Sites of Contact of mRNA with 16S rRNA and 23S rRNA in the *Escherichia coli* Ribosome†

Paul Wollenzien,*,‡ Alain Expert-Bezançon,§ and Alain Favre§

Edward A. Doisy Department of Biochemistry and Molecular Biology, St. Louis University Medical Center, 1402 South Grand Boulevard, St. Louis, Missouri 63104, and Groupe de Photobiologie Moléculaire, Institut Jacques Monod, CNRS-Université Paris VII, 2 Place Jussieu, 75251 Paris, France

Received July 9, 1990; Revised Manuscript Received September 17, 1990

ABSTRACT: The locations of close encounter between ribosomal RNA (rRNA) and messenger RNA (mRNA) were determined by photochemical cross-linking experiments that employ an artificial mRNA, 51 nucleotides long, containing 14 U residues that were randomly substituted by 1–4 4-thiouridine (s^4U) residues. The mRNA was bound to 70S ribosomes or 30S subunits and then was irradiated at 366 nm to activate cross-linking between the s^4U residues and rRNA. Cross-linking occurred to both 16S rRNA and 23S RNA. The rRNA was then analyzed by a series of reverse transcriptase experiments to determine the locations of cross-linking. Twelve sites in the 16S rRNA and two sites in the 23S rRNA have been detected. In the 16S rRNA, two of the sites (U1381, C1395) are in the middle part of the secondary structure close to position C1400, and the remaining sites (G413, U421, G424; A532; G693; U723; A845; G1131/C1132; G1300; G1338) are distributed between six regions that are peripheral in the secondary structure. In the 23S rRNA, one site (U1065) is located in the GTPase center close to A1067, the site of thiostrepton-resistance methylation in domain II, and the other site (U887) is located a short distance away also in domain II. The distribution of these rRNA sites in the ribosome specifies an mRNA track that is consistent with other information. In addition, some of the contact points represent new constraints for the three-dimensional folding of the rRNA.

The arrangement of the mRNA in the ribosome is critical in the initiation step of protein synthesis as well as in the decoding process that occurs during the elongation phase (Dahlberg, 1989). Evidence that there is a specific binding site for the mRNA comes from several types of experiments. Titration experiments indicate that mRNA specifically binds to 30S subunits or 70S ribosomes in a 1:1 stoichiometry (Katunin et al., 1980). Affinity-labeling experiments identify proteins S1, S3, S4, S5, S7, S9, S12, S18, and S21 as being in contact with the mRNA (Cooperman, 1980; Liljas, 1982; Spirin & Vasiliev, 1989; Vladimirov et al., 1990). The distribution of these proteins indicates that the mRNA is situated in the middle part of the 30S subunit in the indentation between the body (the lower two-thirds) and head (the upper one-third) of the 30S subunit. This placement is consistent with the localization of the tRNAs on the ribosome as determined by affinity-labeling experiments and immunoelectron microscopy (Ofengand et al., 1986). Recently, Stade et al. (1989) used a small synthetic mRNA that contained a residue at a unique location for a photoaffinity reagent. When this mRNA was used to form a complex with the 70S ribosome in the presence of cognate tRNA in the P site, protein S7 was cross-linked to a significant extent. This indicated that this protein was close and situated 5′ to the codon–anticodon decoding site.

The details of the track of the mRNA across the rRNA will tell us a great deal more about the arrangement of the mRNA in the ribosome, the order and location of the tRNAs during decoding, and the arrangement of the rRNA in the ribosome. The interaction between the Shine–Dalgarno sequence of mRNA and the complementary sequence in the 3′-terminal region of the 16S rRNA during translation initiation establishes one rRNA site contacted by mRNA (Gold et al., 1981). Stade et al. (1989) demonstrated this interaction in their photoaffinity experiments. Three other affinity experiments have been done to determine the location of additional sites contacted by the mRNA in the ribosomal RNA segments. Wagner et al. (1976) used a glyoxylyl derivative of the trinucleotide U3 to form covalent attachments to two sites in the 16S rRNA at positions G462 and G474. Stiege et al. (1988) photo-cross-linked poly(A) to 16S rRNA and determined the location of one cross-linking site to be in the region from 1394 to 1399. In these cases, the identification of the rRNA cross-linking site was done by isolation of fragments and sequence analysis (Wagner et al., 1976; Stiege et al., 1988). Graifer et al. (1990) have recently reported the chemical cross-linking of a derivative of hexauridylic acid to human placental 80S ribosomes. They concluded that a region in the 18S rRNA was the main site of cross-linking; the region corresponds in the *Escherichia coli* 16S rRNA to the interval between nucleotides 703 and 784 in the middle RNA domain.

To determine if there are additional sites in the rRNA in contact with mRNA, we have undertaken a new cross-linking experiment that uses an mRNA substituted with the intrinsic photoaffinity probe 4-thiouridine (s^4U).[1] 4-Thiouridine is a close analogue of U, which can easily be incorporated both in vivo and in vitro into RNA chains (Favre, 1990). s^4U was already used to analyze the conformation of tRNA form I and form III (Lemaigre Dubreuil et al., 1986). It is highly photoreactive, forming covalent adducts with nucleic acids and proteins, and it can be activated with 365-nm light, a wave-

† This work was supported by NATO (0109/88 to A.E.-B. and P.W.), the Université Paris VI (Grants MRT 88.T.0501 and ANRS 88N60/0823 to A.F.), and the National Science Foundation (DMB 89-04500 to P.W.).

* Address Correspondence to this author.
‡ St. Louis University Medical Center.
§ Institut Jacques Monod.

[1] Abbreviations: s^4U, 4-thiouridine; Tris, tris(hydroxymethyl)aminomethane; EFG, elongation factor G; EFTu, elongation factor Tu; IF, initiation factor.

0006-2960/91/0430-1788$02.50/0

length at which there is no other photochemistry in the usual bases (Hajnsdorf et al., 1986). The sites of cross-linking in the rRNA are determined by reverse transcription analysis. Photoinduced modifications (Denman et al., 1988) and photochemically induced modifications (Ericson & Wollenzien, 1988) cause alterations in the bases that are complete blocks to the enzyme reverse transcriptase. Since this approach does not necessitate the isolation of cross-linked complexes for sequence analysis, it should be much more sensitive in detecting the locations of cross-linking sites. In addition, since reverse transcriptase synthesizes up to the location of the cross-linking site, it indicates precisely the location of nucleotide involved in the contact with mRNA.

The locations of 12 sites in the 16S rRNA and 2 sites in the 23S rRNA contacted by the mRNA have been determined with this procedure. This information indicates part of the path of the mRNA through the ribosome. The distribution of these sites with respect to known functional regions of the rRNA will be discussed.

Materials and Methods

Materials. Nucleotides, deoxynucleotides, and dideoxynucleotides were purchased from Pharmacia. DNA oligonucleotides were synthesized on an Applied Biosystems DNA synthesizer and were gel purified before use. Restriction nucleases were purchased from New England Biolabs. DNase I was from Promega. AMV reverse transcriptase (type 007) was purchased from Life Sciences. Sequenase reagents for DNA sequencing and T4 polynucleotide kinase were purchased from U.S. Biochemicals. $tRNA^{Phe}$ and $tRNA^{Tyr}$ were purchased from Boehringer Mannheim and [^{14}C]Phe was purchased from Amersham. tRNA was charged according to Traub et al. (1971). [^{32}P]Orthophosphate was purchased from ICN, and [γ-^{32}P]ATP was made from it by the method of Walseth and Johnson (1979) or was purchased from Amersham. Nitrocellulose filters (BA85) were from Schleicher and Schuell. 70S ribosomes and 30S ribosomal subunits were prepared by the method described by Staehelin and Maglott (1971) or by the method of Makhno et al. (1987); these gave similar results in these experiments. T7 RNA polymerase was either purchased from Boehringer or purified from cells that contain a cloned copy of the polymerase gene (Davanloo et al., 1984) according to procedures provided by Dr. F. W. Studier.

Construction of the Plasmid 19U-mRNA1 and in Vitro Transcription. The sequence of the synthetic mRNA used in these experiments is written under Results. To make the plasmid DNA that will act as the template for in vitro transcription of this RNA, complementary DNA strands beginning with a *Hin*dIII site and ending in a *Bam*HI site were annealed and then ligated into plasmid pTZ19U (USB) that had been digested with *Hin*dIII and *Bam*HI. The ligation mix was used to transform *E. coli* HB101 cells; double-stranded plasmid DNA was purified from individual colonies and was sequenced. Plasmid DNA from a colony containing a plasmid with the correct inset was purified on ethidium bromide/cesium chloride density gradients, was digested with *Bam*HI, and then was used for in vitro transcription. The transcription reactions contained 25 μg of digested 19U-mRNA1 in a total volume of 100 μL containing 40 mM Tris-HCl (pH 8.0), 8 mM $MgCl_2$, 24 mM NaCl, 2 mM spermidine, 10 mM dithiothreitol, 1.25 mM each A, C, G, and UTP + s^4UTP, and 30 μg of T7 RNA polymerase. The ratio s^4UTP:s^4UTP + UTP equal to 0.5 gives 1.5 s^4U/mRNA molecule. Approximately 100 μCi of [α-^{32}P]UTP or [α-^{32}P]CTP was included to give around 10000 cpm/pmol of mRNA (Cerenkov counting). Transcription was for 1 h at 37 °C, and then the mRNA was purified on an affinity column (Quiagen) according to instructions provided by the manufacturer. The typical yield was 10 μg of RNA containing from 1.5 to 4 s^4U/mRNA molecule, determined by absorbance measurements at 257 and 330 nm (ϵ_{330} of s^4U = 21000 mol^{-1} cm^{-1}).

Synthesis of 4-Thiouridine Triphosphate. 4-Thiouridine triphosphate was chemically synthesized according to Scheit (1968) with modifications (gift of J. L. Fourrey to A.F.) or alternatively enzymatically synthesized by the method described in Tanner et al. (1988).

Ribosome–mRNA Complex Formation and Cross-Linking. Complex formation between activated 70S ribosomes or 30S subunits and mRNA was done in 10 mM Tris-HCl, pH 7.4, 50 mM KCl, and 10 mM $MgCl_2$ (mRNA binding buffer; Katunin et al., 1980) or in the same buffer but at 300 mM KCl (activation buffer). Ten picomoles of 70S or 30S in a total volume of 100 μL was incubated for 15 min at 37 °C with a 5-fold molar excess of mRNA and, in some experiments, a 2–5-fold molar excess of $tRNA^{Phe}$, $tRNA^{Tyr}$, or Phe-$tRNA^{Phe}$. The samples were cooled to ice temperature. In one set of experiments, half of the mixtures were saved as nonirradiated controls. In another set of experiments, ribosomes that had not been complexed with mRNA and ribosome samples that had been complexes with ^{32}P-labeled mRNA but not containing s^4U were irradiated as control samples. The samples were irradiated at 366 nm for 10 min in a thermostated high-intensity mercury light irradiator (Isaacs et al., 1977); the intensity of light in this device (after filtration with the $CoNO_3$ filter) is about 100 mW/cm^2. Alternatively, photoirradiation was processed at 366 nm with a Bausch & Lomb monochromator as described in Favre et al. (1986). The cross-linked samples and control samples were digested with proteinase K, phenol extracted, ethanol precipitated, and redissolved at a final concentration of 150 ng/μL (total rRNA) or 50 ng/μL (16S rRNA).

Reverse Transcription Analysis. Primer extension analysis of the mRNA–rRNA cross-links using reverse transcriptase was done as previously described (Wollenzien, 1988). In the present experiments, 200 ng of 16S rRNA (or an equivalent amount of total RNA) from the cross-linking experiment, 1 pmol of ^{32}P-labeled oligonucleotide primer, and 2.2. units of AMV reverse transcriptase were used in each 20-μL reaction. For the 16S rRNA, 10 primers described previously were used (Ericson et al., 1989). Five primers were used to examine the first 40% of the 23S rRNA molecule (the 13S fragment); their sites of complementarity are 23S-One (1197–1219), 23S-Two (874–896), 23S-Three (720–732), 23S-Four (300–322), and 23S-Six (940–959).

Results

Characteristics of the Synthetic mRNA. The mRNA for this experiment, was made by in vitro transcription of a plasmid that contained a T7 promoter linked to a DNA sequence that encodes a synthetic mRNA that is a total length of 51 nucleotides. The sequence of this mRNA is 5′-GGGAAAGCUCCCGCGCCUCCUCU,GUC,UUC,UAC,-UAG,CCUCCUUGAUCGGAUC-3′.

The middle three triplets in the mRNA, indicated by commas, are codons for valine, phenylalanine, and tyrosine, and the fourth triplet is a stop codon. The position of the U residues inside the mRNA is distributed throughout the molecule except in the first eight nucleotides, which are part of the optimum T7 RNA polymerase promoter and *Hin*dIII site. Because of a potential stem–loop in the beginning of this sequence, the first uridine residue may not be available for

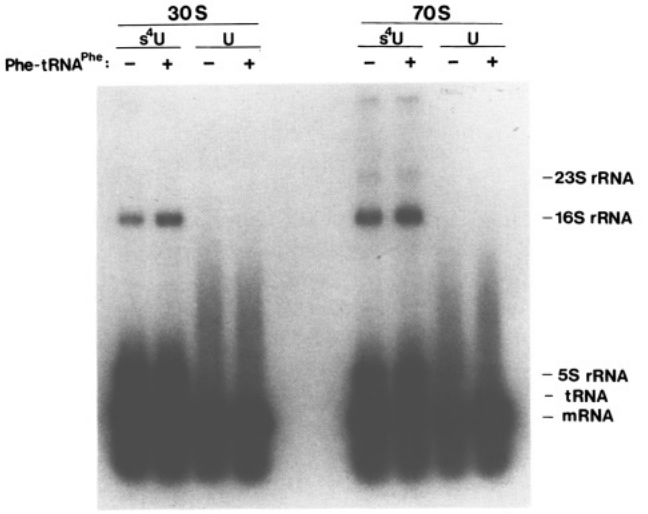

FIGURE 1: Gel electrophoresis of rRNA after cross-linking to mRNA. 30S subunits or 70S ribosomes were incubated with ^{32}P-labeled mRNA to form complexes and then were irradiated. These samples were digested with proteinase K and deproteinated, and then RNA was denatured in formamide and electrophoresed on agarose gels. An autoradiogram of the dried gel is shown here. The lanes marked s^4U or U were experiments performed with mRNA containing 4-thiouridine or uridine. In addition, the complexes were formed without or with Phe-tRNA^{Phe} as indicated. The locations of the RNA species were determined from a photograph of the ethidium-stained gel.

reaction with ribosome components. Since the RNA used in this experiment is intended to imitate an mRNA during the elongation phase of protein synthesis, the Shine–Dalgarno homology was purposely not included in its sequence.

Plasmid DNA containing this insert was digested with *Bam*HI, and T7 RNA polymerase was used to transcribe this RNA in the presence of the four normal nonradioactive nucleotide triphosphates, 4-thiouridine triphosphate, and uridine or cytidine [α-^{32}P]triphosphate. The concentration of 4-thiouridine triphosphate was adjusted to achieve an average of 1.5–4 mol of s^4U/mol of mRNA, as determined by the UV spectra. In most experiments, mRNA containing 3–3.5 mol of s^4U/mol of mRNA was used.

Conditions for optimal and specific interaction between poly(U) mRNA and 70S ribosomes or 30S subunits have been described by Katunin et al. (1980) and Kirillov et al. (1980). Under the same ionic conditions (see Material and Methods), with the 70S ribosomes and 30S subunits we are using, a 5-fold molar excess of mRNA over 30S subunits and 70S ribosomes achieved a binding of 50% of the subunits or ribosomes as determined by filter binding. The inclusion of deacylated tRNA^{Phe} (or deacylated tRNA^{Tyr}) did not alter the amount of mRNA bound to the 30S subunit or 70S ribosome. With a 5-fold molar excess of mRNA and a 5-fold molar excess of Phe-tRNA^{Phe}, the stoichiometry of Phe-tRNA^{Phe} binding was 40%. The amount of bound Phe-tRNA^{Phe} reactive with puromycin is 30%. This indicates that most of the bound mRNA is capable of directing tRNA binding and that some of the bound tRNA is in the functionally defined P site.

Analysis of Cross-Linking by Electrophoresis on Agarose Gels. The level of cross-links of the ^{32}P-labeled mRNA to 16S RNA and 23S RNA was determined by the analysis on 1% agarose gels of proteinase K, phenol-extracted samples (Figure 1). RNA samples were denatured before electrophoresis by heating in formamide. A fraction of the labeled mRNA migrates with 16S and 23S RNA only when thiolated mRNA is used (see Figure 1). In addition, no comigration of the labeled mRNA with 16S or 23S RNA was obtained in the control lanes when the samples were not irradiated (results not shown). There is no addition of the mRNA to 5S rRNA as judged from the similarity in the lanes in which 30S or 70S ribosomes were used, nor is there addition of the mRNA to tRNA as judged from the similarity in the lanes in which tRNA was or was not used.

In one set of experiments, in which the 23S rRNA was partially degraded into 18S and 13S fragments, cross-linking occurred only on the 13S fragment (results not shown). Since the 13S fragment constitutes the first 40% of the 23S rRNA (Noller et al., 1981), only that part was subsequently examined by primer extension.

The fraction that is covalently cross-linked is proportional to s^4U substitution level, and this does not depend on the presence of nonacylated tRNA^{Phe} or tRNA^{Tyr} during the incubation step. Therefore, in our conditions the only effect of the presence of tRNA could be a phasing of the mRNA. As is shown later, the presence of tRNA^{Phe} has a significant effect on the site of interactions of the mRNA with ribosomal RNA at one site in the 16S rRNA.

Localization of rRNA–mRNA Binding Sites. In these experiments we compare the pattern of elongation of the reverse transcriptase obtained on 16S RNA and 23S RNA using a suitable set of oligodeoxynucleotide primers. Control samples were usually 30S or 70S ribosomes complexed and irradiated with mRNA not containing s^4U and 30S or 70S ribosomes irradiated without any added mRNA. These control for a random pattern of stops in the RNA template, any possible UV-induced changes, and any changes that might be induced by the presence of mRNA. Alternatively, control samples that had been assembled with mRNA but were not irradiated were used. The samples were hybridized to the primer, elongated with AMV reverse transcriptase, and analyzed on sequencing gels. The RNA was hydrolyzed to avoid the presence of radioactive bands due to the presence of excess un-cross-linked mRNA.

When a one-band stop is observed, the actual cross-linking point must be the next adjacent nucleotide (in the 5′-position) in the rRNA. In Figure 2 and in the subsequent discussion, the actual cross-linking sites will be used. When a two-band stop is observed, one cross-linking point must be the 5′ nucleotide of the two nucleotides. In addition, a cross-link could also be present on the next nucleotide as well (on the 5′ side of the two bands). Because of this ambiguity, these possible sites will be listed with question marks in the tabulation of the cross-linking sites in Table I.

Relevant portions of autoradiograms of reverse transcription experiments are shown in Figure 2. The yield of the cross-link at a specific site is dependent on a large number of parameters, among which only two can be controlled: the level of mRNA–ribosome interaction and the average number of s^4U per mRNA molecule. Natural stops of the reverse transcriptase can be detected for each nucleotide, and this background sets the lower limit for the detection of sites (see Figure 2). The use of multiple controls and their arrangement in the gel electrophoresis allow faint signals present in the experimental samples to be detectable. A signal was obtained for each detectable site at 1.5 s^4U/mRNA molecule (lowest level tested), and the same sites were obtained with a better signal to background level at 4.5 s^4U/mRNA molecule (highest level tested).

Table I lists the sites that were repeatedly found in at least three of four different cross-linking experiments. These sites are distributed in the three major domains of the 16S RNA, indicating that the mRNA spans a large portion of the 30S particle. On the contrary, only two sites were obtained on the

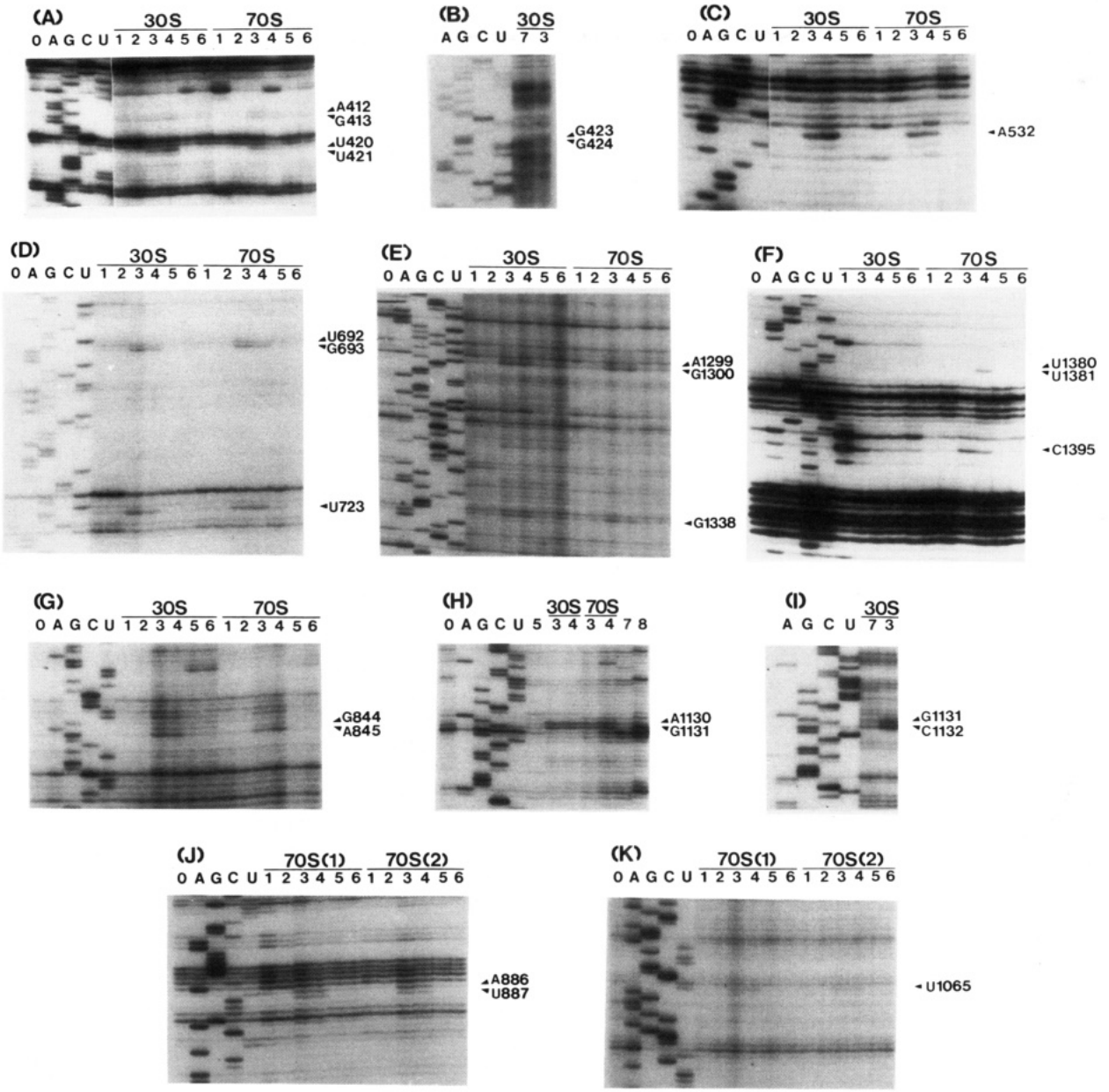

FIGURE 2: Reverse transcription analysis of rRNA cross-linking sites. ^{32}P-Labeled oligonucleotide primers were mixed with total rRNA from 70S ribosomes for sequencing lanes or with RNA isolated from the cross-linking experiments; reverse transcriptase was used to extend the primers. The regions shown are (A) 16S, 433–400; (B) 16S, 441–398; (C) 16S, 540–525; (D) 16S, 731–682; (E) 16S, 1354–1280; (F) 16S, 1412–1368; (G) 16S, 857–829; (H) 16S, 1148–1113; (I) 16S, 1151–1112; (J) 23S, 900–869; and (K) 23S, 1078–1049. Sequencing reactions and a control reaction are marked A, G, C, U, and 0. Other lanes are 30S or 70S ribosomes (1 and 2) irradiated without mRNA (without or with Phe-tRNA^{Phe}), (3 and 4) irradiated with s^{4}U–mRNA (without or with Phe-tRNA^{Phe}), or (5 and 6) irradiated with mRNA containing U only (without or with Phe-tRNA^{Phe}). Control lanes marked 7 and 8 were done with RNA from 30S subunits and 70S ribosomes incubated with s^{4}U–mRNA but not irradiated. Positions indicated on the right of each panel are the nucleotides on which cross-linking occurs. In panels J and K, samples labeled 70S(1) and 70S(2) were from experiments in which activation buffer and mRNA binding buffer were used, respectively.

23S RNA, indicating very limited interaction. In three cases (A532, U1381, and C1395) there are differences in the efficiency of cross-linking that are dependent upon experimental conditions, as can be seen from the relative intensity of the stops (Figure 2 and Table I). In two regions there are differences in the exact locations of the cross-links: one set of samples contained cross-links at G424, G1131, and C1130, whereas similar samples with other ribosomes contained cross-links at G413, U421, and G1131. We attribute these to differences in the ribosomes. All other cross-links were identical.

The 3′-terminal section between 1530 and 1542 cannot be scanned by the primer extension method. However, a primer that initiated extension at position 1510 failed to detect any cross-links until position C1395 was reached.

In addition to the listed cross-linking sites, two other sites in the 16S rRNA are worth mentioning. In reverse transcription experiments, samples that were irradiated in the presence of mRNA that did not contain s^{4}U consistently showed a stop at position U834. At a second site, G347, samples that were irradiated in the presence of mRNA either containing or not containing s^{4}U both showed stops. Since we do not expect the mRNA without s^{4}U to give rise to cross-links, both of these effects may be indirect. Therefore, we are not attributing cross-links to either of these sites.

Discussion

Strategy of Using This mRNA and Its Properties. The 51-nucleotide synthetic mRNA was chosen because this size is only slightly longer than the size protected by the ribosome

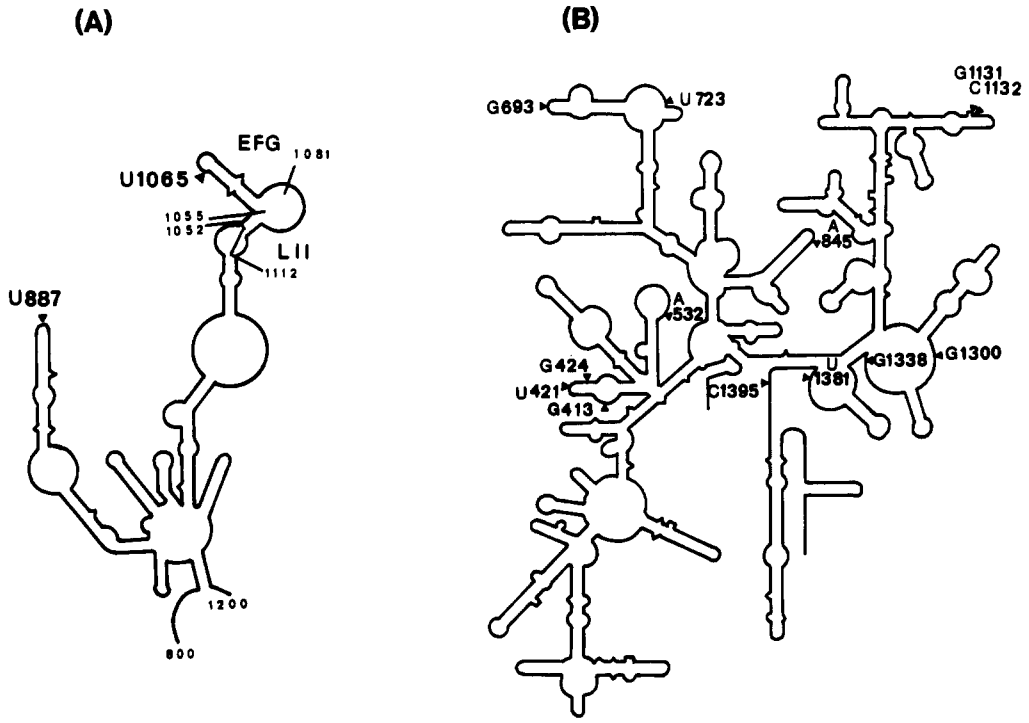

FIGURE 3: Locations of the mRNA cross-linking sites in the rRNA secondary structures. Part of 23S RNA domain II (A) and the entire 16S RNA secondary structure (B) are shown. In panel A, the sections of the RNA associated with elongation factor G (EFG) and protein L11 are marked.

Table I: Sites of mRNA Cross-Linking in 16S and 23S rRNA

cross-linking positions[a]	30S	70S	comments
16S rRNA			
-G413, A412?	+	+	
-U421, U420?	+	+	
-G424, G423?	+	ND[b]	
-A532	++	+	increased in activation buffer
-G693, U692?	+	+	
-U723	+	+	
-A845, G844?	+	+	
-G1131, C1132, A1130?	+	+	
-G1300, A1299?	+	+	
-G1338	+	+	
-U1381, U1380?	+	++	increase in 70S with tRNA
-C1395	-	+	only in 70S
23S rRNA			
-A887		+	
-U1065		+	

[a] The nucleotides adjacent to the reverse transcription stops are listed as the cross-linking sites. Additional sites where there may be cross-linking as judged from double reverse transcription stops are listed with question marks. [b] ND, not determined.

in nuclease digestion experiments (Kang & Cantor, 1985). This should prevent the binding of two mRNAs to the same ribosome while still allowing us to determine the entire track of the mRNA through the ribosome. In addition, this length is not so long as to promote the aggregation of subunits or ribosomes along the polynucleotide or to allow extensive structure within the mRNA that would interfere with its interaction with the ribosome. The strategy of randomly incorporating s^4U throughout the mRNA is economical experimentally in that only one experimental sample needs to be constructed and probed during the primer extension experiment, but it does not give information as to the order or polarity of the rRNA contact sites along the mRNA. Nevertheless, the experiment has revealed many more rRNA sites than were previously known, and it indicates that primer extension will be suitable in analyzing experiments in which the cross-linking is accomplished by site-directed techniques.

The amount of mRNA needed to saturate the 30S and 70S ribosomes is significantly larger than the amounts of poly(U) described by Katunin et al. (1980) probably because of differences in the behavior of the synthetic mRNA we are using compared to poly(U). There may as well be differences in the behavior of the ribosomes and ribosome subunits since the level of protein S1 has been shown to be necessary to reach a 1:1 stoichiometry between 30S or 70S and poly(U) (Katunin et al., 1980; Makhno et al., 1987).

Several aspects of the experiment indicate that the interaction between the ribosome and the 51-nucleotide RNA molecule used here to represent the mRNA is specific, functionally relevant, and not due to regions of fortuitous base complementarity. First, binding saturation of the RNA on the ribosome occurs at less than 1, indicating a specific and limited interaction. Second, the 51-nucleotide mRNA does direct tRNA binding. Third, the sites of cross-linking in the rRNA do have regions of potential complementarity with the mRNA; however, there is no correlation between the intensity of cross-linking and the extent of potential complementarity. In addition, there are many other sites in the rRNA that also have complementarity to the mRNA and are not at all reactive. Last, the exact pattern of cross-linking does not indicate base pairing between the rRNA and the mRNA. If the interactions between the rRNA and mRNA were occurring through base pairing, there should be several cross-linking sites around each sensitive area in the rRNA, and these are not seen.

Distribution of mRNA Sites on the 23S RNA Secondary Structure. Figure 3A shows the part of 23S domain II (Noller et al., 1981) that is relevant to the mRNA cross-links we have obtained at positions U887 and U1065. Position U887 cannot be linked at present to any function or location, but position U1065 is interestingly in the section of 23S RNA that has been related to the GTPase activity. Factor EFG was cross-linked with diepoxybutane to the section 1055–1087, which is inside

Table II: Characteristics of mRNA–16S rRNA Cross-Linking Sites Compared to Poly(U)–tRNA Protection Sites

	poly(U)–tRNA protection[c,d]		characteristic of sites	
mRNA cross-link[a,b]	A site	P site	common	distinct[e]
-G413, U421, G424				mRNA
-A532	**G529, G530, U531**	***A532***	A, P	
-G693		***G693***	P	
-U723				mRNA
		A794, C795		tRNA
-A845				mRNA
		G926		tRNA
		G966		tRNA
-G1131, C1132				mRNA
-G1300				mRNA
-G1338		***G1338, A1339***	P	
-U1381		***U1381***	P	
-C1395		***C1399***, C1400, G1401	P	
	G1405, A1408			tRNA
	A1492, A1493, G1494			tRNA

[a] Cross-links are from this study. [b] Nucleotides that are detected by both cross-linking and chemical probing (class i sites) are indicated in boldface. Common or neighboring nucleotides correlated with A site or P site tRNA binding are underlined or italicized, respectively. [c] Sites of chemical protection are from Moazed and Noller (1986, 1990). [d] Three other sites detected in the chemical protection experiments (A790, G791, and A909) have not been assigned to the A or P sites. [e] Sites detected only by chemical probing experiments (class ii sites) are noted as tRNA. Sites detected only in the cross-linking experiments (class iii sites) are noted as mRNA.

the L11-binding site (C1052–G1112) (Sköld, 1983). Nucleotide A1067 is the site of methylation that confers resistance to the GTPase inhibitor thiostrepton (Thompson et al., 1982). Thiostrepton prevents a number of interactions that are strongly correlated with the A tRNA-binding site: it inhibits IF2, EFTu-dependent tRNA binding, and EFG–GTP complex interactions with the 70S ribosome. The thiostrepton site of interaction was partly analyzed by Egebjerg et al., (1989, 1990) and was shown to correspond to loop 1064–1074, which is also part of L11- and EFG-binding sites (Moazed et al., 1988; Ryan & Draper, 1989; Egebjerg et al., 1990).

Because the EFG-binding site in the 70S ribosome has been determined (Girshovich et al., 1981, 1986; Langer & Lake, 1986), we can make a preliminary conclusion concerning the location of this 23S rRNA–mRNA interaction: it is on the 50S surface, facing the 30S subunit on the base of the rodlike appendage (L7/L12). This site is probably close to the site of another interaction at 16S rRNA position A532. A contact of the mRNA at U1065 and the interaction of this segment of the 23S rRNA with EFG suggest the involvement of this loop with the A site codon–anticodon interaction. It is not possible to know yet if these are direct, close-range interactions or if they are indirect interactions. Immunoelectron microscopy studies so far indicate that the EFG center and EFTu center are not overlapping with the decoding site on the 30S subunit [see Kastner et al. (1990)]. Determining the distance along the mRNA from the A and P sites to the cross-link at U1065 will be important in determining how the GTPase reactions are mechanistically related to translocation and to the proofreading control mechanism (Bilgin et al., 1988).

Distribution of mRNA Sites on the 16S RNA Secondary Structure. Twelve sites in the 16S rRNA have been cross-linked by mRNA. Figure 3B shows the distribution of these sites in the 16S rRNA secondary structure. Evidence exists for the location of several of these sites in the 30S particle. C1395 is very close to C1400, where the anticodon of $tRNA^{Val}$ was cross-linked while it was functionally in the P site (Ofengand et al., 1986). This site has been located by immunoelectron microscopy to be in the base of the cleft (Keren-Zur et al., 1979; Gornicki et al., 1984). U1381, G1338, and C1300 are all in the region associated with protein S7; S7 has been located in the head facing the cleft region by immunoelectron microscopy (Stoeffler & Stoeffler-Meilicke, 1986) and neutron-scattering experiments (Capel et al., 1987). A845 is located in the region cross-linked to IF3 (Ehresmann et al., 1986), and IF3 has been efficiently cross-linked to S7 (Van Duin et al., 1975). This suggests a location of A845 in the cleft region. G693 and U723 are close to sites that were recently shown to be footprinted by IF3 (Muralikrishna & Wickstrom, 1989), and these two sites are associated with S6 and S18 [see Stern et al. (1988)]. The location of this region of the 16S rRNA has been determined by electron microscopy to be in the platform, although there are significant differences in the placement of these sites in existing three-dimensional models (Expert-Bezançon & Wollenzien, 1985; Brimacombe et al., 1988; Stern et al., 1988; Oakes et al., 1990). The fact that these two sites are directly in contact with mRNA places additional constraints on their location. The location of A532 is somewhat different from that of the preceding sites. A532 is in the loop containing m^7G527. m^7G527 and the entire loop region (nucleotides 518–533) have been located by immunoelectron microscopy on the cytoplasmic side of the 30S particle between the head and body (Trempe et al., 1982; Oakes et al., 1990). The locations of the remaining four sites (G413, U421, and G424 in the body of the subunit and G1131/C1132 in the head of the subunit) are less certain. Three-dimensional arrangements for the rRNA that place these sites in reasonable locations for mRNA contact will be described elsewhere.

Comparison of mRNA Cross-Linking Sites and Sites Protected from Chemical Modification by tRNA and Poly(U). Moazed and Noller (1986, 1989) used chemical probing to determine which nucleotides in the 16S and 23S rRNA are protected when tRNA was bound to 70S ribosomes with or without poly(U). These results can now be compared to sites at which there is direct mRNA interaction as evidenced by cross-linking. There are three classes of sites: (i) those seen by cross-linking and chemical probing, (ii) those seen only by chemical probing, and (iii) those seen only by cross-linking.

Five sites (or clusters of adjacent sites) were detected in both experiments (Table II). Of these, only the sites in the region 529–532 were poly(U) dependent or were enhanced in the presence of poly(U). The fact that the remaining sites in this class (G693, G1338, U1331, and C1395) are cross-linked to mRNA indicates that the effects of tRNA binding and mRNA binding may be overlapping. This raises the possibility that there is a close spatial proximity of the rRNA at these sites to the anticodon loop of the tRNA.

Chemical probing detects changes at numerous sites where there was no evidence for mRNA interactions from cross-linking (Table II, class ii). This discrepancy may be due to the failure of this cross-linking experiment to detect part of the mRNA track. Experiments with mRNAs containing different sequences may identify additional rRNA sites, and these experiments are underway. Alternatively, the chemical probing experiment may detect changes in reactivity in the rRNA due to conformational changes in the subunit that are a significant distance away from the actual mRNA track.

Lastly, the fact that seven mRNA–rRNA cross-linking sites did not result in detectable chemical protection with poly(U) deserves comment. This could result from a low affinity of poly(U) to the corresponding sites that allows for structural fluctuations and/or from a peculiar stereochemistry of interaction that does not result in protections against chemical reagents.

Correlation of mRNA Cross-Linking Sites to the A and P tRNA Sites. Moazed and Noller (1990) correlated chemical protection sites to occupancy of tRNA in the A and P binding sites. The mRNA cross-linking site at A532 is the same as the A532 protection when tRNA is in the P site; this site is also adjacent to G529, G530, and U531, which are protected when tRNA is in the A site. These sites lie in the m^7G527 loop where a streptomycin revertant was obtained (Melançon et al., 1988). Taken together, these data strengthen the idea that the m^7G527 loop is part of the proofreading site as has been suggested by Moazed and Noller (1986). Relevant to this observation is the suggestion of Trifonov (1987) that the three-base periodicity NNC at ^{517}GCC AGC AGC CGC^{528} could be part of an mRNA frame-monitoring mechanism. The A532 mRNA cross-linking is outside of these sequences, but the mRNA we used here did not contain the necessary periodicity. It will therefore be interesting to determine if an mRNA bearing the "framing" periodicity (G, non-G, N)$_n$ has a different pattern of cross-linking in the 517–528 section. The other A site protected bases are not next to mRNA cross-linking sites, so they may arise from direct tRNA protection or from indirect effects (see Table II).

Four groups of sites are cross-linked to mRNA and are protected when tRNA is in the P site. One site (C1395) is very close to the protections seen at C1399, C1400, and G1401. As described above, this region must be close to the actual codon–anticodon interaction at the P site (Ofengand et al., 1986). The section ^{1395}CAC ACC UCC^{1403} was also suggested as a potential mRNA framing site (Trifonov, 1987). Two mRNA cross-linking sites in domain III (G1338 and U1381) are correlated with protected sites at G1338, A1339, and U1381. The last site in this group is G693; this is both cross-linked to mRNA and protected when tRNA is in the P site. Other sites associated with A site protection (particularly in the 1405–1308 and 1500 regions) are not yet correlated to mRNA cross-linking.

Acknowledgments

Raman Bhangu and Kalpa Vithalani are thanked for providing the 19U-mRNA1 plasmid. Raman Bhangu is thanked for making the Phe-tRNA^{Phe} binding measurements with the synthetic mRNA. Jean-Louis Fourrey is thanked for providing chemically synthesized 4-thiouridine triphosphate.

References

Bilgin, N., Kirsebom, L. A., Ehrenberg, M., & Kurland, C. G. (1988) *Biochimie 70*, 611–618.

Brimacombe, R., Atmadja, J., Stiege, W., & Schuler, D. (1988) *J. Mol. Biol. 199*, 115–136.

Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgaard, M., Langer, J. A., Ramakrishnan, V., Shindler, D. G., Schneider, D. K., Schoenborn, B. P., Yabuki, S., & Moore, P. B. (1987) *Science 238*, 1403–1406.

Cooperman, B. S. (1980) in *Ribosomes. Structure, Function, and Genetics* (Chamblis, G., Craven, G. R., Davies, J., Davies, K., Kahan, L., & Nomura, M., Eds.) pp 531–554, University Park Press, Baltimore, MD.

Dahlberg, A. E. (1989) *Cell 57*, 525–529.

Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.

Denman, R., Colgan, J., Nurse, K., & Ofengand, J. (1988) *Nucleic Acids Res. 16*, 165–178.

Egebjerg, J., Douthwaite, S. R., & Garrett, R. A. (1989) *EMBO J. 8*, 607–611.

Egebjerg, J., Douthwaite, S. R., Liljas, A., & Garrett, R. A. (1990) *J. Mol. Biol. 213*, 275–288.

Ehresmann, C., Moine, H., Mougel, M., Dondon, J., Grunberg-Manago, M., Ebel, J. P., & Ehresmann, B. (1986) *Nucleic Acids Res. 14*, 4803–4821.

Ericson, G., & Wollenzien, P. (1988) *Anal. Biochem. 174*, 215–223.

Ericson, G., Chevli, K., & Wollenzien, P. (1989) *Biochemistry 28*, 6446–6454.

Expert-Bezançon, A., & Wollenzien, P. L. (1985) *J. Mol. Biol. 184*, 53–66.

Favre, A. (1990) in *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids* (Morrison, H., Ed.) Vol. 1, J. Wiley & Sons, Inc., New York.

Favre, A., Bezerra, R., Hajnsdorf, E., Lemaigre Dubreuil, Y., & Expert-Bezançon, A. (1986) *Eur. J. Biochem. 160*, 441–449.

Girshovich, A. S., Bochkareva, E. S., & Ovchinnikov, Y. A. (1981) *J. Mol. Biol. 151*, 229–243.

Girshovich, A. S., Bochkareva, E. S., & Vasiliev, V. D. (1986) *FEBS Lett. 197*, 192–198.

Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S., & Stormo, G. (1981) *Annu Rev. Microbiol. 35*, 365–403.

Gornicki, P., Nurse, K., Hellman, W., Boublik, M., & Ofengand, J. (1987) *J. Biol. Chem. 259*, 10493–10498.

Graifer, D. M., Zenkova, M. A., Malygin, A. A., Mamaev, S. V., Mundas, D. A., & Karpova, G. G. (1990) *J. Mol. Biol. 214*, 121–128.

Hajnsdorf, E., Favre, A., & Expert-Bezançon, A. (1986) *Nucleic Acids Res. 14*, 4009–4023.

Isaacs, S. T., Shen, C. K. J., Hearst, J. E., & Rapoport, H. (1977) *Biochemistry 16*, 1058–1064.

Kang, C., & Cantor, C. R. (1985) *J. Mol. Biol. 181*, 241–251.

Kastner, B., Trotman, C. N. A., & Tate, W. P. (1990) *J. Mol. Biol. 212*, 241–245.

Katunin, V. I., Semenkov, Yu. P., Mahko, V. I., & Kirrillov, S. V. (1980) *Nucleic Acids Res. 8*, 403–421.

Keren-Zur, M., Boublik, M., & Ofengand, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1054–1058.

Kirillov, S. V., Makhno, V. I., & Semenkov, Y. P. (1980) *Nucleic Acids Res. 8*, 183–196.

Langer, J. A., & Lake, J. A. (1986) *J. Mol. Biol. 187*, 617–621.

Lemaigre Dubreuil, Y., Lamine, K., Hajnsdorf, E., Favre, A., & Le Bret, M. (1986) *Biochemistry 25*, 5726–5735.

Liljas, A. (1982) *Prog. Biophys. Mol. Biol. 187*, 617–621.

Makhno, V. I., Pechin, N. N., Semenkov, Yu. P., & Kirillov, S. V. (1987) *Mol. Biol. 22*, 528–537.

Melançon, P., Lemieux, C., & Brakier-Gingras, L. (1988) *Nucleic Acids Res. 16*, 9631–9639.
Moazed, D., & Noller, H. F. (1986) *Cell 47*, 985–994.
Moazed, D., & Noller, H. F. (1990) *J. Mol. Biol. 211*, 135–145.
Moazed, D., Robertson, J. M., & Noller, H. F. (1988) *Nature 334*, 362–364.
Muralikrishna, P., & Wickstrom, E. (1989) *Biochemistry 28*, 7505–7510.
Noller, H. F., Kop, J. A., Wheaton, V., Brosius, J., Gutell, R. R., Kopylov, A. M., Dohme, F., Stahl, D. A., Gupta, R., & Woese, C. R. (1981) *Nucleic Acids Res. 9*, 6167–6189.
Oakes, M. I., Kahan, L., & Lake, J. A. (1990) *J. Mol. Biol. 211*, 907–918.
Ofengand, J., Ciesiolka, J., Denman, R., & Nurse, K. (1986) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., & Kramer, G., Eds.) pp 473–494, Springer-Verlag, New York.
Ryan, P. C., & Draper, D. E. (1989) *Biochemistry 28*, 9949–9956.
Scheit, K. H. (1968) *Chem. Ber. 101*, 1147.
Sköld, S. E. (1983) *Nucleic Acids Res. 11*, 4923–4932.
Spirin, A. S., & Vasiliev, V. D. (1989) *Biol. Chem. 66*, 215–223.
Stade, K., Rinke-Appel, J., & Brimacombe, R. (1989) *Nucleic Acids Res. 17*, 9889–9908.
Staehelin, T., & Maglott, C. R. (1971) *Methods Enzymol. 20*, 449–456.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447–481.
Stiege, W., Stade, D., Schuler, D., & Brimacombe, R. (1988) *Nucleic Acids Res. 16*, 2369–2388.
Stoeffler, G., & Stoeffler-Meilicke, M. (1986) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., & Kramer, G., Eds.) pp 28–46, Springer-Verlag, New York.
Tanner, N. K., Hanna, M. M., & Abelson, J. (1988) *Biochemistry 27*, 8852–8861.
Thompson, J., Schmidt, F., & Cundliffe, E. (1982) *J. Biol. Chem. 257*, 7915–7917.
Traub, P., Mizushima, S., Lowry, C. V., & Nomura, M. (1971) *Methods Enzymol. 20*, 391–407.
Trempe, M. R., Ohgi, K., & Glitz, D. G. (1982) *J. Biol. Chem. 257*, 9822–9829.
Trifonov, E. N. (1987) *J. Mol. Biol. 194*, 643–652.
Van Duin, K., Kurland, C. G., Dondon, J., & Grunberg-Manago, M. (1975) *FEBS Lett. 59*, 287–290.
Vladimirov, S. N., Babkina, G. T., Venijaminova, A. G., Gimautdinova, O. I., Zenkova, M. A., & Karpova, G. G. (1990) *Biochim. Biophys. Acta 1048*, 245–256.
Wagner, R., Gassen, H. G., Ehresmann, Ch., Stiegler, P., & Ebel, J. P. (1976) *FEBS Lett. 67*, 312–315.
Walseth, T. F., & Johnson, R. A. (1979) *Biochim. Biophys. Acta 562*, 11–31.
Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., & Gesteland, R. F. (1988) *EMBO J. 7*, 1503–1507.
Wollenzien, P. L. (1988) *Methods Enzymol. 164*, 319–329.

Properties of a U1/mRNA 5′ Splice Site Duplex Containing Pseudouridine As Measured by Thermodynamic and NMR Methods†

Kathleen B. Hall*
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 South Euclid Avenue, St. Louis, Missouri 63110
Larry W. McLaughlin*
Department of Chemistry, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02167
Received September 5, 1990; Revised Manuscript Received November 27, 1990

ABSTRACT: Three RNA undecamers, $^{5'}$AUACΨΨACCUG (Ψ = pseudouridine), $^{5'}$AUACUUACCUG, and their complementary 11-mer $^{5'}$CAGGUAAGUAU, have been chemically synthesized by phosphite triester chemistry on a controlled-pore glass (CPG) support. The two duplexes formed with these molecules, $^{5'}$AUACΨΨACCUG/$^{5'}$CAGGUAAGUAU and $^{5'}$AUACUUACCUG/$^{5'}$CAGGUAAGUAU, represent the 5′ end of human U1 snRNA paired to the mRNA consensus 5′ splice site. In one undecamer, pseudouridines are incorporated at those positions corresponding to the native in vivo U1 snRNA, while the other (control) undecamer contains only uridine. Surprisingly, the NMR data show that the extra imino proton of the pseudouridines, which is found in the major groove and is presumably not hydrogen bonded, is clearly visible in the imino proton NMR spectrum at pH 6. This result suggests that the structure of the RNA restricts access of solvent to the major groove, slowing the exchange of the pseudouridine NH1 imino proton. A comparison of the thermodynamic properties of the two duplexes show that the free energy of duplex formation is unchanged by the substitution of pseudouridine for uridine.

Most RNA molecules contain modified nucleotide residues at specific locations within a given sequence. The functional and/or structural importance of these modifications is not well understood.

Pseudouridine (Ψ) is a regioisomer of uridine in which the heterocyclic base uracil is bound to the ribose sugar through the carbon atom at the C5 position rather than through the normal N1 nitrogen (Figure 1). With the formation of a C-nucleotide in place of the more common N-nucleotide, two

†Supported by the Lucille P. Markey Charitable Trust (88-30 to K.B.H.) and by grants to L.W.M. from the National Science Foundation (DMB-8904306) and the National Institutes of Health (GM-37065).
*Correspondence can be addressed to either author.

0006-2960/91/0430-1795$02.50/0