} or in LiCl core particles (presumably by S4), strongly support the notion that protein S4 specifically stabilizes features of secondary structure that are already present in naked rRNA.

Another protein present in LiCl core particles, S8, binds to a single 65-base hairpin structure in domain II of 16S rRNA (Muller et al., 1979; Zimmerman, 1980; Wower & Brimacombe, 1983). The size of this feature is at the limit of resolution of this technique, but S8 binding may account for the slight decrease in the average length of the spread LiCl core particles and 30S subunits compared to the 16S rRNA.

The third protein prominent in LiCl core particles, protein S7, has been cross-linked to the 3′ domain, with the suggestion that this protein binds near three base-paired regions (corresponding to the 280-, 400-, and 460-nucleotide loops in the 3′ domain of 16S) brought into close proximity (Moller et al., 1978; Zweib & Brimacombe, 1979; Zimmerman, 1980; Wower & Brimacombe, 1983). Although domain III is the most difficult to analyze here, the results in Figure 6 (panel E) are consistent with those of this binding site for S7, with the observed stabilization of a 300-nucleotide loop and perhaps also of a 231-nucleotide loop.

16S and 23S rRNAs Show Structural Similarities. Figure 8 shows three-dimensional histogram representations of the most stable loops among 16S and 23S rRNA assessed by electron microscopy. In these histograms the relative frequency of a given loop is plotted as a function of both the loop midpoint position and the loop size. Because of the angle of perspective and scales used, the exact locations and Gaussian distributions of loops are less accurate than in Figures 5 and 6, but these displays show the overall domains as well as the

position, size, and reltive frequency of all observed loops in a single figure.

Some striking structural similarities of 16S and 23S rRNA are evident. For both molecules, the major loops are less than 600 nucleotides long and tend to occur near the ends of the molecules. Both 16S and 23S rRNAs have dominant 5′-terminal loops and also have a less frequent smaller loop within the terminal loop. In both cases there is another prominent domain near the 3′ end, and in contrast to the 5′ region, this domain contains a broad distribution of loop sizes for both RNA species. In both molecules, the 3′-end loops are all subterminal.

Although we have seen that loops can be differentially stabilized by factors like Mg^{2+} ions (Klein et al., 1983; unpublished results), a major determinant appears to be the strength of the stem at the base of each loop. For example, in each case one of the two predicted 5′-terminal loops dominates in micrographs, and in both cases, it is the loop with the lower predicted free energy. Also in both cases, the most stable loops in the entire predicted structures are in the outside domains (four out of four for 16S and three out of four for 23S), with the most stable free energies in the 3′ domain.

From Figure 8 one can imagine a possible evolutionary relationship of 16S and 23S rRNA, with additional internal sequences originally present in one of two duplicate genes. In subsequent evolution, sequences have diverged greatly; here and in other studies [for example, Tissières et al. (1959), Blair et al. (1981), and personal communication] 16S rRNA and 30S ribosomes show much more labile structure than that of 23S rRNA and 50S ribosomes. However, the 5′-terminal loops and other features have been retained, suggesting their functional significance. For example, might the 5′ loops in 16S and 23S rRNA somehow serve comparable roles at intervals during the cycle of ribosome function?

Added in Proof

Recent work by P. Wollenzein (personal communication) using electron microscopy of 16S rRNA cross-linked with psoralen shows several structural features that correspond well to those described here: a 360-nucleotide loop at the 5′ end, a 260-nucleotide loop in the center of the molecule, and five loops in the 3′ region ranging from 200 to 500 nucleotides in length.

Acknowledgments

We are indebted to Grady Phillips and Tom Rucinsky for sustained help with electron microscopy. Art Staden constructed the DNA probe used in Figure 3, and Monty Brandenberg developed the programs both to interface the digitizer with the VAX 11/780 computer and to draw the three-dimensional histograms of Figure 8.

Registry No. Mg, 7439-95-4.

References

Blair, D. P., Heilmann, L., & Hill, W. E. (1981) *Biophys. Chem. 14*, 81–89.

Brosius, J., Palmer, M. L., Kennedy, P. J., & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4801–4805.

Cole, M. D., Beer, M., Koller, Th., Strycharz, W. A., & Nomura, M. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 270–274.

Dijk, J., & Littlechild, J. (1979) *Methods Enzymol. 59*, 481–502.

Edlind, T. D., & Bassel, A. R. (1980) *J. Bacteriol. 141*, 365–373.

Ehresmann, C., Stiegler, P., Carbon, P., Ungewickell, E., & Garrett, R. A. (1980) *Eur. J. Biochem. 103*, 439–446.

Fraser, R. D. B., & Suzuki, E. (1970) in *Spectral Analysis* (Blackburn, J. A., Ed.) pp 171–211, Marcel Dekker, New York.

Glotz, C., Zweib, C., Brimacombe, R., Edwards, K., & Kossel, T. (1981) *Nucleic Acids Res. 9*, 3287–3306.

Homann, H. E., & Nierhaus, K. H. (1971) *Eur. J. Biochem. 20*, 249–257.

Jacobson, A B. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 307.

King, T. C., & Schlessinger, D. (1983) *J. Biol. Chem. 258*, 12034–12042.

Klein, B. K., King, T. C., & Schlessinger, D. (1983) *J. Mol. Biol. 168*, 809–830.

Laemmli, U. G., & Favre, M. (1973) *J. Mol. Biol. 80*, 575–599.

Moller, K., Zweib, C., & Brimacombe, R. (1978) *J. Mol. Biol. 126*, 489–506.

Muller, R., Garrett, R. A., & Noller, H. F. (1979) *J. Biol. Chem. 254*, 3873–3878.

Noll, M., Hapke, B., Schreier, M. H. & Noll, H. (1972) *J. Mol. Biol. 75*, 281–295.

Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R. R., Kopylov, A., Dohme, F., Herr, W., Stahl, D. A., Gupta, R., & Woese, C. R. (1981) *Nucleic Acids Res. 9*, 6167–6189.

Robakis, N., & Boublik, M. (1981) *Biochem. Biophys. Res. Commun. 103*, 1401–1408.

Schlessinger, D. (1960) *J. Mol. Biol. 2*, 92–95.

Stiegler, P., Carbon, P., Ebel, J.-P., & Ehresmann, C. (1981) *Eur. J. Biochem. 120*, 487–495.

Thompson, J. F., & Hearst, J. E. (1983a) *Cell (Cambridge, Mass.) 32*, 1355–1365.

Thompson, J. F., & Hearst, J. E. (1983b) *Cell (Cambridge, Mass.) 33*, 19–24.

Tissières, A., Watson, J. D., Schlessinger, D., & Hollingworth, B. R. (1959) *J. Mol. Biol. 1*, 221–233.

Traub, P., & Nomura, M. (1969) *J. Mol. Biol. 40*, 391–413.

Woese, C. R., Gutell, R., Gupta, R., & Noller, H. F. (1983) *Microbiol. Rev. 47*, 621–669.

Wollenzien, P. L., & Cantor, C. R. (1982) *Proc. Natl. Acad. Sci. (U.S.A.) 79*, 3940–3944.

Wollenzien, P., Hearst, J. E., Thammana, P., & Cantor, C. R. (1979) *J. Mol. Biol. 135*, 255–269.

Wower, I., & Brimacombe, R. (1983) *Nucleic Acids Res. 11*, 1419–1437.

Wray, W., Boulikas, T., Wray, V. P., & Hancock, R. (1981) *Anal. Biochem. 118*, 197–203.

Zimmerman, R. A. (1980) in *Ribosomes* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kohan, L., & Nomura, M., Eds.) pp 135–169, University Park Press, Baltimore, MD.

Zweib, C., & Brimacombe, R. (1979) *Nucleic Acids Res. 6*, 1775–1790.

Zweib, C., Glotz, C., & Brimacombe, R. (1981) Nucleic *Acids Res. 9*, 3621–3640.