Four Newly Located Pseudouridylate Residues in *Escherichia coli* 23S Ribosomal RNA Are All at the Peptidyltransferase Center: Analysis by the Application of a New Sequencing Technique

Andrey Bakin and James Ofengand*

Roche Institute of Molecular Biology, Roche Research Center, Nutley, New Jersey 07110

Received May 28, 1993®

ABSTRACT: A new technique has been developed for the facile location of pseudouridylate (Ψ) residues in any RNA molecule. The method uses two known modification procedures which in combination uniquely identify U residues which have been converted into Ψ. The first procedure involves reaction of all U-like and G-like residues with *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium)ethylcarbodiimide *p*-tosylate (CMC), followed by alkaline removal of all CMC groups except those linked to the N_3 of Ψ. This stops reverse transcription, resulting in a gel band which identifies the U residue. The second procedure is uridine-specific hydrazinolysis which cleaves the RNA chain at all U residues and produces a gel band upon reverse transcription. Ψ residues, being resistant to hydrazinolysis, are not cleaved and do not stop reverse transcription. This leads to the absence of a band at Ψ residues. The combined method can also distinguish Ψ from 5-methyluridine, 4-thiouridine, uridine-5-oxyacetic acid, and 2-thio-5-methylaminomethyluridine as shown by treating rRNA and tRNA species known to contain these modified bases at defined sites. By this procedure, four new sites for Ψ in *Escherichia coli* 23S RNA were discovered, and one was disproven. The four new sites are at positions 2457, 2504, 2580, and 2605. The erroneous site is at position 2555. These four new Ψ residues, which are all in or within 2–3 residues of the peptidyltransferase ring, are thus in a position to play a functional and/or structural role at the peptidyltransferase center. In addition, this work has shown the presence of a new unidentified modified C residue at position 2501, also in the peptidyltransferase ring.

Pseudouridylic acid (Ψ), the carbon–carbon glycoside isomer of uridylic acid, was detected as a new nucleotide by Davis and Allen (1957), and its structure proposed shortly thereafter (Yu & Allen, 1959; Scannell *et al.*, 1959). The structure shown in Figure 1 was definitively proven by Cohn (1959, 1960). Ψ is found in tRNA (Sprinzl *et al.*, 1991), rRNA (Maden, 1990), and snRNA (Reddy, 1988) but not so far in mRNA or viral RNAs. Enzymes which form Ψ in tRNA have been purified from several sources (Arena *et al.*, 1978; Green *et al.*, 1982; Samuelsson & Olsson, 1990) and an enzyme from *Escherichia coli* has been cloned (Kammen *et al.*, 1988). In all cases studied, Ψ is formed at the polynucleotide level by conversion of a specific uridine residue in a reaction which requires no energy source.

In spite of the fact that this unusual nucleotide has been known for over 30 years and is ubiquitously distributed, no specific function of any class of RNA has yet to be shown to be dependent on Ψ formation. For example, although Ψ is found in most cytoplasmic elongator tRNAs as part of the highly conserved GrTΨC sequence (Sprinzl *et al.*, 1991) as well as frequently elsewhere in the molecule, notably in the anticodon loop and stem, tRNA species with all the U and U-derived residues replaced by 5-fluorouracil were shown to be fully active in both aminoacylation and protein synthesis functions (Horowitz *et al.*, 1974; Ofengand *et al.*, 1974; Chinali *et al.*, 1978). More recently, transcripts of tDNA made by T7 RNA polymerase which lack all modified bases, including Ψ, have been shown to be functional both in aminoacylation [reviewed in Schulman (1991)], although in some cases with decreased specificity (Perret *et al.*, 1990), and in ribosomal protein synthesis (Samuelsson *et al.*, 1988). It should be noted, however, that a role in regulation of the synthesis of certain enzymes has been described for the Ψ residues at position 38, 39, and 40 in tRNA (Johnston *et al.*, 1980; Tsui *et al.*, 1991).

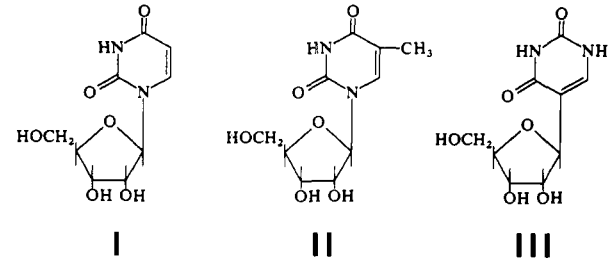

FIGURE 1: Pyrimidine nucleoside structures. I, uridine; II, ribothymidine; III, pseudouridine.

Considerably less information is available about Ψ residues in ribosomal RNAs. The number of Ψ residues in *E. coli* 16S RNA has been variously estimated to be between 1 and 2 (Dubin & Günalp, 1967; Nichols & Lane, 1967; Krzyzosiak *et al.*, 1987; Gehrke & Kuo, 1989), and in 23S RNA, values of 4–5 (Nichols & Lane, 1967), 8 (Gehrke & Kuo, 1989), or 9 (Dubin & Günalp, 1967) have been reported. The site(s) in 16S RNA are unknown, and in 23S RNA only four of the Ψ positions have been localized. Eukaryotic ribosomal RNAs contain many more Ψ residues, some of which have been localized as a result of the efforts primarily of Maden and his colleagues [reviewed in Maden (1990)]. No enzymes have been identified which form Ψ in rRNA, and the Ψ residues have no known function, although a role for Ψ at the peptidyltransferase center has been postulated (Lane *et al.*, 1992). The time of synthesis of Ψ is not known for *E. coli*, but in yeast and other eukaryotes it occurs during the early precursor stages of rRNA synthesis (Brand *et al.*, 1979; Maden, 1990).

* To whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts*, August 15, 1993.

0006-2960/93/0432-9754$04.00/0

A prerequisite to understanding the function of Ψ is knowledge of where the residues are located. This has been up to now a laborious process, especially in large RNAs, and has hindered progress in understanding the role of Ψ. In this work, we describe two complementary sequencing methods based on primed reverse transcription which, in combination, make the location of Ψ residues in any RNA a straightforward process. We illustrate the methodology by an analysis of *E. coli* 23S RNA. This analysis has confirmed the presence of three of the four sites previously described (Noller, 1984; Raué *et al.*, 1988) and has identified four new Ψ residues. Remarkably, all seven of the Ψ sites cluster in the vicinity of the peptidyltransferase center.

MATERIALS AND METHODS

Materials. $tRNA^{Val}_1$(1420 pmol/A_{260} unit) and $tRNA^{Glu}_2$ (1500 pmol/A_{260} unit) were from Subriden, Rolling Bay, WA. $tRNA^{fMet}_3$ (m^7G46 replaced by A; 1750 pmol/A_{260} unit)[1] was a gift from Dr. A. D. Kelmers, Oak Ridge National Laboratories. Natural and synthetic 23S RNA were prepared as described previously (Weitzmann *et al.*, 1990). We thank K. Nurse for preparation of the synthetic 23S RNA. Hydrazine and *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium)-ethylcarbodiimide *p*-tosylate (CMC) were obtained from Aldrich. Primers for reverse transcription were made on an Applied BioSystems Model 394 synthesizer, deprotected in 30% NH_3 (14.8 M) at 55 °C for 18 h, desalted by 1-butanol treatment (Sawadogo & Van Dyke, 1991), and precipitated twice from 0.2 M KOAc, pH 5.0, with 2 volumes of EtOH. AMV reverse transcriptase was from Promega.

Reaction with CMC. The dried rRNA or tRNA sample (8 μg) was treated with 30 μL of 0.17 M CMC in 50 mM Bicine, pH 8.3, 4 mM EDTA, and 7 M urea at 37 °C for 20 min. Reaction was stopped with 100 μL of 0.3 M NaOAc and 0.1 mM EDTA, pH 5.6 (buffer A) and 700 μL of EtOH. After being chilled for 5 min in dry ice, the pellet was recovered, washed with 70% EtOH–water, dissolved in 100 μL of buffer A, and reprecipitated with 300 μL of EtOH. After being washed as above, the pellet was dried, dissolved in 40 μL of 50 mM Na_2CO_3 buffer, pH 10.4 (except as indicated), and incubated at 37 °C for 2 and 4 h (except as indicated). The RNA was precipitated by addition of 100 μL of buffer A and 700 μL of EtOH and placed on dry ice for 5 min. The pellet was washed twice with 70% EtOH–water, dried, and dissolved in 40 μL of water.

Reaction with Hydrazine. This method is modified from that described by Peattie (1979). The rRNA or tRNA sample (10 μg) was dried in an Eppendorf tube and then treated at 0 °C with 10 μL of a 1:1 mixture of hydrazine and water. The hydrazine solution can be kept for 12 weeks or longer at 4 °C, whereas anhydrous hydrazine discolors rapidly upon exposure to air. Reactions were run for 2–8 min as indicated before being quenched with 200 μL of buffer A. After addition of 750 μL of EtOH and chilling for 5 min on dry ice, the precipitate was recovered by centrifugation. The pellet was washed with 70% EtOH–water, dissolved in 200 μL of buffer A, reprecipitated with 600 μL of EtOH, and washed as above. The pellet was dried and 20 μL of 1 M aniline adjusted to pH 4.5 with HOAc was added. After 20 min at 60 °C, the solution was lyophilized to dryness. Following two cycles of solution in 20 μL of water and relyophilization, the sample was dissolved in 40 μL of water and used for reverse transcription as described below.

Reverse Transcription and Gel Analysis. The procedure is modified from Bakin and Ofengand (1992). For screening 23S RNA, primers complementary to residues 817–837, 1667–1684, 1896–1914, 1961–1980, 2115–2125, 2550–2569, and 2647–2665 were used. For tRNAs, the primers used are indicated in the figure legends. Reverse transcription was performed using [α-^{32}P]dATP or 5′-^{32}P-labeled primers. For hybridization, the mixture contained 0.15–0.3 pmol of RNA, 45 pmol of unlabeled primer or 2–4 pmol of 5′-labeled primer [(0.2–0.4) × 10^6 cpm/pmol] and hybridization buffer (50 mM Tris-HCl, pH 8.5, 20 mM KCl) in a total volume of 5 μL. The mixture was heated at 70 °C for 3 min, then at 37 °C for 5 min, and placed on ice for 2 min or longer. For labeling with [α-^{32}P]dATP, 5 μL of hybridization mixture and 5 μL of a mixture containing 0.8 pmol [α-^{32}P]dATP (1.5 μCi/pmol), 40 pmol each of dGTP, dCTP, and dTTP, 100 mM Tris-HCl, pH 8.5, 20 mM $MgCl_2$, 20 mM DTT, and 1.6 units of reverse transcriptase was incubated at 25 °C for 5–8 min. Reaction was stopped by chilling to 0 °C. The extension reaction was initiated by addition of 1 μL of 4 mM each dATP, dGTP, dCTP, and dTTP. Samples were incubated at 37 °C for 20 min. When 5′-labeled primers were used, the labeling step was omitted and 5 μL of hybridization mixture was combined with 5 μL of a mixture containing 100 mM Tris-HCl, pH 8.5, 20 mM $MgCl_2$, 20 mM DTT, 4 nmol each of dATP, dGTP, dCTP, and dTTP, and 1.6 units of reverse transcriptase and incubated at 37 °C for 30 min. Sequencing lanes were prepared using 0.32 pmol of synthetic 23S RNA, 7.5 units of reverse transcriptase, and 1.2 pmol of [α-^{32}P]-dATP (1.5 μCi/pmol) as described previously (Bakin & Ofengand, 1992), except that the labeling reaction was done at 25 °C for 5–8 min and the sequencing reaction time was decreased from 45 to 20 min. Reverse transcription was stopped by the addition of 2 μL of a solution containing 0.5 mg/mL RNase A and 0.1 M EDTA, pH 6.5. Samples were incubated at 37 °C for 45 min and precipitated with 3 volumes of ethanol. After being dried, the pellet was dissolved in 10–16 μL of 75% formamide, 1.4 mM Tris-HCl, pH 7.4, 1.4 mM EDTA, and 0.08% each xylene cyanol FF and bromophenol blue and analyzed on an 8% polyacrylamide–7 M urea gel. For tRNAs, 10% polyacrylamide gels were used.

RESULTS

The Concept behind the Method. Two complementary approaches were taken to the sequence localization of Ψ residues in RNA by reverse transcription. The first approach takes advantage of the unique stability of N_3-CMC-Ψ to alkaline hydrolysis (Ho & Gilham, 1971). Whereas U and G adducts of CMC are readily cleaved by alkaline conditions (Ho & Gilham, 1971), cleavage of N_3-CMC-Ψ requires 7 M NH_4OH at 100 °C for 8 min (Ho & Gilman, 1971). The CMC group at the N_3 position of Ψ efficiently blocks reverse transcription as shown for N_3-CMC-uridine (Moazed *et al.*, 1986). Therefore, after treatment, bands should appear one base downstream from Ψ sites, and possibly at the site itself as a result of "stuttering" (Denman et al., 1988), while bands at U and modified U sites should be absent (Figure 2A). Partial modification should be detectable as a positive band, although the intensity should be less.

The second approach depends upon the greater resistance of Ψ and m^5U residues to hydrazinolysis compared to U and

[1] Modified residues are abbreviated as follows: a lowercase letter (or letters) indicates the type of modification, a superscript number indicates the position of the modification, an uppercase letter represents the original residue, and a number following the uppercase letter indicates the position of the residue in the RNA molecule. m, methyl; s, thio; o, oxyacetic acid; mam, methylaminomethyl. m following the residue denotes a 2′-OCH_3 nucleoside.

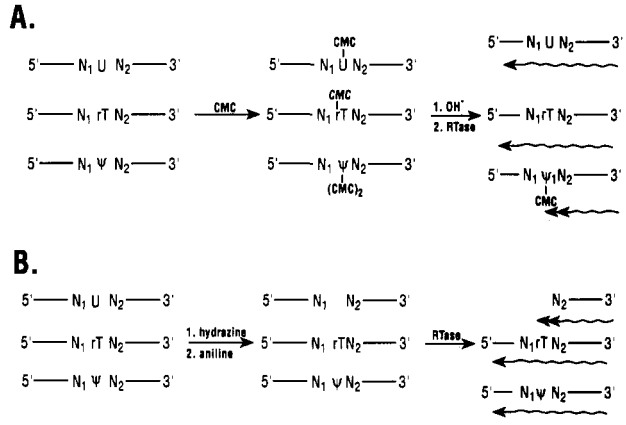

FIGURE 2: Schematic diagram of the two complementary procedures for detecting Ψ residues in any RNA molecule. (A) CMC–alkali treatment; (B) hydrazine–aniline treatment.

most other U-derived bases (Davis & Allen, 1957; Verwoerd & Zillig, 1963; Peattie, 1979; Lankat-Buttgereit *et al.*, 1987). The reaction with hydrazine leads to pyrimidine ring opening (Brown, 1967) with consequent loss of the ability to base-pair. This causes termination of reverse transcription one base 3′ to the ring-opened residue and results in a strong gel band one base downstream from the affected site and a weak band at the site itself due to "stuttering" (Figure 2B). Since reverse transcription is sensitive to the loss of base-pairing potential, we expected that cleavage of the polynucleotide chain by aniline treatment, as in the Peattie (1979) procedure, would be unnecessary. However, it was found empirically that inclusion of the aniline step resulted in an improved reproducibility of the band pattern. The method detects Ψ, m^5U, and other resistant bases as the absence of a band upon gel electrophoresis. This readout is the opposite of the first method. Partial modification can be difficult to detect since a band would still be present, although at a lesser intensity. On the other hand, the absence of a band should be a good indicator of virtually complete modification.

Application of the Method to 23S RNA. (A) Reaction with CMC. Four Ψ residues in *E. coli* 23S rRNA have been described in the literature, at positions 746, 1911, 1917, and 2555 (Noller, 1984; Raué *et al.*, 1988). An additional 4–5 sites remained to be determined, based on total Ψ analyses of 8–9 residues per chain (Gehrke & Kuo, 1989; Dubin & Günalp, 1967). We initiated our search in the vicinity of the above-mentioned Ψ positions with the aim of both validating the methodology and verifying the "established" sites.

We first examined the 1900 region of 23S RNA (Figure 3, panel 1) because Ψ residues were reported at positions 1911 and 1917 and an m^5U residue was reported at position 1939. Since it was known that CMC-m^5U hydrolyzes more slowly than CMC-U (Ho & Gilham, 1967) this region provided a good test of the specificity of the method. As shown in Figure 3, panel 1, strong stops were observed at positions one base 3′ to 1911 and 1917 in natural 23S RNA treated with CMC which did not decrease with incubation at either pH 10.4 or 10.6. In contrast, the weaker stops at these positions in CMC-treated synthetic RNA disappeared with increasing time of alkali treatment. The maximum effect can be seen by comparing the 4-h lanes of natural and synthetic RNA. These stops were not seen in natural RNA reacted in the absence of CMC. There is, in addition, a strong stop at position 1916 corresponding to the unknown modified U at 1915. This stop is seen with or without CMC treatment and is also present in natural RNA not exposed to alkali (Figure 4, panel 1) but is absent from treated or untreated synthetic RNA. Although m^5U1939 in natural RNA appeared more resistant than U1939 in synthetic RNA after 2 h of incubation in alkali, both stops were essentially gone after 4 h at pH 10.4 or 10.6. That there is m^5U1939 in this natural RNA preparation is shown below (Figure 4, panel 1). Figure 3, panel 1, also shows the presence of a stop at 1836, corresponding to the known modified residue m^2G1835. Additional gels run for longer times confirm that the stop occurred at 1836 (data not shown). Since 4 h at pH 10.4 appeared sufficient to remove all unwanted CMC and the control RNA at pH 10.6 showed a slight increase in degradation, pH 10.4 was chosen for all subsequent studies.

Panel 2 of Figure 3 shows the same approach applied to the 750 region which is known to contain the sequence m^1-G745Ψ746m^5U747. The m^1G residue was detected as a strong stop in the control lane as well as the CMC-treated natural RNA lanes. However, only the CMC-treated natural RNA gave a strong band at 747 corresponding to Ψ746. There was also a weak band at 746 itself. No other band of comparable intensity was found anywhere else, either in the control or CMC-treated lanes. In particular, note that the weak band at 748, corresponding to m^5U747, disappeared upon further incubation at pH 10.4.

The same approach was used to verify the location of Ψ2555, the fourth known Ψ residue (Figure 3, panel 3). To our surprise, although two new Ψ sites were detected, at 2580 and 2605, no evidence for a Ψ residue at 2555 could be found even though a weak stop in the control lane corresponding to Um2552 was detected. The strong bands at 2605 and 2606 could indicate the presence of adjacent Ψ residues at 2604 and 2605 or stuttering at a single Ψ at 2605. The hydrazine reaction (see Figure 4, panel 3, below) showed that there was only one Ψ residue at 2605. A third new Ψ residue was found at position 2504. This is more clearly seen in panel 4 of Figure 3. Here the double bands corresponding to Ψ2504 are clearly visible. Another Ψ residue, at position 2457, was also found. Note that while Ψ746, 1911, 1917, and 2580 gave predominantly single bands one base 3′ to the Ψ residue, Ψ2457, 2504, and 2605 gave double bands even though the sequence shows that there is only one available U residue at 2457 and 2504. This phenomenon can add uncertainty in the case of adjacent U residues. It is therefore advisable to confirm the identification of Ψ sites in such cases by means of the hydrazine reaction. There are also several other interesting features in this panel. No reverse transcriptase pause corresponding to m^6A2503 (Smith *et al.*, 1992) was found, but a very strong double stop which should be due to a modified C2501 was observed. The stop corresponding to Cm2498 was also seen in all of the natural RNAs.

A somewhat complicated set of stops at positions 2446, 2448, 2449, and 2450 were also seen. Careful analysis in conjunction with other gels (not shown) has demonstrated that the 2446 stop corresponds to a modified G at 2445 which according to Smith *et al.* (1992) is an m^2G. The three remaining bands all appear to be due to the modified U at 2449 when it is exposed to pH 10.4 for 4 h, since these latter bands were not seen when the alkaline treatment was omitted (data not shown).

The conclusion from this analysis is that four new sites for Ψ have been found in 23S RNA and that Ψ2555 is an erroneous assignment.

(B) Reaction with Hydrazine. A similar series of reverse transcriptions were done using 23S RNA treated with hydrazine (Figure 4). Comparison of the U sequencing lane with reacted synthetic, unmodified, 23S RNA in each of the

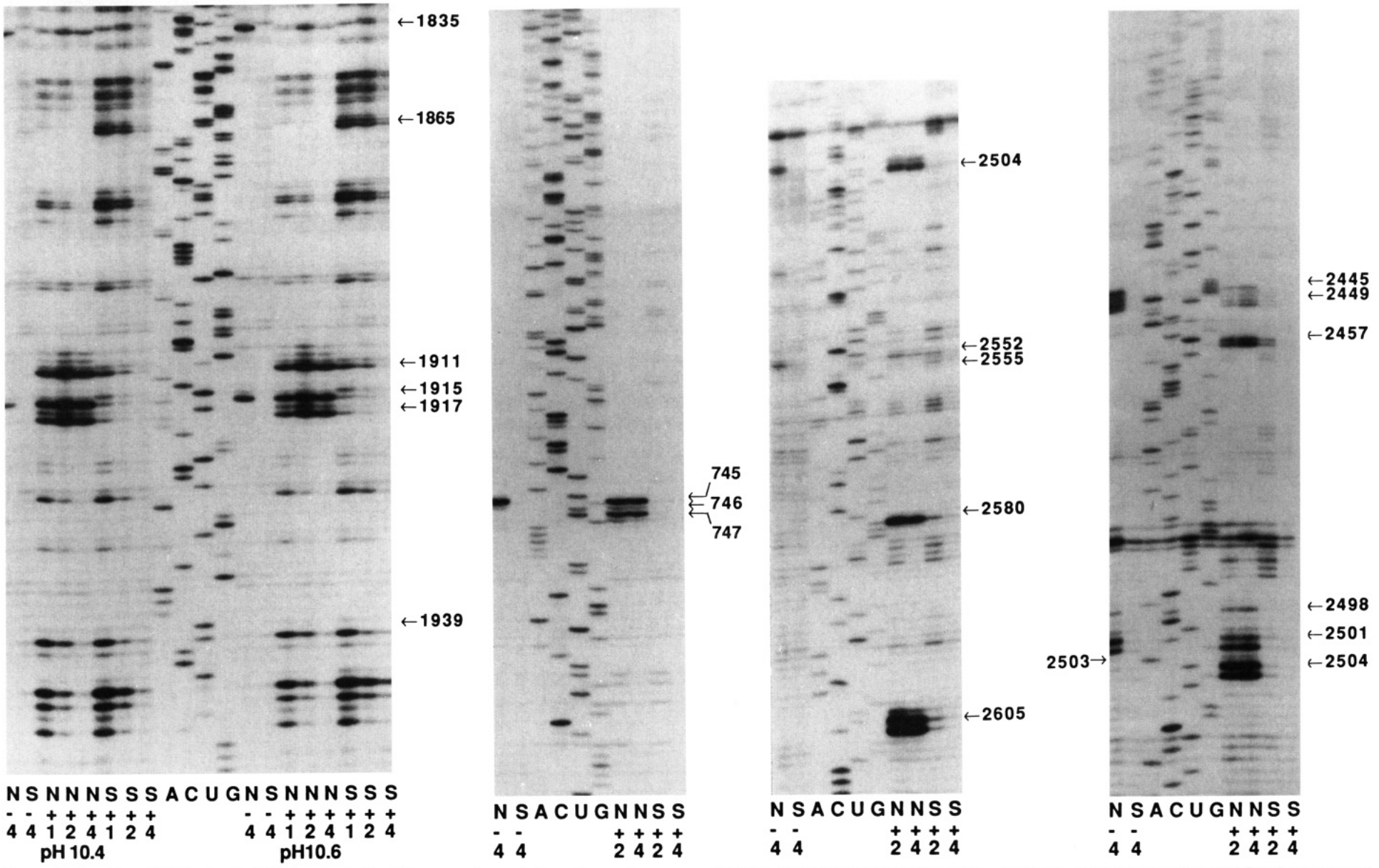

FIGURE 3: Location of Ψ sites by the CMC method. Panel 1, reverse transcription using a primer corresponding to residues 1961–1980. Samples were treated with (+) or without (–) CMC for 20 min at 37 °C, exposed to pH 10.4 or 10.6 for 1–4 h at 37 °C, and reverse transcribed. N, natural 23S RNA; S, synthetic 23S RNA. A, C, U, and G are sequencing lanes using synthetic RNA. Residue numbers of positions of interest are indicated. Panel 2, reverse transcription using primer 817–837. Samples were treated with (+) or without (–) CMC for 20 min at 37 °C and then exposed to pH 10.4 for 2 or 4 h at 37 °C. N, natural 23S RNA; S, synthetic 23S RNA. A, C, U, and G are sequencing lanes with synthetic RNA. Panel 3, reverse transcription using primer 2647–2665. Legend as in panel 2. Panel 4, reverse transcription using primer 2550–2569. Legend as in panel 2.

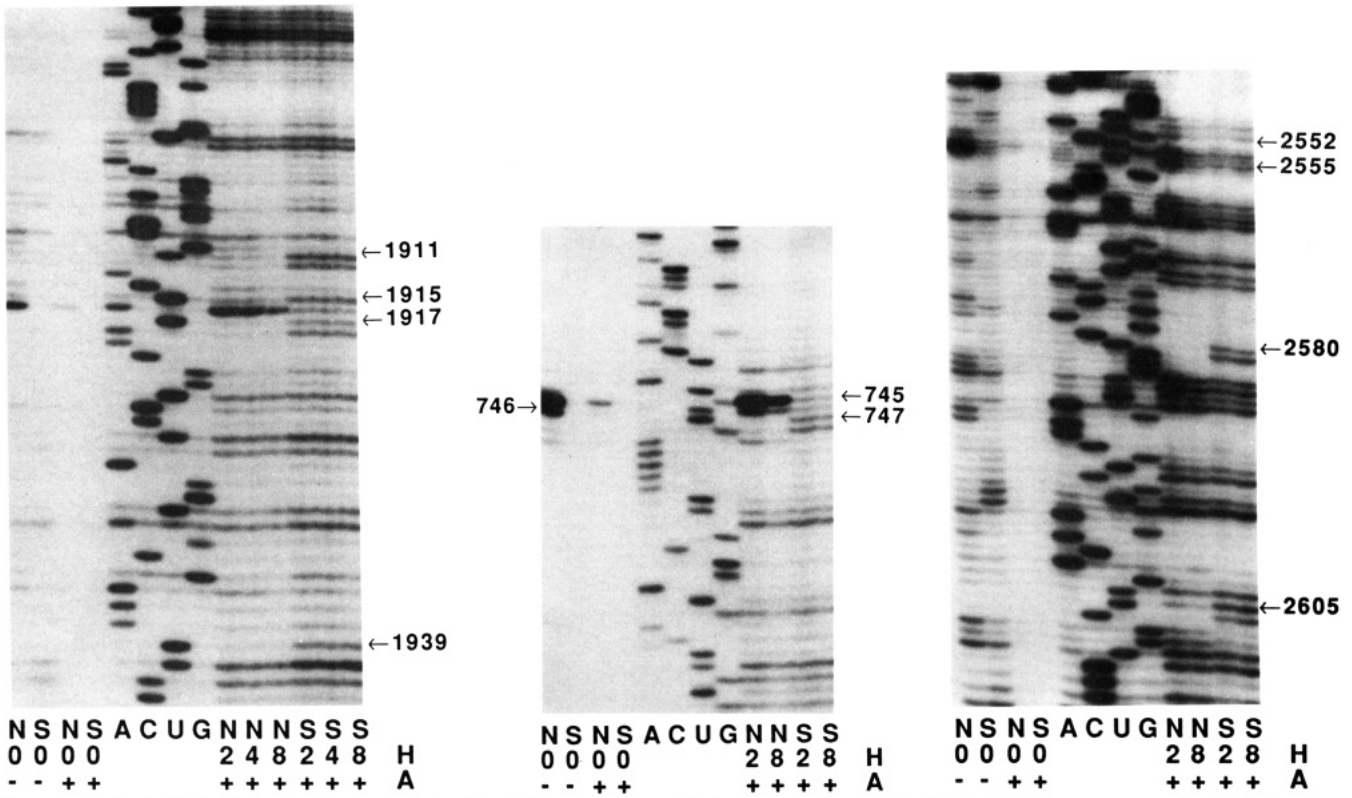

FIGURE 4: Analysis by the hydrazine procedure. Panel 1, reverse transcription using primer 1961–1980. Samples were treated with hydrazine (H) for the indicated times in minutes, followed by aniline (A) as indicated, and reverse transcribed as described in Materials and Methods. N, natural RNA; S, synthetic RNA. A, C, U, and G are sequencing lanes using synthetic RNA. Sequence numbers of modified bases are indicated. Panel 2, reverse transcription using primer 817–837. Legend as in panel 1. Panel 3, reverse transcription using primer 2647–2665. Legend as in panel 1.

panels shows that at every uridine residue there is a strong band one base 3′ to the U which corresponds to the 5′-most base still capable of base pairing plus a weaker band at the site of the degraded pyrimidine ring itself. Since base pairing should not be possible at this site, we attribute this band to some type of stuttering effect, similar to that described previously (Denman *et al.*, 1988).

Comparison of the hydrazine-treated natural and synthetic RNA lanes in Figure 4, panel 1, show only four places where the bands due to U residues are absent. Both bands at 1911, 1912 and 1917, 1918 are absent, corresponding to the two Ψ residues (Figure 3, panel 1). Only one band at 1939, due to m^5U1939, is missing because U1940 generates the usual two bands at 1940 and 1941. One band, at 1866, is also missing, suggesting a Ψ or m^5U at 1865. Since there was no stable adduct with CMC (Figure 3, panel 1), this position should correspond to m^5U. However, sequencing of the isolated rRNA rather than the *in vitro* transcript of the *rrnB* gene revealed that while the base was U in the transcript, it was C in the RNA isolated from *E. coli* MRE600 (data not shown). A C/U heterogeneity at this site has been noted previously (Smith *et al.*, 1992). Overall, essentially the same results were obtained when the reaction with hydrazine was carried out for 2 or 8 min. The strong band at 1916 seen also in untreated natural RNA arises from the unknown modified U at 1915 (see also Figure 3, panel 1). Thus, the hydrazine reaction detects both Ψ and m^5U under these conditions.

Figure 4, panel 2, shows the 750 region. The strong stop at m^1G745 is clearly visible in both treated and untreated natural RNA. In the hydrazine-treated samples, bands are absent only in the natural RNA lanes and only at positions 747 and 748. These missing bands correspond to Ψ746 and m^5U747. Band 746 should also have been missing but is obscured by the band due to m^1G745. Note that here also all other U residues generate a double band in both natural and synthetic RNAs.

Similar results were obtained in the 2600 region (Figure 4, panel 3). The bands at 2580 and 2581, present in synthetic RNA, were absent in natural RNA, whereas U residues at 2584, 2585, 2586, 2593, and 2596 all gave double bands with both RNAs. Thus, U2580 is either Ψ or m^5U by this analysis. CMC analysis (Figure 3, panel 3) showed it to be Ψ. U2604 and U2605 generate three bands in synthetic RNA but only two bands in natural RNA. This identifies U2604 as being unmodified and shows that only U2605 is Ψ or m^5U. The combined results from Figure 3, panel 3, and Figure 4, panel 3, establish U2605 conclusively as a Ψ residue.

In the 2500 region, all of the U residues produced double bands in both natural and synthetic RNA except for U2504 and U2457, where the double bands were absent only from the natural RNA (data not shown). Thus these two sites were either Ψ or m^5U and the data of Figure 3, panel 4, established them both to be Ψ.

Reaction of Other Modified Uridine Residues. The reactivity of the modified uridines s^4U8, o^5U34, and mam^5s^2U34, so far only found in tRNA, to the CMC–alkali procedure was tested by analysis of purified *E. coli* tRNAs with known modified nucleotide sites (Sprinzl *et al.*, 1991). $tRNA_1^{Val}$ (s^4U8, o^5U34), $tRNA_2^{Glu}$ (Ψ13, mam^5s^2U34), and RNA_3^{fMet} (s^4U8) were used (Figure 5). The o^5U34 residue is representative of all 5′-OR-substituted uridines, whereas mam^5s^2U34 is the archetype of all 5′-CHR_2-substituted uridines as well as of s^2U derivatives. It was not possible to test the dihydrouridine (D) residues in these tRNAs because exposure to pH 10.4 for 2 h at 37 °C was sufficient to open the dihydrouridine ring as evidenced by strong reverse tran-

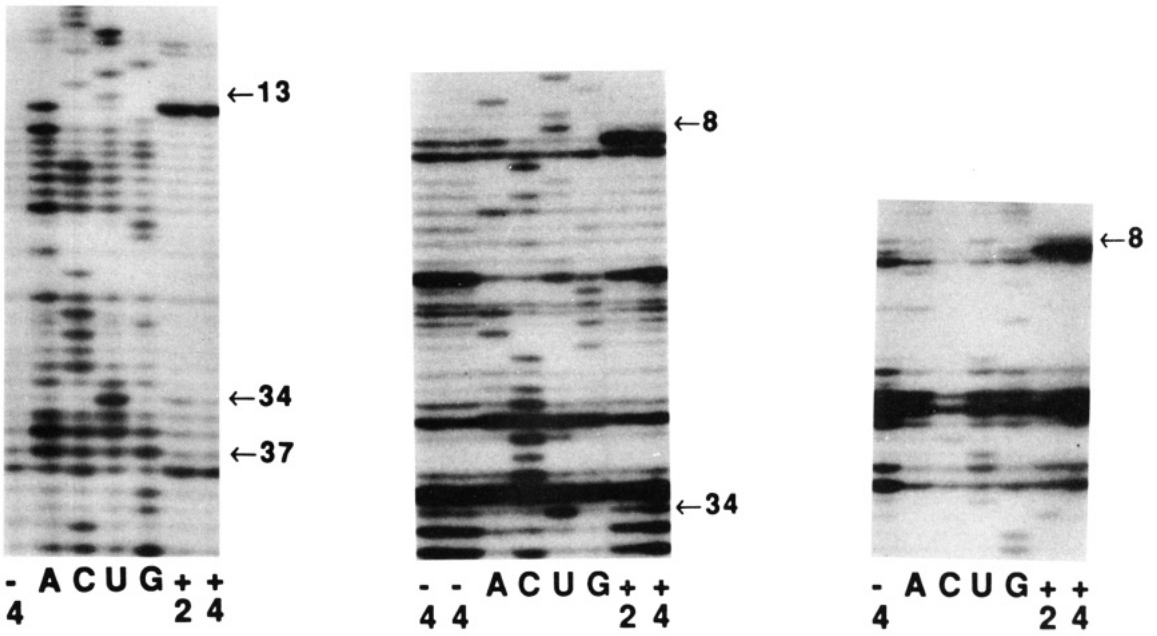

FIGURE 5: Reaction of modified bases in tRNA with CMC. Panel 1, reverse transcription of $tRNA^{Glu}_2$ using a primer corresponding to residues 62–79. Legend as in Figure 3, panel 2. Positions of interest are marked. Panel 2, reverse transcription of $tRNA^{Val}_1$ using primer 61–77. Legend as in Figure 3, panel 2. Panel 3, reverse transcription of $tRNA^{fMet}_3$ using primer 61–76. Legend as in Figure 3, panel 2.

scriptase stops in the absence of CMC (Figure 5, panels 2 and 3). Figure 5, panel 1, shows that while Ψ13 of $tRNA^{Glu}_2$ was readily detected, no such bands could be found corresponding to the mam^{5}s^{2}U34 residue. The only other band is due to a reverse transcriptase stop one base 3′ to m^{2}A37. Similarly, Figure 5, panel 2, shows that the CMC adduct of s^{4}U8 of $tRNA^{Val}_1$ is sufficiently stable to these alkaline conditions to yield a strong stop. The o^{5}U34 residue does not appear to yield a stable CMC adduct since the band one base 3′ to U34 in the sequencing lane is not stronger than that in the minus CMC lanes, but as the bands in the control lanes were relatively strong, a small extent of stable product cannot be ruled out. Figure 5, panel 3 ($tRNA^{fMet}_3$) confirms that the s^{4}U8 residue yields a stable adduct.

The hydrazine–aniline reaction was not tested since Lankat-Buttgereit *et al.* (1987) had already reported that D, s^{4}U, o^{5}U, and mam^{5}s^{2}U all reacted, although not as strongly as U. They also stated that m^{5}U reacted weakly and Ψ was unreactive. Under our conditions, m^{5}U as well as Ψ was unreactive.

DISCUSSION

The Method. Up to now the location of Ψ residues in large RNA molecules has been a daunting task because there is no convenient way to label the modified residue, as for example is the case with methylated or thiolated bases. The only useful assay for Ψ formation, release of ^{3}H from the C_5 position of the uracil ring, does not lend itself to site localization. We have solved this problem by coupling two previously described chemical reactions selective for Ψ with primed reverse transcription, which allows any area of a large RNA molecule whose sequence is known to be probed.

Although it has been claimed that only Ψ is completely unreactive to hydrazine–aniline, m^{5}U being weakly reactive (Lankat-Buttgereit *et al.*, 1987), we were unable to find an extent of treatment which would on the one hand react with the m^{5}U base and on the other hand still preserve a sufficient fraction of unreacted U residues so as to allow a reasonable amount of chain extension by reverse transcriptase. Whenever we increased the degree of exposure to hydrazine in order to react the m^{5}U residues, chain termination became so severe that no useful information could be obtained. Therefore, while U residues could be distinguished from Ψ, m^{5}U residues could not.

Fortunately, a well-characterized but largely ignored reaction, CMC addition followed by alkaline hydrolysis, had earlier been found to be quite specific for Ψ (Ho & Gilham, 1971). CMC-m^{5}U was labile under the conditions employed here, as were also mam^{5}s^{2}U and, probably, o^{5}U. Although s^{4}U also yielded a stable CMC adduct, it can be distinguished from Ψ because it is reactive to hydrazine–aniline (Lankat-Buttgereit *et al.*, 1987). Another advantage of using both methods is the complementary nature of their readout. The CMC-OH procedure produces a gel band at sites of Ψ residues, whereas the hydrazine–aniline procedure results in the absence of a band. Consequently, partial modification of U to Ψ would show up as Ψ in the former procedure but would probably be scored as U in the latter one. Thus, while the CMC-OH method is more specific, the hydrazine–aniline procedure is useful as a way to estimate the extent of modification.

It is very helpful in these analyses to have completely unmodified RNA as a control. This is shown quite clearly in Figures 3 and 4. Fortunately, with the ability to make virtually any RNA by T7 or SP6 RNA polymerase once the DNA gene sequence is known and available, this is not a limiting factor. Certainly our interpretation of the gels of Figure 3 and 4 was greatly assisted by the availability of synthetic 23S RNA (Weitzmann *et al.*, 1990).

The combination of the two methods is also useful in cases of sequence errors. For example, we initially believed there was an m^{5}U residue at U1865 because of the absence of bands in the hydrazine–aniline method at that position in the synthetic unmodified RNA sequencing lane as well as the absence of bands in the CMC-OH reaction. However, upon sequencing the natural RNA, we discovered that U1865 in the *rrnB* gene was almost entirely C1865 in the RNA isolated from ribosomes of the MRE600 strain.

Location of Ψ Residues in E. coli 23S RNA. By the method described here, 941 bases, or about 32% of the 23S RNA, have been screened for Ψ residues (Figure 6). Although the

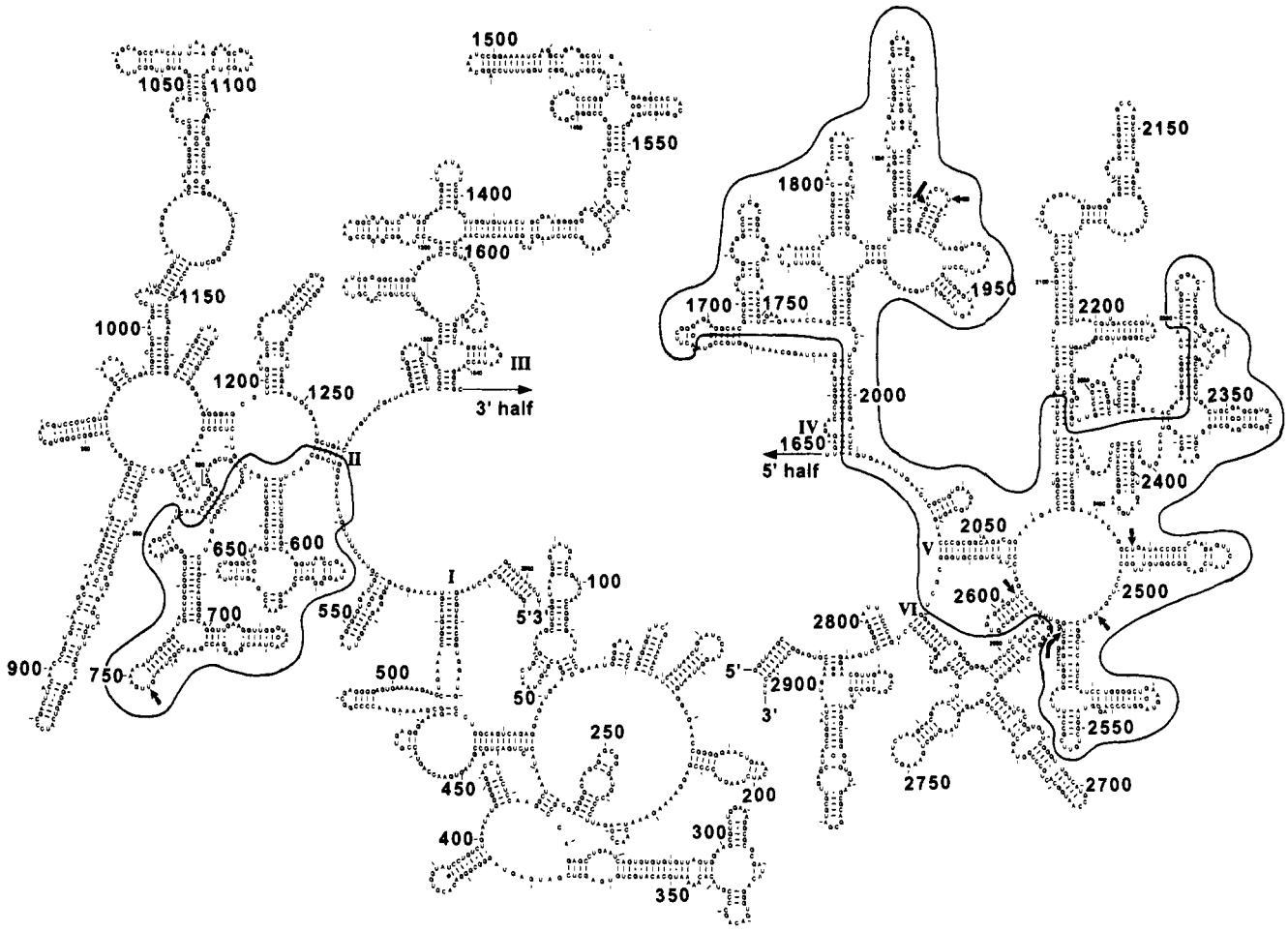

FIGURE 6: Secondary structure of 23S RNA showing the regions screened for Ψ residues. The secondary structure, redrawn from Gutell (1993), was generously provided by Robin Gutell. The outlined regions were screened for Ψ residues by reverse transcription after treatment with hydrazine and with CMC. The locations of the seven Ψ residues are indicated by arrows.

Table I: Location of Pseudouridine Residues in 23S RNA

position	reference
746	Branlant *et al.* (1981)
1911	Branlant *et al.* (1981)
1917	Branlant *et al.* (1981)
2457	this work
2504	this work
2580	this work
2605	this work
2555[a]	Noller (1984)

[a] Not detected in this study.

percentage covered is not large, as Figure 6 shows, the three known peptidyltransferase regions were completely checked. The position of each of the Ψ residues currently known is listed in Table I. The four new sites for Ψ in 23S rRNA plus the three previously known and confirmed in this work may account for all of the Ψ residues present in *E. coli* 23S RNA. There could, however, be one or two more given the inherent error in total Ψ analyses. In order to determine this rigorously, it would be necessary to screen the rest of the molecule either by the method outlined in this work or else by cutting out the screened regions by site-directed RNase H cleavage and doing a total analysis for Ψ on the remaining unscreened segments.

Nevertheless, the positions already detected show a remarkable and unexpected clustering around the peptidyltransferase center (PTC). As shown in Figure 7, all four of the new sites, 2457, 2504, 2580, and 2605, are at or immediately adjacent to the PTC. In addition, the stem containing Ψ1911 and Ψ1917 has been localized near the PTC by RNA–RNA cross-linking within the 23S RNA [reviewed in Brimacombe *et al.* (1993)], by cross-linking of G1945 to the 3′-end of tRNA (Wower *et al.*, 1989), and by protection of C1941 by the 3′-end of tRNA (Moazed & Noller, 1989, 1991). The Ψ746-containing stem is also near the PTC according to RNA–RNA cross-linking [reviewed in Brimacombe *et al.* (1993)] and vernamycin footprinting of A752 (Figure 7a), as well as A2058, A2062, and G2505 (Figure 7c) (Moazed & Noller, 1989). Thus all seven of the identified Ψ residues are at or near the site of peptide bond formation. Moreover, at least some of the Ψ residues, namely, 746, 2457, 2504, 2580, and 2605, are in intimate proximity to nucleotides which are directly involved in functional aspects of the peptide bond formation process (Figure 7).

Other Modified Bases in 23S RNA. Recent work has identified four previously unknown methylated nucleotides in *E. coli* 23S RNA, namely, m^2G1835, m^5C1962, m^2G2445–7, and m^6A2503 (Smith *et al.*, 1992). These and others previously identified and localized are listed in Table II. mA2071 has not been further identified. In addition, two residues of m^4C and one each of m^2A and m^3U have been reported but not localized (Gehrke & Kuo, 1989). It is believed that mU1915 is not m^3U, m^5U, or Um and that there are no methyl groups in U*2449 (Smith *et al.*, 1992).

Our analyses confirmed the presence of modified nucleotides at positions corresponding to m^1G745, m^2G1835, mU1915, m^7G2069, m^2G2445–7, U*2449, Cm2498, and Um2552 by finding reverse transcriptase stops at the appropriate sites

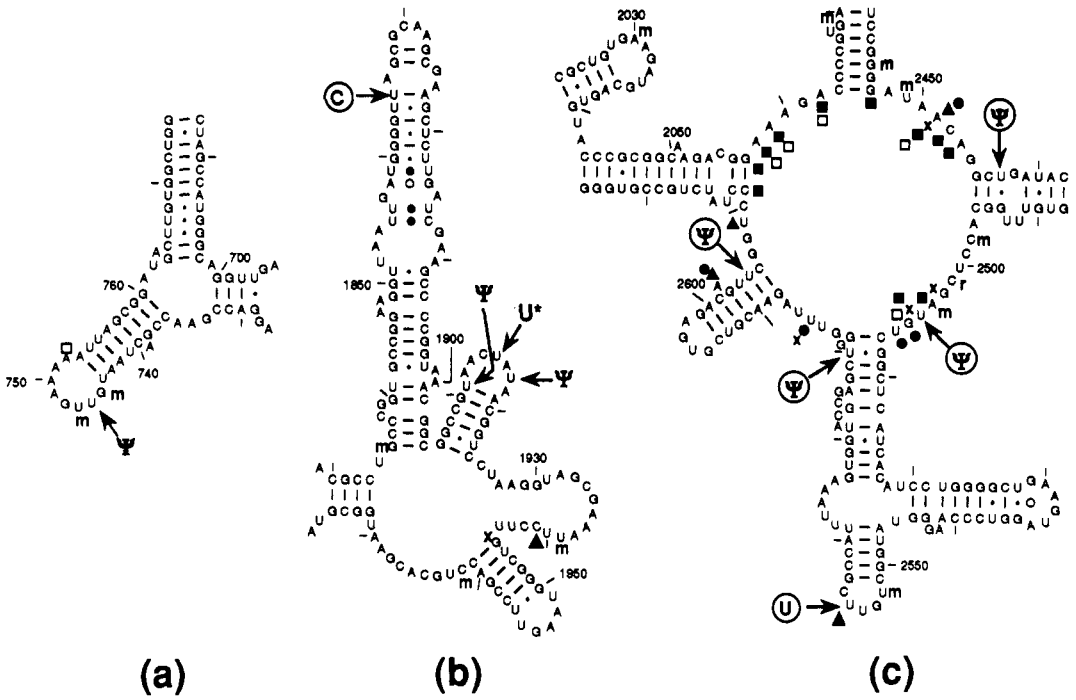

FIGURE 7: Secondary structure of selected segments of 23S RNA showing the location of the Ψ residues. Secondary structures were reproduced from Gutell (1993). The seven Ψ sites are indicated by arrows. U2555, which was misidentified as Ψ, C1865, which is U in the *rrnB* gene, U*1915, which is an unidentified modified uridine, are also indicated. The circles denote sites which were determined in this work. ×, sites of cross-linking to the aminoacyl end of tRNA or its puromycin analog (Steiner *et al.*, 1988; Hall *et al.*, 1988; Wower *et al.*, 1989); **m**, sites of methylation (Smith *et al.*, 1992); **r**, base modification which causes reverse transcriptase double stop (this work, Table II); ● (P site) and ▲ (A site), sites of footprinting by the aminoacyl end of tRNA (Moazed & Noller, 1989, 1991); ■, antibiotic resistance sites (Douthwaite, 1992); □, antibiotic protection sites (Moazed & Noller, 1987).

Table II: Location of Modified Bases Other than Pseudouridine in 23S RNA

position	reported[a]	hydrazine[b]	CMC[c]	RvTase[d]	RvTase stops[e]
G745	m^1G			+	745
U747	m^5U	+	–	–	
A1618	m^6A			–	
G1835	m^2G			+	1836
U1915	mU			+	1916
U1939	m^5U	+	–	–	
C1962	m^5C			–	
A2030	m^6A			–	
G2069	m^7G			+	2069, 2070
A2071	mA			–	
G2251	Gm			nt[f]	
G2445–7	m^2G			+ (2445)	2446
U2449	U*			+[g]	2448–2450
C2498	Cm			+	2498
C2501	–			+	2501, 2502
A2503	m^6A			–	
U2552	Um	–	–	+	2553

[a] Smith *et al.* (1992) and Noller (1984) and references cited therein. [b] Uridines only. Failure to open the pyrimidine ring and produce a reverse transcriptase stop is scored as +. [c] Failure to make an alkali-stable adduct which produces a reverse transcriptase stop is scored as –. [d] Detected as a gel band at or one base 3′ to the modified base. [e] Location of gel bands after reverse transcription. [f] nt, not tested. [g] Bands seen only after the alkali treatment as used for the CMC reaction. No bands were seen with untreated RNA.

(Table II) and, in the case of the 2445–7 site, showed that the modified G residue was located at position 2445. The data also confirmed the presence of m^5U747 and m^5U1939 as U-derived residues which do not react with hydrazine (scored +) but which do not make an alkali-stable CMC adduct (scored –). These properties are characteristic of m^5U residues. In our hands, reverse transcriptase was not hindered by m^6A1618, m^6A2030, or m^6A2503, although weak pauses were seen at positions 2033 and 2532. These sites might also be modified in some way. There is in addition a strong double stop indicating the presence of a modified C2501. This base could be one of the two m^4C residues reported by Gehrke and Kuo (1989). We found no evidence for a stop at or next to A2071, suggesting that if this base is truly methylated, it is unlikely to be the m^2A residue reported by Gehrke and Kuo (1989).

Function of Ψ Residues in 23S RNA. Both Smith *et. al* (1992) and Brimacombe *et al.* (1993) have called attention to a clustering of the modified nucleotides at the PTC and their possible important role in ribosomal function. However, it is important to note that the two types of modification, methylation and pseudouridylation, differ markedly in their effect. Methylation, except for m^7G, which introduces a positive charge, increases local hydrophobicity whereas Ψ formation introduces an additional functional group at N_1 which is capable of H-bonding as well as facile acylation (Spector & Keller, 1958). In this connection it is worth pointing out that the conventional view that Ψ formation is an exotic event may be incorrect. In fact, it is a thrifty way for the cell to introduce modifications into RNA. While methylation requires not only a specific enzyme but also *S*-adenosylmethionine made from ATP and methionine, pseudouridylation requires no added energy source.

The clustering of Ψ residues around the PTC may be for structural reasons, as the extra hydrogen-bonding potential could be involved in a variety of ways. However, their close association with the PTC also suggests the alternate possibility that they may participate directly in the peptide bond formation process (Lane *et al.*, 1992). The presence of so many residues at the PTC need not negate this possibility if it is considered that peptide bond formation, being such an essential process, may require multiple Ψ residues at the active site to ensure the availability of at least one residue at all times. That is, all (or some) of the Ψ residues may be tightly clustered at the PTC so that at least one is always available for each round of peptide transfer. In our view, the best way to test this

premise is by purifying the enzymes which form Ψ so that the activity of 50S ribosomes with and without specific Ψ residues can be directly tested for their ability to catalyze peptide bond formation. Purification of these enzymes is currently underway (Wrzesinski *et al.*, 1993).

REFERENCES

Arena, F., Ciliberto, G., Ciampi, S., & Cortese, R. (1978) *Nucleic Acids Res. 5*, 4523–4536.

Bakin, A., & Ofengand, J. (1992) *BioTechniques 13*, 682–684.

Brand, R. C., Klootwijk, J., Sibum, C. P., & Planta, R. J. (1979) *Nucleic Acids Res. 7*, 121–134.

Branlant, C., Krol, A., Machatt, M. A., Pouyet, J., & Ebel, J. P. (1981) *Nucleic Acids Res. 9*, 4303–4324.

Brimacombe, R., Mitchell, P., Osswald, M., Stade, K., & Bochkariov, D. (1993) *FASEB J. 7*, 161–167.

Brown, D. M. (1967) *Methods Enzymol. 12A*, 31–34.

Chinali, G., Horowitz, J., & Ofengand, J. (1978) *Biochemistry 17*, 2755–2760.

Cohn, W. E. (1959) *Biochim. Biophys. Acta 32*, 569–571.

Cohn, W. E. (1960) *J. Biol. Chem. 235*, 1488–1498.

Davis, F. F., & Allen, F. W. (1957) *J. Biol. Chem. 227*, 907–915.

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

Douthwaite, S. (1992) *J. Bacteriol. 174*, 1333–1338.

Dubin, D., & Günalp, A. (1967) *Biochim. Biophys. Acta 134*, 106–123.

Gehrke, C. W., & Kuo, K. C. (1989) *J. Chromatog. 471*, 3–36.

Green, C. J., Kammen, H. O., & Penhoet, E. E. (1982) *J. Biol. Chem. 257*, 3045–3052.

Gutell, R. R. (1993) in *The Translational Apparatus* (Nierhaus, K. H., Subramanian, A. R., Erdmann, V. A., Franceschi, F., & Wittman-Liebold, B., Eds.), Plenum Press, New York (in press).

Hall, C. C., Johnson, D., & Cooperman, B. S. (1989) *Biochemistry 27*, 3983–3990.

Ho, N. W. Y., & Gilham, P. T. (1967) *Biochemistry 6*, 3632–3639.

Ho, N. W. Y., & Gilham, P. T. (1971) *Biochemistry 10*, 3651–3657.

Horowitz, J., Ou, C.-N., Ishaq, M., Ofengand, J., & Bierbaum, J. (1974) *J. Mol. Biol. 88*, 301–312.

Johnston, H. M., Barnes, W. M., Chumley, F. G., Bossi, L., & Roth, J. R. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 508–512.

Kammen, H. O., Marvel, C. C., Hardy, M., & Penhoet, E. E. (1988) *J. Biol. Chem. 263*, 2255–2263.

Krzyzosiak, W., Denman, R., Nurse, K., Hellmann, M., Boublik, M., Gehrke, C. W., Agris, P. F., & Ofengand, J. (1987) *Biochemistry 26*, 2353–2364.

Lane, B. G., Ofengand, J., & Gray, M. W. (1992) *FEBS Lett. 302*, 1–4.

Lankat-Buttgereit, B., Gross, H. J., & Krupp, G. (1987) *Nucleic Acids Res. 15* 7649.

Maden, B. E. H. (1990) *Prog. Nucleic Acid Res. Mol. Biol. 39*, 241–303.

Moazed, D., & Noller, H. F. (1987) *Biochimie 69*, 879–884.

Moazed, D., & Noller, H. F. (1989) *Cell 57*, 585–597.

Moazed, D., & Noller, H. F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3725–3728.

Moazed, D., Stern, S., & Noller, H. F. (1986) *J. Mol. Biol. 187*, 399–416.

Nichols, J. L., & Lane, B. G. (1967) *J. Mol. Biol. 30*, 477–489.

Noller, H. F. (1984) *Annu. Rev. Biochem. 53*, 119–162.

Ofengand, J., Bierbaum, J., Horowitz, J., Ou, C.-N., & Ishaq, M. (1974) *J. Mol. Biol. 88*, 313–325.

Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1760–1764.

Perret, V., Garcia, A., Grosjean, H., Ebel, J.-P., Florentz, C., & Giegé, R. (1990) *Nature 344*, 787–789.

Raué, H. A., Klootwijk, J., & Musters, W. (1988) *Prog. Biophys. Mol. Biol. 51*, 77.

Reddy, R. (1988) *Nucleic Acids Res. 16* (Suppl.), r71–r85.

Samuelsson, T., & Olsson, M. (1990) *J. Biol. Chem. 265*, 8782–8787.

Samuelsson, T., Borén, T., Johansen, T.-I., & Lustig, F. (1988) *J. Biol. Chem. 263*, 13692–13699.

Sawadogo, M., & Van Dyke, M. W. (1991). *Nucleic Acids Res. 19*, 674.

Scannell, J. P., Crestfield, A. M., & Allen, F. W. (1959) *Biochim. Biophys. Acta 32*, 406–412.

Schulman, L. H., (1991) *Prog. Nucleic Acid Res. Mol. Biol. 41*, 23–87.

Smith, J. E., Cooperman, B. S., & Mitchell, P. (1992) *Biochemistry 31*, 10825–10834.

Spector, L. B., & Keller, E. B. (1958) *J. Biol. Chem. 232*, 185–192.

Sprinzl, M., Dank, N., Nock, S., & Schön, A. (1991) *Nucleic Acids Res. 19*, 2127–2171.

Steiner, G., Kuechler, E., & Barta, A. (1988) *EMBO J. 7*, 3949–3955.

Tsui, H.-C. T., Arps, P. J., Connolly, D. M. & Winkler, M. E. (1991) *J. Bacteriol. 173*, 7395–7400.

Verwoerd, D. W., & Zillig, W. (1963) *Biochim. Biophys. Acta 68*, 484–486.

Weitzmann, C., Cunningham, P. R., & Ofengand, J. (1990) *Nucleic Acids Res. 18*, 3515–3520.

Wower, J., Hixson, S. S., & Zimmermann, R. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5232–5236.

Wrzesinski, J., Lane, B. G., & Ofengand, J. (1993) *FASEB J. 7*, A1092.

Yu, C.-T., & Allen, F. W. (1959) *Biochim. Biophys. Acta 32*, 393–406.