Structure and topology of 16S ribosomal RNA. An analysis of the pattern of psoralen crosslinking

Charles R.Cantor, Paul L.Wollenzien and John E.Hearst

Barth Laboratory, Departments of Chemistry and Biological Sciences, Columbia University, New York, NY 10027, and Department of Chemistry, University of California, Berkeley, CA 94720, USA

Received 28 December 1979

ABSTRACT

All of the eleven psoralen crosslinking features previously mapped by electron microscopy of the 16S rRNA can be matched by energetically stable duplex regions found by a computer search of the sequence. One additional particularly stable duplex found in the sequence has subsequently been seen in the microscope. There is no indication that a best or unique fit of the sequence to the crosslinking data has been found. However, it is encouraging that all twelve assignments involve non-overlapping regions of the sequence which allows their simultaneous construction. The resulting general pattern of secondary structure is very different from previous suggestions for rRNAs or other complex RNAs. It is an RNA chain folded in space much like a typical protein chain.

Because of the interwound nature of double helices, base pairing between distant regions of the sequence might result in topologically knotted structures. However, examination of available electron microscopic data suggests that the 16S rRNA does not contain any knots.

1. INTRODUCTION

Psoralen crosslinking of *Escherichia coli* 16S rRNA free in solution¹ and packaged within the 30S ribosomal subunit² suggests the presence of numerous contacts between regions of the RNA distant in the primary structure. By analogy with psoralen reactions with more defined systems³ these contacts presumably represent double stranded helical regions. The quantitative results obtained on eleven crosslinked features suggest that most or all can occur simultaneously and that many of the same features seen in the free RNA are also present in the ribosome.

In this paper we explore the extent to which the available crosslinking data on the 16S rRNA, considered in concert with the known sequence^{4,5} and simple rules for the energetics of nucleic acid conformation⁶ restricts the choice of plausible secondary structure schemes. Each known crosslink, in principle, fixes two positions in the 16S rRNA so that with eleven crosslinks, the average stretch of RNA between two crosslinks would

be about the size of a tRNA. Depending on the length of the duplex region responsible for each crosslink the resulting single-stranded regions which must be searched for all inter and intra region pairing might approach practical levels. In fact, our analysis of 16S rRNA secondary structure is not nearly complete enough to propose a meaningful total secondary structure pattern. However, we will show that the general organization of the 16S rRNA appears to be totally different from the sorts of pictures that have been presented previously⁷. The schemes that emerge are of particular interest since they raise basic topological questions about the folding of RNA chains. We will show that 16S rRNA isolated from 30S ribosomal subunits is not a topologically knotted structure.

2. METHODS

(a) Electron Microscopy

Electron microscopy of HMT-crosslinked 16S rRNA was performed as described previously¹. Images shown here are selected but by no means unique examples of figures seen in the microscope. Previous work had already assigned the approximate location⁷ and absolute polarity⁸ of each of eleven crosslinked features. It is important to note that this was done, by design, in the absence of any direct knowledge of the correct nucleotide sequence. No sequence information was employed in making the original electron microscopic assignments.

(b) Sequence Analysis by Computer

A computer program which selects all possible base pairing schemes meeting preset criteria for length and G-C base pair content was employed. The 16S rRNA sequence used for this procedure was that of the *rrnB* gene of Brosius et al⁴. However, this sequence is so similar to that obtained by Carbon et al⁵ that it would not matter which sequence was used. The computer scan itself was also conducted without direct access to any crosslinking data. All structures which were four or more base pairs in length and having formation free energies of approximately -8kcal or more were listed for further examination.

(c) Comparison of Sequence and Crosslinking

The approximately 1700 base-paired regions listed in part (b) were scanned for all stable duplex regions which might correspond to a cross-linked feature, using as a guide the following principles:

- (i) Find the most stable long uninterrupted stretch of perfect base pairing that occurs between residues within +50 of a known approximate

crosslink position.

(ii) Attempt to extend each duplex found in (i) as far as possible by limited interruptions with short bulge or interior loops.

(iii) Compute the stability of the interrupted stretches in (ii) and keep only the most stable pattern or patterns for future consideration.

(iv) Repeat the above steps for each crosslink independently.

3. RESULTS

(a) A Match of the Sequence to Observed Crosslinking Positions

The nucleotide sequence of the 16S rRNA was examined by a computer program for regions of stable secondary structure. A complete scan of all possible perfect double-helical sections containing six or more G-C base pairs was performed for the entire 16S rRNA sequence. Examination of this scan neglecting loop closure entropies indicated 20 base-paired regions with free energies of formation of -20 to -25kcal and one (of nine contiguous G-C base pairs) with a free energy of -38.2kcal.

This complete search indicated that the RNA did not have a significantly greater potential for forming stable duplex regions than a random sequence. It also clearly indicated that there was only one duplex region so much more stable than average that there was a compelling reason to select it as a potential element of the real structure. Thus the obstacle to selecting the correct or the most stable intramolecular base pairing scheme for the 16S rRNA is a combinatorial problem of formidable magnitude. In principle, all simultaneous non-overlapping base pairing schemes must be evaluated. To try to circumvent this problem we asked whether the known locations of psoralen crosslinking sites would restrict the choice of possible secondary structures in a useful way.

Eleven psoralen-crosslinked features, which we designate by I, II, ..., XI, were located by electron microscopy with a typical standard deviation of 50 bases in the size of the loop or stems generated by each crosslink¹. An obvious question to ask is whether any or all of these features correspond to particularly stable or long double helical regions in the rRNA. Stable regions are likely targets because we observed that most crosslink yields were not seriously depressed by lowered salt or magnesium. Long regions are likely targets simply because they possess more potential psoralen binding and crosslinking sites.

Very few of the most stable base-paired features fall within or close to regions indicated by the psoralen crosslinking data. Those which do

cannot be extended significantly and are, for the most part, pre-empted by more stable alternative extended structures which do fit the psoralen data. The one exception is the -38.2kcal region (729-739; 924-933) which is close enough to feature XI to warrant further consideration. These observations suggest that a credible free energy of formation for any potential structure must therefore be more stable than -20 to -25kcal if it is to play a determinative role in the secondary structure of the molecule.

Potential psoralen crosslinking sites were then examined as described in Methods using more permissive criteria for the selection of base paired regions. For each nominal crosslink location in Table 1, regions of the sequence 50 bases on either side of the location were examined. Within the resulting 100 x 100 square of possible matched perfectly complementary sequences, only those interactions with at least one region having a formation free energy of -9 to -10kcal were given further consideration (except feature V which is only -8kcal). This insured that each potential secondary structure had at least one section with sufficient stability to

Table 1: Duplex Regions Which Might Correspond to Crosslinking Sites

Feature	Frequency (%) [*]	EM Nominal [†]	Complementary Sequences		Base Pairs
I	20.4 ± .1	965-1525	864-911	1437-1490	28
XI	17 ± 4	649-947	655-738	924-993	41
VI	14 ± 3	277-490	245-291	439-475	25
VII	14.0 ± .4	147-522	152-176	503-523	17
II	13 ± 1	531-1525	524-552	1496-1524	16
III	9 ± 1	0-676	80-94	632-652	11
V	7 ± 2	52-1523	9-18	1528-1536	8
IV	7.0 ± .2	0-237	20-63	203-235	24
IX	6.6 ± .3	1239-1435	1201-1241	1385-1413	22
X	4.5 ± .1	739-1370	784-825	1332-1381	23
VIII	~3	136-638	144-149	608-613	6
XII	2.5 ± .5	353-1344	334-354	1274-1314	19

** GU pairs not included

* As observed after crosslinking in reconstitution buffer

† As assigned by previous work¹

act as a nucleus.

In some cases like features VIII and X only one sequence which met the above requirements was found. The duplex region contained in this sequence was extended where possible through small bulges or interior loops to make the longest possible interrupted duplex stem likely to be an independently stable entity. Structures having at least three base pairs and containing two or more G-C pairs were used in selecting the optimal pattern of extension of the nucleation sequence. Final results are summarized in Table 1.

In the other cases (such as features II, IV, VI, VII, IX and XI) two or three sequences met the requirements for nucleation centers. Attempts were made to extend these as just described above and in each case one secondary structure pattern clearly emerged as the most stable (Table 1); the others were discarded.

A close fit to all the crosslinks except I and III was generated by this procedure. By broadening the search to include slightly more distant residues, unique regions of extensive complementarity were discovered which seem to bracket these two features.

In performing a search of the sequence two additional regions of extended complementarity were noted. These need not have any statistical significance a priori but several factors led us to give them some consideration. XI' was an energetically favorable long stem (561-585; 837-863) which corresponds exactly to feature XI with the 3' and 5' ends of the rRNA reversed. Although we are confident in our general assignment of XI, the feature is nearly symmetrical. Thus it is difficult to exclude, on experimental grounds, that some molecules in which we assign feature XI by microscopy, actually contain XI' instead. However, experiments reported by Müller, et al⁹ argue that this region of the rRNA is involved in stable local hairpin structure. On this basis we provisionally exclude feature XI'.

Sequence region XII had not previously been seen in the electron microscope, but it involved outstanding complementarity between regions not implicated in any of the above eleven features. Scanning fields of crosslinked 16S rRNA molecules revealed a significant number of molecules (2.5% of all molecules crosslinked in reconstitution buffer) containing a feature which corresponded to XII. Two examples are shown in Figure 1.

Given the wide latitude in comparing sequence and electron microscope data, the final match between them in these hypothesized hydrogen bonding regions is encouraging. This is best seen by comparing observed loop

Figure 1. Two 16S rRNA molecules that are consistent the crosslinking of feature XII. The polarity of the tails can be deduced from photographs such as those in the third column of Figure 4. The bar indicates 0.1 μ m.

sizes in the electronmicroscope with the range of possible loop sizes that could be generated by crosslinks within the complementary stems of each sequence region. As shown in Table 2, the fit is quite good; in some cases strikingly so.

An estimate of the thermodynamic stability of each of the postulated

Table 2: Sizes of Crosslinked Loops Predicted and Observed

Feature	Observed Loop Size	Predicted Loop Size Range*
I	559	526-626
XI	298	186-338
VI	213	148-230
VII	375	327-371
II	994	944-1000
III	676	538-572
V	1483	1510-1527
IV	237	140-215
IX	196	144-212
X	631	507-597
VIII	502	459-469
XII	991	920-980

* Expected for crosslinking in any duplex site of the stem closing this particular feature

secondary structure features is shown in Table 3. These estimates are based on the method of Tinoco et al⁶ and therefore strictly apply only to moderately high ionic strength at 25°C. The presence of magnesium ion in all the samples we have examined could have a profound effect at increasing the stability of duplex or other helical nucleic acid structures but there is no accurate way at present to estimate these effects. However, the values shown in Table 3 are presumably lower bounds. Given in the Table are both the net attractive energies of duplexes and any small unpaired loops as well as the total energies including loop closure entropies for the large loops contained in each feature. These are calculated by considering the closure of each loop independently and assuming all residues within the loop are single stranded. Thus the loop closure entropies are surely overestimates.

The total energies are a measure of whether each loop feature could form spontaneously and be stable as an independent unit. Clearly all the features shown in Table 3 are stable relative to the unfolded single strand except V which is marginal. However, this does not mean these features are stable relative to all possible alternate pairings of the two single stranded regions with other parts of the 16S rRNA.

The duplex energies are an estimate of the potential contribution of each feature to an overall folded 16S rRNA. Some of these energies are impressively large. All but two features, V and VIII, are stable enough

Table 3: Energetics of Loops Formation (kcal/mol)

Feature	ΔG_f^0 (not including loop entropy)	ΔG_f^0 (including loop entropy)
I	-39.6	-30.7
XI	-71.2	-63.2
VI	-46.8	-39.0
VII	-28.8	-20.3
II	-32.0	-20.6
III	-23.4	-14.4
V	-10.8	- 0.9
IV	-38.6	-30.8
IX	-42.8	-35.0
X	-37.8	-28.9
VIII	-19.0	- 8.2
XII	-34.0	-24.9

to compete effectively against the most stable smaller individual base paired segments found by the computer scan. It would be rather difficult to dissociate these features without resorting to powerful denaturants such as elevated temperatures or organic solvents. This is consistent with the observed stabilities of structures involved in psoralen crosslinking. All the features identified in Table 1 have many potential psoralen crosslinking sites. This is not surprising since half of all duplex binding sites for a psoralen can potentially lead to crosslinking.

(b) A General Picture of 16S rRNA Secondary Structure

The occurrence of the stable cores of duplex regions like those shown in Table 1 is not unexpected statistically. The extensions of these cores into larger duplex regions are probably not significantly better than expected in random sequences. Even the ability to match all of the crosslinked features with plausible secondary structure regions is not surprising given the uncertainty in the electron microscopic mapping of the crosslink positions. The 16S rRNA, like any random RNA sequence, has an extraordinarily large number of reasonably stable potential base paired regions.

The frequencies of most types of multiple crosslinked molecules seen in the electron microscope were consistent with the simultaneous occurrence of all of the sites which generate these features¹. This suggests that psoralen crosslinks are sampling a relatively homogeneous set of molecular structures. Although heterogeneity, particularly at the 3' end, cannot be excluded by our data, we will make the simplifying assumption that any molecule may potentially contain all of the duplex regions responsible for psoralen crosslinking. When the 16S rRNA sequence was examined it was impressive to find that most features involved residues completely distinct from those in any other features. A few overlap at their edges in such a way as to make almost continuous sections of double strand with just a slight ambiguity in how to assign the points of contact at various duplexes.

The probability of finding by chance such an extensive set of complementary sequences which can form simultaneously and match all observed crosslinks seems small. However, there is no reason to expect that the structures shown in Table 1 are the best fit to the psoralen crosslinking data or the most stable set of secondary structures. We shall make no such claims. However, we can learn quite a bit about possible 16S rRNA structures from the general nature of the fit between the sequence and

the psoralen data.

The secondary structure regions indicated by psoralen crosslinks cannot be perfect long duplexes because the sequence allows no possibility for such regions in the 16S rRNA. However, large sections of secondary structure can form by interrupted duplex regions. All of these sections appear to be able to form simultaneously as judged by both the statistics of psoralen crosslinking and the sequences used to match each crosslink. A schematic structure with all twelve of the duplexes assigned to crosslinks is shown in Figure 2. It is useful to examine the general features of this structure while remaining aware that specific details are not necessarily correct.

The structure contains a convoluted RNA chain folding pattern. The presence of the long duplex stems condenses the RNA chain into a volume only two to three times that of the 30S particle. The overall appearance of the structure is like that of a typical tracing of the backbone of a large

Figure 2. A schematic picture of the 16S rRNA containing duplex regions assigned to all twelve mapped psoralen crosslinks. The molecule is drawn in its simplest possible topology though more complex arrangements cannot be excluded. Note that all duplexes are interrupted with loops. The lengths shown schematically correspond roughly to the total number of base pairs.

globular protein. Like most protein structures the chain folding is somewhat difficult to visualize. Note, however, that the 3'-terminal hundred residues appear to suture widely spaced regions of the molecule together by forming features I, II and V.

The picture of the 16S rRNA shown in Fig. 2 is only an approximation. Surely wrong in some details, it is still likely to be a fair qualitative representation of the structure of 16S rRNA in solution. Folding in the general way shown in Fig. 2 explains a number of observed properties of the 16S rRNA. The folded structure is condensed into a small volume. This is consistent with recent hydrodynamic measurements of the 16S rRNA¹⁰. Other electron microscopic studies have indicated that 30S ribosomes heavily depleted in proteins still retain an overall morphology similar to the intact particle¹¹. When prepared under suitably mild conditions even the bare 16S rRNA itself may retain a structure similar in appearance to the 30S particle¹². Such a folded structure implies that, in large part, the RNA can serve as an organizing element of the ribosome. During assembly, proteins may bind to pre-existing sites on the RNA. The similarity of most psoralen crosslinking sites in the 16S rRNA and 30S ribosome² suggests that protein binding does not change the organization of the structural domains of the 16S rRNA. The complex linkages postulated in Fig. 2 may also explain some of the complex patterns of nucleic acid protein interaction seen when detailed RNA binding sites of individual proteins are mapped. For example, protein S4 binds to several distant regions of the 16S rRNA sequence¹³. These regions appear to be able to form a stable tertiary structure core¹⁴.

(c) Possible Knotted Topologies in 16S rRNA

The nature of the path traced out by the 16S rRNA backbone is similar to that of a folded protein but there is one potential difference of fundamental importance. Two sections of a protein chain that form extensive β sheet contacts can do so without necessarily becoming topologically interwound. There is a tendency for β sheets to twist and, in principle, strands involved in long β sheets could become topologically knotted. However, in practice only a single knotted topology has been found in all the proteins of known three-dimensional structure¹⁵.

The topology of nucleic acid double helices is potentially much more complicated. For RNA, every eleven consecutive base pairs will lead to one helix turn and thus one potential topological link. The structural consequences of this linkage to circular double-stranded DNA are well ex-

explored both theoretically¹⁶⁻¹⁹ and experimentally²⁰⁻²². Some theoretical considerations of linkage problems in RNA structure have also recently been described²³.

The pattern of crosslinks seen in the 16S rRNA and the sequence matches found for these crosslinks raise some potentially serious topological consequences for linear RNAs. The problem can best be appreciated by viewing some of the pairs of features which contribute to the structure of the 16S rRNA. If double helices had no twist, the two structures shown in Fig. 3a could simply be pulled apart at their ends into linear chains. However, if each crosslinked feature corresponds to one or more net helix turns then each structure is, topologically, a knot (Fig. 3b). Pulling at the two ends will extend most of the molecule but will leave a bolus of RNA trapped somewhere internally.

In fact, when the number of base pairs assigned to each crosslinked feature is considered, there is a potential for forming one or more helix turn in almost every case. Therefore, a real possibility exists that 16S rRNA in solution, and in the ribosome, is a topologically knotted structure. Here we explore the structural evidence that knots may actually exist; the biological implications will be given in the discussion section.

A single covalent linear strand, given the freedom to move, can always untie any knot, in contrast, a circle, containing any knots within it, must remain knotted unless the chain is broken. The implications of this for electron microscopy of crosslinked nucleic acids are considerable. Subjected to rigorous formamide denaturation conditions, an uncrosslinked RNA can unknot and untwist from any noncovalent structure to become a linear molecule. The presence of crosslinks can alter this result. Suppose a psoralen crosslink occurs in a duplex region of one of the molecules shown in Fig. 3. Upon denaturation the second duplex region in each case will unwind leaving a simple circular molecule with two single stranded tails. Many such molecules have been seen electron microscopically¹.

Suppose, instead, that both duplex regions of a structure in Fig. 3 become crosslinked. Denaturation and spreading will then lead to a structure resembling a letter theta. Observation of such a structure in the electron microscope indicates that the potential for knotted topology exists but it does not confirm the actual presence of a knot. The ladder-like duplexes shown in Fig. 3a, even if they did not possess one or more helical turns per ladder, would lead to a psoralen crosslinked "theta" structure.

Figure 3. Pairs of duplexes which can lead to knotted topology.

- (a) Schematic of molecules with features IV and VII, and I and XI.
 (b) Topologically equivalent if each duplex region contains a single helix turn.

When fields of crosslinked 16S rRNA molecules were observed in the electron microscope, a significant number of theta structures were seen. Typical individual molecules are shown in Figure 4. These involve pairs of features such as IV and VII, and II and III. A statistical analysis of the frequencies of observed theta figures was carried out. The frequencies of eleven single crosslinked features (I through XII, excluding VIII) were determined by examining molecules crosslinked in reconstitution buffer. The results are shown in Table 1. If a theta figure arises simply from two single crosslinks, the expected frequency is just the product of the frequencies of the two crosslinks observed individually. If, on the other hand, some theta figures arise as the result of knots trapped in the molecule by one or more crosslinks, then a higher frequency is expected. The expected frequency of total theta figures in a knotted and unknotted

Figure 4. Typical examples of theta features seen in psoralen crosslinked 16S rRNA molecules. The bar indicates 0.1 μ m.

topology and the observed frequency of theta figures are compared in Table 4. There is excellent agreement between values calculated for an unknotted topology and measured values. This provides strong support for the relative positions of the psoralen crosslinks. It argues that, although the crosslink positions raise the potential of knotted conformations, in practice few if any knots exist in the molecule.

As a second test, consider what will happen if a crosslink is formed between two residues which flank a region of the RNA involved in a knot. Now the knot is contained inside a covalent circle. Quite a few such features are seen, especially in molecules containing very large loops. For example, molecules with feature V frequently contain additional looped structures. A few selected molecules of this type are shown in Figure 5. The extra loops are either additional psoralen crosslinked features or trapped knots. To decide between these two possibilities we computed the

Table 4: Tests of Knotted 16S rRNA Topology

Frequency (% of All Molecules)			
Molecules Observed	Expected for Unknotted Topology†	Expected for Knotted Topology†	Observed
Total Theta Figures	11.6± 0.5	>11.6± 0.5	9.9± 0.3
Feature V Molecules Which Contain Theta Figures	0.8± 0.2	> 7.0± 2	0.8± 0.2
Features I,VI,and XII or I,VII,and XII	0.14± 0.1	6.0± 1	0.05± 0.05*

†Expected frequencies are computed from values of the frequencies of individual crosslinks seen in this sample. All values are averages of two independent tabulations.

*First tabulation: 0 of 274 molecules; second tabulation: 1 of 944 molecules.

frequency of molecules expected to contain feature V and, simultaneously, pairs of other crosslinks which generate a theta feature, for either an unknotted topology or a knotted topology. The results shown in Table 4 indicate that there is no need to invoke knots to explain the extra loops in molecules containing feature V.

A third test for knots was performed by looking for molecules which might contain double theta features formed by the simultaneous occurrence of features I, VI and XII or I, VII and XII. Such molecules are shown schematically in Figure 6. If the central feature (XII) really contains a net topological link, then all molecules containing the pairs of features I and VI or I and VII would also have to contain feature XII. However, in practice, as shown in Table 4, only one molecule in more than a thousand contains features I, VI and XII or I, VII and XII. This frequency is consistent with the simultaneous crosslinking of three features. Many molecules were found containing only the pair of features I and VI or I and VII. This test specifically rules out the possibility of knots involving feature XII. The overall impression from this statistical analysis is that knots do not exist in the 16S rRNA.

Figure 5. Additional loops seen in molecules containing the end-to-end crosslink feature V.

4. DISCUSSION

(a) How are knots avoided?

While the 16S rRNA isolated from the 30S subunits appears to have no knots, an inspection of the sequence shows a considerable potential for forming base-paired sections which would lead to knotted topologies. Two aspects of rRNA primary structure may reflect mechanisms that are used to avoid forming knots. For a structural rRNA with a high overall extent of base pairing the paucity of perfect duplex complementary sequence regions is really amazing. However, well-matched duplex regions are very stable structures and thus represent severe risks of trapping metastable folding intermediates. Details of two possible features, VI and X are given in Figure 7. These are representative of all of the others. Note that they

Figure 6. An example of a structure which can be used to test for knots. If a knot exists, crosslinking the outer duplexes will trap the inner one and molecules on the left should not be seen in the microscope. The molecule on the right is also an example of a structure which can only be formed in certain ordered pathways. The center duplex cannot be formed last. Note that the length of the outer duplexes is irrelevant. A single base pair is sufficient to establish the topological constraint so long as the inner duplex consists of a full helix turn.

contain highly interrupted duplex regions. Short stretches of base pairs will not make a full helix turn.

By forming a complex structure from very short blocks of base pairs the 16S rRNA facilitates a trial and error searching of duplex regions

Figure 7. Two examples of sequence regions used to match crosslinked features.

during assembly. In addition, by a much more direct mechanism, the short lengths of duplexes in the final folded structure can minimize the topological linking. Blocks of 4 to 6 base pairs produce only half a helix turn, and intervening loops could easily unwind this half turn to make the net linking zero. In this way a long stretch of sequence could form many base pairs without ever forming a knot. A similar argument has recently been introduced to explain structures seen in some single strand viral DNAs²⁴.

Young and Steitz have recently shown that nucleotide sequences in the precursor of the 16S rRNA just outside the structural gene are complementary and form a nearly perfect 40 base pair duplex²⁵. This extremely stable structure effectively renders the precursor of the 16S rRNA structural gene a covalent circle. It prevents any change in the topological linking of the 16S rRNA structural gene until the long duplex is cleaved during RNA processing. After cleavage, the remaining duplex (feature V) is sufficiently weak that, even in the presence of all other structures, it could open to allow changes in structure (Table 3). Thus the design of the precursor RNA may serve to help regulate topological aspects of assembly.

(b) Implications for Ribosome Structure

Psoralen crosslinking results provide a very different picture of the 16S rRNA than that held previously. Instead of an amorphous series of hairpins, the 16S rRNA can fold into an organized globular structure. As such it may really be the framework for the structure of the ribosome. Most proteins may simply bind to pre-existing RNA structure. Evidence for such an RNA core is now available both from hydrodynamic measurements¹⁰ and electron microscopy¹².

The model illustrated schematically in Figure 2 is consistent with much of what is known about the 30S ribosome. Given the potential complexity of secondary structure in large RNAs the overall picture which has emerged thus far is a surprisingly simple one. Yet it leads to a highly condensed RNA molecule.

Several of the crosslinked features near the 3' end, including II and V, cannot be seen in the active 30S particle². This raises the possibility that within the ribosome significant RNA structural changes may be possible and may even have functional significance. Since the initiation site recognition region²⁶ of the 16S rRNA is partly base paired in the model of Fig. 2, a structure change in this region upon ribosome activation is sensible.

ACKNOWLEDGEMENTS

This work was supported by grants from the NIH to C.R. Cantor (GM 19843), to J.E. Hearst (GM 11180) and an NIH postdoctoral fellowship to P. Wollenzien (AI10061). Additional support was obtained from the American Cancer Society, grant NP 185 to J.E. Hearst. We thank Dr. Pallaiah Thammana for helpful discussions, and Dr. Ann Jacobson for help with the electron microscopy.

PUBLICATIONS

1. Wollenzien, P.L., Hearst, J.E., Thammana, P. and Cantor, C.R. 1979. J. Mol. Biol. in press.
2. Thammana, P., Cantor, C.R., Wollenzien, P.L. and Hearst, J.E. 1979. J. Mol. Biol. in press.
3. Isaacs, S.T., Shen, C.J., Hearst, J.E. and Rapoport, H. 1977. Biochemistry 16, 1058-1064.
4. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. 1979. Proc. Natl. Acad. Sci. 75, 4801-4805.
5. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.P. 1978. FEBS Lett., 94, 152-156.
6. Tinoco, Jr., I., Borer, P.N., Dengler, B., Levins, H.D., Uhlenbeck, O.D., Crothers, D.M. and Gralla, J. 1973. Nature New Biol. 246, 40-41.
7. Ehresmann, C., Stiegler, P., Mackie, G.A., Zimmermann, R.A., Ebel, J.P. and Fellner, P., 1975. Nucleic Acids Res. 2, 265-278.
8. Wollenzien, P.L., Hearst, J.E., Squires, C.L. and Squires, C.L. 1979. J. Mol. Biol., in press.
9. Muller, R., Garret, R.A., and Noller, H.F. 1979. J. Biol. Chem. 254, 3873-3878.
10. Allen, S.H. and Wong, K.P. 1978. J. Biol. Chem., 253, 8759-8766.
11. Vasiliev, V.D., Koteliensky, V.E., Shatsky, I.N. and Rezapkin, G.V. 1977. FEBS Lett. 84, 43-47.
12. Vasiliev, V.D., Selivanova, O.M. and Koteliensky, V.E. 1978. FEBS Lett. 95, 273-276.
13. Cole, M.D., Beer, M., Ko-ler, T., Strycharz, W.A. and Nomura, M. 1978. Proc. Natl. Acad. Sci. 75, 270-274.
14. Garrett, R.A., Ungewickell, E., Newberry, V., Hunter, J. and Wagner, R. 1977. Cell Biol. Internat. Reports 1, 487-502.
15. Richardson, J. 1977. Nature 268, 495-500.
16. Glaubiger, D. and Hearst, J.E. 1967. Biopolymers 5, 691-696.
17. Fuller, F.B. 1971. Proc. Natl. Acad. Sci. 68, 815-819.
18. Fuller, F.B. 1972. Rev. Roum Math. Pures Appl. 17, 1329-1334.
19. Crick, F.H.C. 1976. Proc. Natl. Acad. Sci. 73, 2639-2643.
20. Vinograd, J., Lebowitz, J., Radloff, R., Watson, R. and Lacpis, P. 1965. Proc. Natl. Acad. Sci. 53, 1104-1111.
21. Liu, L.F., Depew, R.E. and Wang, J.C. 1976. J. Mol. Biol. 106, 439-452.
22. Bauer, W.R. 1978. Ann. Rev. Biophys. Bioengin. 7, 287-313.
23. Studnicka, G.M., Rhan, G.M., Cummings, I.W., and Salser, W.A. 1978. Nucleic Acids Res. 5, 3365-3387.
24. Stettler, U.H., Weber, H., Koller, T., and Weissmann, C. 1979. J. Mol. Biol. 131, 21-40.
25. Young, R.A. and Steitz, J.A. 1978. Proc. Natl. Acad. Sci. 75, 3593-3597.
26. Steitz, J.A. and Jakes, K. 1975. Proc. Natl. Acad. Sci. 72, 4734-4738.