Nucleic Acids Research, 1994, Vol. 22, No. 18 3681–3684

The single pseudouridine residue in *Escherichia coli* 16S RNA is located at position 516

Andrey Bakin, Jeffrey A.Kowalak[1], James A.McCloskey[1,2] and James Ofengand*
Roche Institute of Molecular Biology, Roche Research Center, Nutley, NJ 07110 and Departments of [1]Biochemistry and [2]Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112, USA

Received June 29, 1994; Accepted July 27, 1994

ABSTRACT

The number and location of pseudouridine residues in *Escherichia coli* 16S ribosomal RNA has been determined by a combination of direct and indirect methods. Only one residue was found, at position 516. This site is at the 5′-end of one of the three most highly conserved long sequences of this RNA molecule. A number of experimental findings have strongly implicated this loop in the fidelity of codon recognition by A-site bound tRNA. By virtue of its location, we suggest that Ψ516 may also play a role in maintaining the fidelity of protein synthesis.

INTRODUCTION

The dominant role of RNA in ribosome function is now well established (reviewed in 1–3). Less well appreciated is the fact that rRNA contains a number of modified nucleosides (4) whose specificity of placement strongly implies that they play a purposeful role. These are nucleosides methylated either on the base or on the 2′-hydroxyl of ribose and pseudouridine (Ψ, 5-β-D-ribofuranosyluracil), the 5-ribosyl isomer of uridine, and its derivatives, such as 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine in eukaryotic small subunit rRNA (5, 6). Other modified nucleosides in rRNA not included in these categories are dihydrouridine in *E.coli* 23S rRNA (7, 8) and N^4-acetylcytidine (9, 10) and its 2′-*O*-methyl derivative (10) in other rRNAs.

By far the predominant single modified nucleoside in eukaryotic rRNA is Ψ, although if one combined all of the 2′-*O*-methyl nucleotides together, they would be almost as numerous (11). The unique glycosidic structure of Ψ results in an additional H bond donor site at the N-1 position, which is potentially acylatable (12). Despite these facts, very little is known about the role of Ψ in maintaining rRNA structure or in ribosome function (11).

A prerequisite to understanding the function of Ψ in the ribosome is knowledge of their number and specific sequence location in the RNA chain. Although the presence of pseudouridine in *E.coli* 16S rRNA has been known for a long time (13, 14), the location of Ψ has not been established. Mass spectrometric analysis of oligonucleotide fractions from complete RNase T_1 hydrolysates of *E.coli* 16S rRNA revealed that Ψ was present as a dinucleotide (ΨG), implying that it occurs in the sequence GΨG (15). Unfortunately, this sequence occurs 32 times in the 16S RNA molecule and is widely distributed.

A new, facile technique for locating Ψ residues in large RNA molecules was recently reported and the method used to locate four new Ψ sites in *E.coli* 23S RNA (16). We have now used this technique in conjunction with more direct methods to determine the number and location of Ψ residues in *E.coli* 16S RNA. We find that there is only one Ψ in this RNA and it is located at position 516. The site is in one of the three long, highly conserved, single-stranded sequences which are characteristic of small subunit ribosomal RNA in all organisms. Of particular interest is the fact that this sequence has been strongly implicated, by diverse criteria, in the fidelity of codon recognition at the ribosomal A site.

MATERIALS AND METHODS

Materials

General. Deoxyoligonucleotides for RNase H digestion and for primed reverse transcription were prepared as described previously (16). 16S rRNA was prepared by SDS–phenol extraction of 30S subunits (17). The colicin fragment, residues 1493–1542, from the ksgA strain of *E.coli* (18) was a gift from C.W.A.Pleij (Leiden University, The Netherlands).

16S RNA fragments. Four nmol (50 A_{260} units) of 16S RNA were incubated with 10 nmol of deoxyoligomer and 25 units of RNase H (Gibco-BRL) in 0.5 ml of 20 mM HEPES–KOH, pH 7.5, 100 mM KCl, 10 mM $Mg(OAc)_2$ and 0.1 mM DTT for 1 h at 37°C. After phenol extraction and ethanol precipitation, the RNA was dissolved in 20 mM HEPES, pH 7.5, 5 mM EDTA, heated to 70°C for 5 min, quickly cooled on ice, and fractionated on a 10–30% sucrose gradient in 20 mM HEPES, pH 7.6, 100 mM NH_4Cl, 1 mM EDTA by centrifugation (Beckman VTi50, 4.5 h, 45,000 r.p.m.). H2S and H1S fragments were purified through a sucrose gradient twice, and the H2L fragment three times. The heating–cooling cycle was repeated each time. The integrity and purity of the RNA fragments were examined by electrophoresis on a 4.5% polyacrylamide gel in 7 M urea.

*To whom correspondence should be addressed

Methods

Nucleoside analysis by directly combined liquid chromatography/mass spectrometry (LC/MS). The colicin fragment of *E.coli* 16S rRNA (25 μg) and RNase H-generated fragments of *E.coli* 16S rRNA (100 pmol) were hydrolyzed to nucleosides using nuclease P_1 and alkaline phosphatase as described (19), and analyzed using thermospray LC/MS. A description of the instrument, experimental procedures, and interpretation of data were previously reported (20, 21).

Sequence localization of pseudouridine. Chemical modification of 16S RNA by the *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium) ethylcarbodiimide *p*-tosylate (CMC)–alkali and hydrazine–aniline methods and reverse transcription analysis were carried out as described previously (16). Primers complementary to 16S RNA residues 604–624 and 413–433 were used.

RESULTS

Pseudouridine content of *E.coli* 16S RNA

Both early measurements of the amount of Ψ in 16S RNA (13, 14) and those made more recently (22, 23) fluctuated between one and two residues per RNA molecule. The existence of at most two Ψ residues suggested that they might be located in one or more of the three known long, single-stranded, highly conserved and functionally significant sequences, namely residues 516–535, 1397–1408, and 1492–1505. Thus, we decided to determine their exact localization.

Pseudouridine content of fragments of *E.coli* 16S RNA

Residues 1492–1505 were ruled out by analysis of the colicin fragment, residues 1493–1542. This fragment was prepared by colicin E3 digestion of 30S ribosomes (24). Analysis by LC/MS revealed the presence of m^3U1498 and m^2G1516 but no Ψ (data not shown). The tandem m^6_2A residues at positions 1518 and 1519 were not expected as the fragment was derived from the ksgA strain which lacks the appropriate methyl transferase for m^6_2A synthesis (18).

In another series of experiments, RNase H digestion was used to prepare the following fragments: H1S [nucleotides 1–(297–308)], H2S [nucleotides 1–(518–529)], and H2L [nucleotides (518–529)–1542]. The values in parentheses refer to the residues covered by the deoxyoligomers. Since the actual site of RNase H cleavage is variable within the hybridized region, we list the entire range of possible cut sites in this manner. A schematic diagram showing the fragments, the location of the methylated bases, and the site of colicin E3 cleavage is presented in Fig. 1. The purity of the three fragments was checked by gel electrophoresis. Only a single band of the expected size was found (data not shown). Nucleosides from enzymatic hydrolysates of each fragment were then analyzed by LC/MS (Fig. 2). Ψ was detected only in fragment H2S, indicating that Ψ is only located between nucleotides (297–308) and (518–529). In particular, analysis of fragment (518–529)–1542 (Fig. 2C) did not reveal any Ψ. The identification of modified residues other than those expected to be found in specific fragments implies that some secondary site hybridization of the deoxyoligonucleotides had occurred. For example, m^2G (known to be at positions 966, 1207, and 1516) and m^6_2A (known to be at positions 1518 and 1519) were identified in the hydrolysate of the H1S fragment [nucleotides 1–(297–308)]. The level of such contamination is estimated to be 10–20%. Location of the Ψ residue(s) between positions 297 and 529 plus restriction to the sequence GΨG identifed three candidate sites, namely GΨG357–359, GΨG404–406, and GΨG515–517.

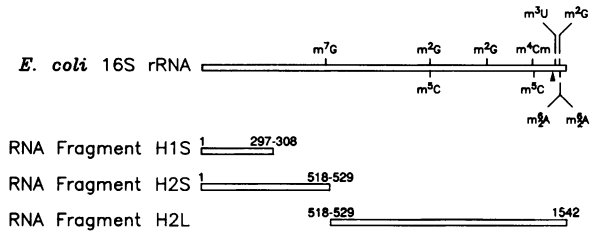

Figure 1. A linear schematic diagram of 16S RNA showing the location of the methylated bases. The site of colicin E3 cleavage is indicated by the solid triangle. The location and size of fragments H1S, H2S, and H2L are as shown.

Exact positioning of the pseudouridine residue(s)

For a precise determination of the location of the Ψ residue(s), a new sequencing method (16) was used which depends on the

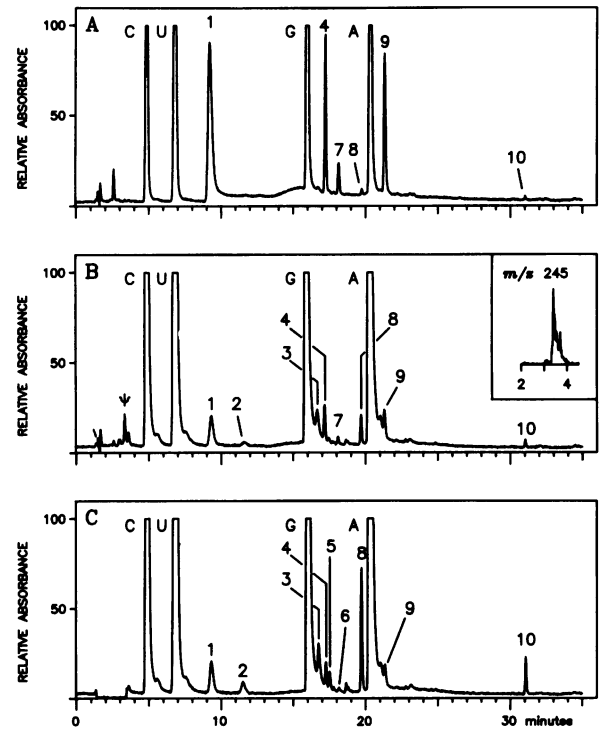

Figure 2. LC/MS analysis of nucleosides from an enzymatic digest of RNase H fragments of *E.coli* 16S rRNA. (**A**, **B**, **C**) Chromatograms corresponding to the analysis of H1S [residues 1–(297–308)], H2S [residues 1–(518–529)], and H2L [residues (518–529)–1542], respectively. Detection by UV absorbance at 254 nm. The major ribonucleosides and pseudouridine (Ψ) are lettered. Other peak identities are: 1, deoxycytidine (dC); 2, 5-methylcytidine (m^5C); 3, 7-methylguanosine (m^7G); 4, deoxyguanosine (dG); 5, 3-methyluridine (m^3U); 6, N^4,2′-*O*-dimethylcytidine (m^4Cm); 7, thymidine (dT); 8, N^2-methylguanosine (m^2G); 9, deoxyadenosine (dA); and 10, N^6,N^6-dimethyladenosine (m^6_2A). Shoulders on the right sides of C, U, G, A peaks (panels B and C) are characteristic chromatographic artifacts resulting from injection of large quantities of the four major nucleosides. Unlabeled peaks were shown by their mass spectra not to be nucleosides.

selective chemical modification of Ψ with CMC, resulting in a block to reverse transcription (Fig. 3B). In addition, the well-known resistance of Ψ to hydrazinolysis was adapted (16) in a similar reverse transcription analysis (Fig. 3A). Analysis of Fig. 3B shows the presence of a strong doublet band at positions 516 and 517 in isolated 16S RNA (N) after CMC treatment (+) but nothing in the *in vitro* synthesized 16S RNA (S), or in the absence of CMC (−). This doublet band corresponds to modification of U516. No other bands were detected with these properties. The strong band at 527−528 seen in N with or without CMC treatment corresponds to m^7G527, which after the alkali treatment causes a strong stop due to opening of the imidazole ring (27). Hydrazine−aniline treatment (Fig. 3A) confirmed the presence of Ψ only at position 516. All of the U residues yielded doublet bands when treated with hydrazine and aniline in the S lanes, and all but one doublet was also present in the N lanes. The missing doublet corresponded to U516. These results unambiguously establish U516 as a Ψ residue. No other Ψ residues were detected in this fragment. A complete scan from 297 to residue 529 did not reveal a second Ψ (data not shown). Since there was no Ψ in fragments 1−(297−308) or (519−529)−1542 by direct analysis (Fig. 2), these results establish that there is only one Ψ residue in *E.coli* 16S RNA and that it is located at position 516.

DISCUSSION

Despite both early (13, 14) and more recent (22, 23) measurements that showed 1−2 residues of Ψ in *E.coli* 16S RNA, Ψ has not been generally listed as a minor nucleotide component of 16S RNA, possibly because a 5% by weight contamination of the rRNA used with tRNA could easily account for the Ψ residues. In this work, we have shown that there is Ψ in 16S RNA, that there is only one residue, and that it is located at position 516. This location is intriguing because it is positioned at the 5′-end of the most highly conserved loop sequence in all small subunit RNAs (Fig. 4). Not only is this sequence conserved but there is ample evidence for the role of this loop and connecting stem region in important functional aspects of the ribosome. As annotated in the legend to Fig. 4, the anticodons of both P- and A-site bound tRNAs protect residues in this loop, a variety of experiments show that this stem−loop structure (residues 500−547) has a prominent role in determining the fidelity of codon recognition (Sm^r, $Supp^{ochre}$, Ram, interaction with S12) and parts of the loop are either physically close to the decoding site (XL), or at least allosterically connected with it (tRNA protection). In addition, mutation of G529 or G530 generates a dominant lethal phenotype *in vivo*, and *in vitro*, G530 mutation affects codon recognition at the A-site (31, 41).

There are also interesting structural aspects of this stem−loop. The stem of the 530 loop consists of two helices, 500−504/541−545 and 511−515/536−540, which have the potential to stack upon one another. Moreover, as noted in Fig. 4, the conserved nucleotides at the apex of the 530 loop (524−526) are involved in a tertiary interaction with nucleotides in a bulge loop (505−507) located in the stem, forming a pseudoknot structure (25, 37) which is required for ribosome function (34). In addition, the pairing of residues 521−522/527−528 creates a two base pair helix which has the potential to form a coaxial stack on the pseudoknot structure (25). The presence of four short helices in this stem−loop allows for

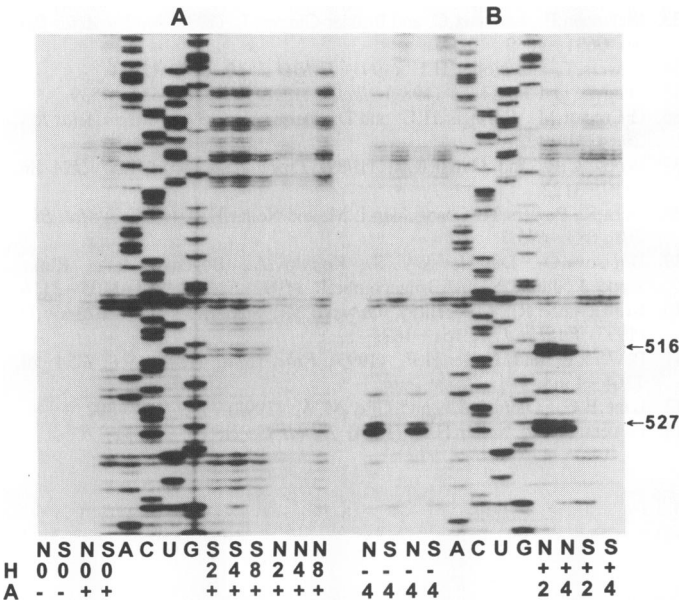

Figure 3. Reverse transcription analysis of the site of Ψ formation. (**A**) Hydrazine (H) treatment was for 0−8 min as indicated followed by treatment in the absence (−) or presence (+) of aniline (A). (**B**) Treatment without (−) or with (+) CMC followed by 2 or 4 h incubation at pH 10.4. The methodology was according to (16). N, isolated, modified 16S RNA; S, unmodified *in vitro* transcript of 16S RNA; A, C, U, G, sequencing lanes using the *in vitro* transcript.

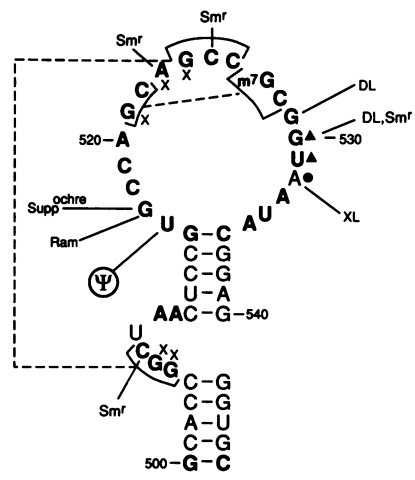

Figure 4. Structural and functional properties of the stem−loop structure containing Ψ516. Bold letters: highly evolutionarily conserved (25). ▲, ●: protection from chemical modification by either intact tRNA or the anticodon stem−loop when bound to the A- (▲) or P-site (●) (28). DL: mutations at position 529 (29) or 530 (30, 31) produce a dominant lethal phenotype *in vivo*. Sm^r: mutations at positions 507 (34), 523 (32, 33), 525 (34), or 530 (31) result in streptomycin-resistant ribosomes. $Supp^{ochre}$: mutation at the corresponding site in yeast mitochondrial small subunit rRNA results in suppression of ochre mutations (35). Ram: base changes result in increased readthrough of, and frameshifting at, stop codons (36). - - -: base changes within the connected boxes in rRNAs from different organisms maintain the potential for base pair formation (25, 37) and are required for function (34). X: sites interacting with protein S12 (38). XL, site of crosslinking to mRNA near its decoding position (39, 40).

a variety of possible coaxial stacking arrangements, although not all can form simultaneously (25). These complex and unusual higher order structural features of the 530 stem−loop suggest that this region may act as a molecular switch to accomodate dynamic changes in ribosome structure as a consequence of translational processes. Indeed, the 530 loop has been proposed to act as a molecular reciprocal toggle switch to control the affinity of the ribosome for cognate versus near-cognate ternary complexes of amino acid−tRNA, EFTu, and GTP, and thus to control the fidelity of codon recognition (43). In view of the striking localization of modifications to other functional sites of the bacterial ribosome (40), it is reasonable to postulate that Ψ may be involved in the fine tuning of elements of the structure and function of the 530 stem−loop.

Ψ residues have several characteristics not shared by U as a result of the relocation of the glycosidic bond from N-1 to C-5. Firstly, the N-1 proton can now act as a hydrogen bond donor and thus stabilize secondary or tertiary structure either intra- or intermolecularly. That is, the H bond feature could serve as a cross-strut to hold the local 16S RNA structure in place, or it could interact with the 23S RNA or a ribosomal protein as part of the subunit association matrix. A second possible function for the N-1 proton is its facile acylation capability (12). This capacity was noted previously and used as the basis for a hypothetical functional role for Ψ at the peptidyl transfer center (42). Thirdly, as pointed out by Maden (11), in either the *syn* or *anti* base pairing configuration the atoms lining the major groove side of Ψ differ from those of U, creating the potential for differential recognition by proteins or other ligands.

Do all organisms have Ψ at the same place in the conserved stem−loop structure? Unfortunately, the available data is limited to eukaryotes, which possess considerably more Ψ in their small subunit RNAs. For example, *Saccharomyces carlsbergensis* has approximately 14, *Xenopous laevis* about 44, while mammals have 36−38 Ψ/mol RNA (11). Exact placement of 16 Ψ residues and approximate (within 1−4 residues) placement of 14 others in mammals did not reveal any Ψ at the site equivalent to 516. However, as there are approximately 8 residues not yet located and it is possible that the equivalent of Ψ516 may yet be found.

Is there a specific enzyme for Ψ516 biosynthesis? Preliminary evidence indicates that there is such an activity in the soluble fraction of the *E.coli* cell which possesses a high degree of specificity (J.Wrzesinski, B.Lane and J.Ofengand, unpublished results). The purification and properties of this enzyme are currently under study with the aim of comparing and contrasting the *in vitro* assembly and function of 30S ribosomes with and without Ψ516.

REFERENCES

1. Noller,H.F. (1991) *Ann. Rev. Biochem.* **60**, 191−227.
2. Noller,H.F. (1993) *FASEB J.* **7**, 87−89.
3. Cunningham,P.R., Nurse,K., Weitzmann,C.J. and Ofengand,J. (1993) *Biochemistry* **32**, 7172−7180.
4. Limbach,P.A., Crain,P.F. and McCloskey,J.A. (1994) *Nucleic Acids Res.* **22**, 2183−2196.
5. Saponara,A.G. and Enger,M.D. (1974) *Biochim. Biophys. Acta* **349**, 61−77.
6. Maden,B.E.H., Forbes,J., de Jonge,P. and Klootwijk,J. (1975) *FEBS Lett.* **59**, 60−63.
7. Johnson,J.D. and Horowitz,J. (1971) *Biochim. Biophys. Acta* **247**, 262−279.
8. Kowalak,J.A. and McCloskey,J.A. (1993) In Nierhaus,K.H., Franceschi,F., Subramanian,A.R., Erdmann,V.A., and Wittman-Liebold,B. (eds), *The Translational Apparatus: Structure, Function, Regulation, Evolution*. Plenum Press, New York, pp. 79−88.
9. Thomas,G., Gordon,J. and Rogg,H. (1978) *J. Biol. Chem.* **253**, 1101−1105.
10. Bruenger,E., Kowalak,J.A., Kuchino,Y., McCloskey,J.A., Stetter,K.O. and Crain,P.F. (1993) *FASEB J.* **7**, 196−200.
11. Maden,B.E.H. (1990) *Progr. Nucl. Acids Res. Mol. Biol.* **39**, 241−303.
12. Spector,L.B. and Keller,E.B. (1958) *J. Biol. Chem.* **232**, 185−192.
13. Dubin,D.T. and Günalp,A. (1967) *Biochim. Biophys. Acta* **134**, 106−123.
14. Nichols,J.L. and Lane, B.G. (1967) *J. Mol. Biol.* **30**, 477−489.
15. Kowalak,J.A., Pomerantz,S.C., Crain,P.F. and McCloskey,J.A. (1993) *Nucleic Acids Res.* **21**, 4577−4585.
16. Bakin,A. and Ofengand,J. (1993) *Biochemistry* **32**, 9754−9762.
17. Cunningham,P.R., Weitzmann,C.J., Nurse,K., Masurel,R., van Knippenberg,P.H., and Ofengand,J. (1990) *Biochim. Biophys. Acta* **1050**, 18−26.
18. Helser,T., Davies,J.E. and Dahlberg,J.E. (1972) *Nature New Biol.* **235**, 6−9.
19. Crain,P.F. (1990) *Methods Enzymol.* **193**, 782−790.
20. Edmonds,C.G., Vestal,M.L. and McCloskey,J.A. (1985) *Nucleic Acids Res.* **13**, 8197−8206.
21. Pomerantz,S.C. and McCloskey,J.A. (1990) *Methods Enzymol.* **193**, 796−824.
22. Krzyzosiak,W., Denman,R., Nurse,K., Hellmann,W., Boublik,M., Gehrke,C.W., Agris,P.F. and Ofengand,J. (1987) *Biochemistry* **26**, 2353−2364.
23. Gehrke,C.W. and Kuo,K.C. (1989) *J. Chromatogr.* **471**, 3−36.
24. Bowman,C.M., Dahlberg,J.E., Ikemura,T., Konisky,J. and Nomura,M. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 964−968.
25. Gutell,R.R. (1993) In Nierhaus,K.H., Franceschi,F., Subramanian,A.R., Erdmann,V.A., and Wittman-Liebold,B., (eds.), *The Translational Apparatus*. Plenum Press, New York NY, pp 477−488.
26. Greenberg,M.E. and Bender,P.T. (1993) In Ausubel,F.A., Brend,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (eds.), *Current Protocols in Molecular Biology*. Greene Publishing and Willey Interscience, New York, pp. 4.0.1−4.10.9.
27. Hall,R.H. (1971) In *The Modified Nucleosides in Nucleic Acids*. Columbia University Press, New York, pp. 146−147.
28. Moazed,D. and Noller,H.F. (1990) *J. Mol. Biol.* **211**, 135−145.
29. Santer,U., Cekleniak,J. and Santer,M. (1993) *Mol. Biol. Cell* **4**, 420a.
30. Powers,T. and Noller,H.F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1042−1046.
31. Santer,M., Santer,U., Nurse,K., Bakin,A., Cunningham,P., Zain,M., O'Connell,D. and Ofengand,J. (1993) *Biochemistry* **32**, 5539−5547.
32. Gauthier,A., Turmel,M. and Lemieux,C. (1988) *Mol. Gen. Genet.* **214**, 192−197.
33. Melançon,P., Lemieux,C. and Brakier-Gingras,L. (1988) *Nucleic Acids Res.* 16, 9631−9639.
34. Powers,T. and Noller,H.F. (1991) *EMBO J.* **10**, 2203−2214.
35. Shen,Z. and Fox,T.D. (1989) *Nucleic Acids Res.* **17**, 4535−4539.
36. O'Connor,M., Göringer,H.U. and Dahlberg,A.E. (1992) *Nucleic Acids Res.* **20**, 4221−4227.
37. Woese,C.R. and Gutell,R.R. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3119−3122.
38. Stern,S., Powers,T., Changchien,L.M. and Noller,H.F. (1988) *J. Mol. Biol.* **201**, 683−695.
39. Dontsova,O., Dokudovskaya,S., Kopylov,A., Bogdanov,A.A., Rinke-Appel,J., Junke,N. and Brimacombe,R. (1992) *EMBO J.* **11**, 3105−3116.
40. Brimacombe,R., Mitchell,P., Osswald,M., Stade,K. and Bochkariov,D. (1993) *FASEB J.* **7**, 161−167.
41. Powers,T. and Noller,H.F. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 1364−1368.
42. Lane,B.G., Ofengand,J. and Gray,M.W. (1992) *FEBS Lett.* **302**, 1−4.
43. Powers,T. and Noller,H.F. (1994) *Trends Genet.* **10**, 27−31.