Minireview

# Ribosomal RNA pseudouridines and pseudouridine synthases

James Ofengand*

*Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, P.O. Box 016129, Miami, FL 33101, USA*

Received 20 November 2001; accepted 27 November 2001

First published online 25 January 2002

Edited by Lev Kisselev

**Abstract Pseudouridines are found in virtually all ribosomal RNAs but their function is unknown. There are four to eight times more pseudouridines in eukaryotes than in eubacteria. Mapping 19 *Haloarcula marismortui* pseudouridines on the three-dimensional 50S subunit does not show clustering. In bacteria, specific enzymes choose the site of pseudouridine formation. In eukaryotes, and probably also in archaea, selection and modification is done by a guide RNA–protein complex. No unique specific role for ribosomal pseudouridines has been identified. We propose that pseudouridine's function is as a molecular glue to stabilize required RNA conformations that would otherwise be too flexible. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.**

*Key words:* Function; Guide RNA; snoRNA; Archaeon; Eukaryon; Eubacterium

## 1. Introduction

Pseudouridine (5-ribosyluracil; Ψ) is the most common modified nucleoside in ribosomal RNA (rRNA). Although the sum of all 2′-O-methyl nucleosides are present in approximately the same amount as Ψ in eukaryotes [1], they are distributed among the four canonical nucleosides, and so are individually considerably less in amount than Ψ. Base-methylated and other modified nucleosides are present in even lower amount. Ψ is made after the polynucleotide chain is synthesized [2,3] by first choosing a particular uridine residue and then isomerizing it to Ψ. Neither cofactors nor energy are required, except in eukaryotes (see below) where the synthase is found in complex with at least three other proteins and a guide RNA. In one other synthase, zinc is an essential cofactor [3]. Both recognition and isomerization are complex processes and are only partly understood. In eubacteria, the selection process is carried out by a set of protein enzymes, Ψ synthases, which have the ability both to recognize specific uridines in an RNA chain and to catalyze the isomerization to Ψ. Although it is known that a set of such synthases exist with the ability to select one or another U residue out of up to 1200 uridines in a large subunit (LSU) RNA (see below), the details of the recognition process are unknown, and may vary, perhaps widely, for different synthases. The mechanism of Ψ formation is known to require cleavage of the *N*-glycosyl link, rotation of the uracil base while still enzyme-bound [4], and re-formation of a *C*-glycosyl bond. It is not known how this is accomplished, although a proposal has been made involving the participation of the γ-carboxyl of an active site aspartate in a particular sequence context [4], and some supporting evidence has been obtained [5,6].

In eukaryotes, it is now known that the uridine selection process is performed by guide RNAs which base-pair with rRNA nucleotides surrounding the designated U, whereas the isomerization step is carried out by a protein enzyme, probably the same one for all Ψ [7–9]. Thus in eukaryotes the recognition function is not only on a separate molecule from the catalytic function but is even in a different molecular class. No pseudouridylation guide RNAs for rRNA Ψ have been found so far in archaea but genomic analysis for putative Ψ synthases suggests that archaeal organisms, at least those whose genomes have been sequenced, use a guide RNA system also (see below).

In this minireview, we discuss the number and location of Ψ residues in rRNA, the Ψ synthases that put them there, and what happens when specific Ψ and Ψ synthases are removed.

## 2. Number and location

The number of Ψ residues in rRNA from representative eubacteria, archaea, eukaryotes, and organelles is listed in Table 1. Most of the data were obtained by sequencing the RNA for Ψ although in a few instances only the total amount was determined. One striking feature for which there is no ready explanation is the four- to eight-fold increase in the relative amount of Ψ in eukaryotes compared to eubacteria whether measured as the percent of U residues or as the percent of total nucleotides. It appears that whatever it is that Ψ does is more in demand in eukaryotes. A second feature is the large variation in Ψ among the archaea. *Halobacter halobium*, *Sulfolobus solfataricus*, and *Sulfolobus acidocaldarius* have a small number like the eubacteria, while *Haloarcula marismortui* has many more, like the eukaryotes. Moreover, it is possible that the final number of Ψ in *H. marismortui* may be even higher. The number shown is the number of strong positive sites. In addition, we found around 40 additional weaker sites. These could have arisen from Us that were incompletely modified to Ψ, or from Ψs that were incompletely derivatized with CMC, the sequence detection agent. They could also be due to some artifact not previously encountered, and not be Ψ at all. However, should they in fact be Ψ sites, the total number would then equal or exceed that found in the higher eukaryotes.

*Fax: (1)-305-243 3955.
*E-mail address:* jofengan@molbio.med.miami.edu (J. Ofengand).

0014-5793/02/$22.00 © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
PII: S0014-5793(02)02305-0

The number of Ψ alone is not very informative without knowledge of where they are in the rRNA secondary and tertiary structure. In small subunit (SSU) RNA, there is only a single Ψ in *Escherichia coli* and *Bacillus subtilis*, whereas they are numerous in the eukaryotes *Saccharomyces cerevisiae*, *Mus musculus*, and *Homo sapiens*. However, mapping does not show any clustering around the known functional centers of the SSU, and when placed in the three-dimensional structure of the *Thermus thermophilus* 30S ribosome, they appear to be distributed throughout the particle [22]. The role of Ψ in the SSU is, therefore, unclear.

The positions of all of the sequenced Ψ residues in LSU RNAs listed in Table 1 are shown in Fig. 1 on the secondary structure backbone of *H. marismortui*. Ψ predominate in the 3′-half of the LSU RNA, and except for two Ψ in *E. coli*, none of the eubacteria-like RNAs have any Ψ in the 5′-half. In the 3′-half, there are many examples of common sites, and numerous examples in different organisms of Ψ occurrence at adjacent sites or across a helix, suggesting a need for Ψ in a given vicinity but not necessarily at an exact site. In some cases, there is no U in the sequence at the same site, but in other cases U is present but not modified [16]. The most highly conserved site is in the loop of helix 69, residues 1952, 1956, and 1958. Ψ1952 is highly but not completely conserved, Ψ1956 is conserved in all organisms examined except for *S. acidocaldarius*, and Ψ1958 is so far totally conserved in cytoplasmic ribosomes. Curiously, although *Zea mays* chloroplast ribosomes have all three Ψ, the *Trypanosoma brucei* kinetoplast ribosomes and the mitochondrial ribosomes from yeast, mouse, and man manage to do without them [16]. In the 70S tRNA structure proposed by Yusupov et al. [24], this stem-loop projects out from the 50S to make a bridge to the 30S at the decoding site, and supports both A and P site-bound tRNAs underneath their elbows. The eukaryote-like behavior of *H. marismortui* is also seen in its secondary structure distribution of Ψ. An unusual feature is the occurrence of Ψ in regions not previously shown to have Ψ, even in eukaryotes. For example, Ψ1784, Ψ1850, and 1860 are in stems not modified in other organisms, and those in the 200–400 and 1500–1600 regions are also new. These Ψ are not generally archaeal-specific since *H. halobium* does not have them although the sequence analysis would not have detected the 200–400 or 1500–1600 region Ψ. Likewise, not enough of the *S. acidocaldarius* sequence was examined to be able to make a statement about any of these new Ψ sites in that organism.

The more interesting question is where the Ψ are in the actual ribosome. So far, only the 50S structure of *H. marismortui* has been determined to sufficient resolution to attempt mapping of Ψ. All of the Ψ shown in Fig. 1 have been so mapped [22]. Here we show only the results for *H. marismortui* and *E. coli* on the *H. marismortui* structure (Fig. 2). As the

Table 1
Number of pseudouridine residues in small and large subunit ribosomal RNAs

| Organism | Number of Ψ residues | Ψ (% of U) | Ψ (% of all residues) | How determined[a] (% RNAReference sequenced) | |
|---|---|---|---|---|---|
| Small subunit cytoplasm | | | | | |
| *Escherichia coli* | 1 | 0.3 | 0.06 | seq. (100) | [10] |
| *Bacillus subtilis* | 1 | 0.3 | 0.06 | seq. (15) | [11] |
| *Halobacter halobium* | 0 | – | – | seq. (70) | [12] |
| *Haloferax volcanii* | 0 | – | – | total | [13] |
| *Sulfolobus solfataricus* | 4–5 | 1.8–2.2 | 0.28–0.35 | total | [14] |
| *Saccharomyces cerevisiae* | 14[b] | 2.7 | 0.78 | seq. (98) | [12] |
| *Xenopus laevis* | ~44 | 10.7 | 2.4 | total | [1] |
| *Mus musculus* | ~36 | 8.9 | 1.9 | seq. (80) | [1] |
| *Homo sapiens* | ~36 | 8.6 | 1.8 | seq. (80) | [1] |
| Large subunit cytoplasm | | | | | |
| *Escherichia coli* | 10[c] | 1.7 | 0.34 | seq. (99) | [15] |
| *Bacillus subtilis* | 5[c] | 0.9 | 0.17 | seq. (56) | [16] |
| *Haloarcula marismortui* | 19[d] | 3.5[d] | 0.65[d] | seq. (90) | unpub. |
| *Halobacter halobium* | 4 | 0.7 | 0.14 | seq. (54) | [16] |
| *Sulfolobus solfataricus* | 4 | 0.8 | 0.13 | total | [14] |
| *Sulfolobus acidocaldarius* | 6 | 1.2 | 0.20 | seq. (20) | [17] |
| *Saccharomyces cerevisiae* | 30 | 3.5 | 0.88 | seq. (99) | [18] |
| *Drosophila melanogaster* | 57 | 4.8 | 1.44 | seq. (59) | [16] |
| *Xenopus laevis* | ~52 | 7.9 | 1.27 | total | [1] |
| *Mus musculus* | 57 | 7.3 | 1.21 | seq. (46) | [16] |
| *Homo sapiens* | 55 | 7.3 | 1.09 | seq. (41) | [16] |
| Large subunit mitochondria | | | | | |
| *Saccharomyces cerevisiae* | 1 | 0.1 | 0.03 | seq. (99) | [18] |
| *Mus musculus* | 1 | 0.2 | 0.06 | seq. (17) | [16] |
| *Homo sapiens* | 1 | 0.3 | 0.06 | seq. (41) | [16] |
| *Trypanosoma brucei* | 6 | 0.8 | 0.35 | seq. (98) | [16] |
| Large subunit chloroplasts | | | | | |
| *Zea mays* | 4[c] | 0.7 | 0.14 | seq. (29) | [16] |
| 5.8S RNA | | | | | |
| *Saccharomyces cerevisiae* | 1 | 2.2 | 0.63 | seq. (100) | [19] |
| *Mus musculus* | 2 | 5.7 | 1.3 | seq. (100) | [20,21] |
| *Homo sapiens* | 2 | 5.5 | 1.3 | seq. (100) | [20,21] |

[a]seq., determination by sequencing the RNA; total, determination by total nucleoside analysis; unpub., unpublished results of S. Jean-Charles and J. Ofengand.
[b]This value includes the hypermodified pseudouridine, $m^1acp^3\Psi1189$.
[c]Includes the $m^3\Psi1915$ in *E. coli* and the equivalent residues in *B. subtilis* and *Z. mays*.

*H. marismortui* data set is eukaryote-like, that of *E. coli* was also mapped as the archetype eubacterium. It was chosen because it has sufficient Ψ to show a trend, whereas there are too few in the other eubacteria. The clustering of the 10 *E. coli* Ψ mapped on the *H. marismortui* LSU is obvious, as are the more diffuse locations of the 19 *H. marismortui* Ψ. In this case also, the *H. marismortui* Ψ locations resemble the distribution found for other eukaryotes [22]. Whereas the location of the *E. coli* Ψ suggests a role in A site stabilization and/or function, the scatter of those in *H. marismortui* suggests a more general role in structural stabilization rather than one specific to ligand binding. It is interesting in this regard that only the two Ψ in the helix 69 loop are coincident in both organisms, and none of the other *H. marismortui* Ψ come within 6–10 Å of the *E. coli* positions.

## 3. Pseudouridine biosynthesis in eubacteria

A set of enzymes specific for single U or subsets of U carry out the reaction in eubacteria. The most well-studied organism is *E. coli*. Table 2 lists all 10 of the known Ψ synthases in this organism. These genes were identified initially by amino acid sequence homology [27,40–42] using RsuA, RluA, TruA, and TruB, identified biochemically, for reference. Coincidentally, these four synthase also define a set of four families according to sequence homology, as indicated in Table 2. Subsequently, as indicated in Table 2, all were identified as bona fide synthases. Three make Ψ only in tRNA and will not be discussed further here. One makes the single Ψ in SSU RNA, and six make the 10 Ψ in LSU RNA. Since single gene deletions cause specific Ψ to disappear and all the Ψ sites can be thus accounted for, no Ψ is made by more than one synthase in vivo. The reverse is not the case. RluC and RluD make three Ψ each. RluD may recognize its Ψ by searching for all U residues in or near the loop of the helix 69 stem-loop as the only unmodified U in the stem-loop is a G-U base pair near the base of helix 69 and away from the loop. RluC recognition is more complex. The three Ψ it makes, 955, 2504, and 2580, are neither in a common sequence context nor in a secondary or tertiary structure that shows common features or is close in three-dimensional space. The recognition mechanism(s) used by RluC will be particularly interesting to decipher. RluA is noteworthy as the first example of a dual-specificity modifying enzyme. It is specific for U746 in LSU RNA and also specific for U32 in the four tRNAs which have Ψ at that position. This recognition mechanism is somewhat better understood since modification appears to favor an unfolded structure (modification is favored with RNA fragments and in the absence of $Mg^{2+}$) and there is a common sequence context in both LSU RNA and tRNA at the site of modification [29]. RluB and RluF are also an interesting pair with regard to recognition as they select adjacent U residues for modification with virtually no cross-reactivity [28].

Table 2 also shows that an essential aspartic acid exists in all of the Ψ synthases. This aspartate is found in a related common sequence context in all four *E. coli* families [22]. As noted in Section 1, Santi and colleagues proposed that this aspartate is involved in the catalytic mechanism [4]. While the data in Table 2 do not prove that the aspartate in question directly participates in Ψ formation, it is certainly clear that this particular aspartate is essential. Similar results were obtained for two yeast synthases which make Ψ in tRNA and Ψ in mitochondrial rRNA [43,44] as well as for the yeast synthase likely responsible for cytoplasmic rRNA Ψ synthesis [45].

All eubacteria appear to have sets of Ψ synthases which group into the same four families (Table 3). Uniformly, there are either one or no members of the TruA and TruB families with the vast majority of organisms having one of each. It can be assumed, therefore, that in these bacteria, most of the tRNAs have both a Ψ38–40 and a Ψ55. For the purposes of this review, however, we focus on the rRNA Ψ synthases in the RsuA and RluA families. There are as few as two and as many as 10 divided between the two families in various ways. As no complete Ψ sequence analysis has been done for any organism except *E. coli*, it is not possible to say to what extent genome searching by sequence homology has uncovered all of the synthase genes, nor to assign a particular synthase to a particular Ψ. It is notable, however, that no sequenced bacterial genome has yet been found which lacks all Ψ synthases of the RsuA and/or RluA type. From this fact, it appears that Ψ in rRNA is an ancient invention.

## 4. Pseudouridine biosynthesis in eukaryotes

rRNA Ψ formation in eukaryotes takes place in the nucleolus on the earliest detectable rRNA precursor, possibly while transcription is still in progress [1]. A guide RNA, in complex with at least four proteins, base-pairs with a segment of rRNA surrounding the U to be isomerized to Ψ [7–9]. One of the proteins, Cbf5p in yeast (Nap57p in mammals), has a strong homology to TruB and mainly for that reason is considered to be the synthase. The other proteins are Gar1p, Nhp2p, and Nop10p. All four are essential for pseudouridylation but the functions of the latter three are unknown [9]. The structural features of the guide RNA–rRNA complex are shown in Fig. 3. Normally, two of these guide RNA elements are joined together in tandem to form a snoRNA and can therefore sometimes specify two distinct Ψ sites. Recently, however, a guide snoRNA has been described in *Leptomonas collosoma* which consists of only a single such element [46]. Since there are 90–100 Ψ in rRNA of higher eukaryotes, there should be 50–100 guide RNA complexes milling about in the nucleolus waiting for their turn to make Ψ as the rRNA is transcribed. The logistic and spatial requirements for this would appear to be daunting. It is not clear whether all of the Ψ are made this way or whether a subset may be made by distinct synthases, as in bacteria. Except for the ultra-streamlined *Guillardia theta*, there are other synthase genes available in the sequenced genomes of eukaryotes to allow for this possibility (Table 3).

Guide RNAs have been experimentally shown to function in the way just described in both yeast and human cells. More extensive studies have been done in yeast. So far 15 yeast snoRNAs have been experimentally shown, by deletion, to make specific Ψ, and four more have been putatively identified with specific known Ψ sites, based on their sequence [7]. Twenty Ψ sites, out of the 44 total in yeast (Table 1), remain to be connected to a snoRNA. As yet, no guide RNAs for the conserved Ψ in the helix 69 loop have been identified in yeast. The effect of deletion of such a snoRNA on yeast cell growth would be most interesting in view of the effects found in *E. coli* (see below).

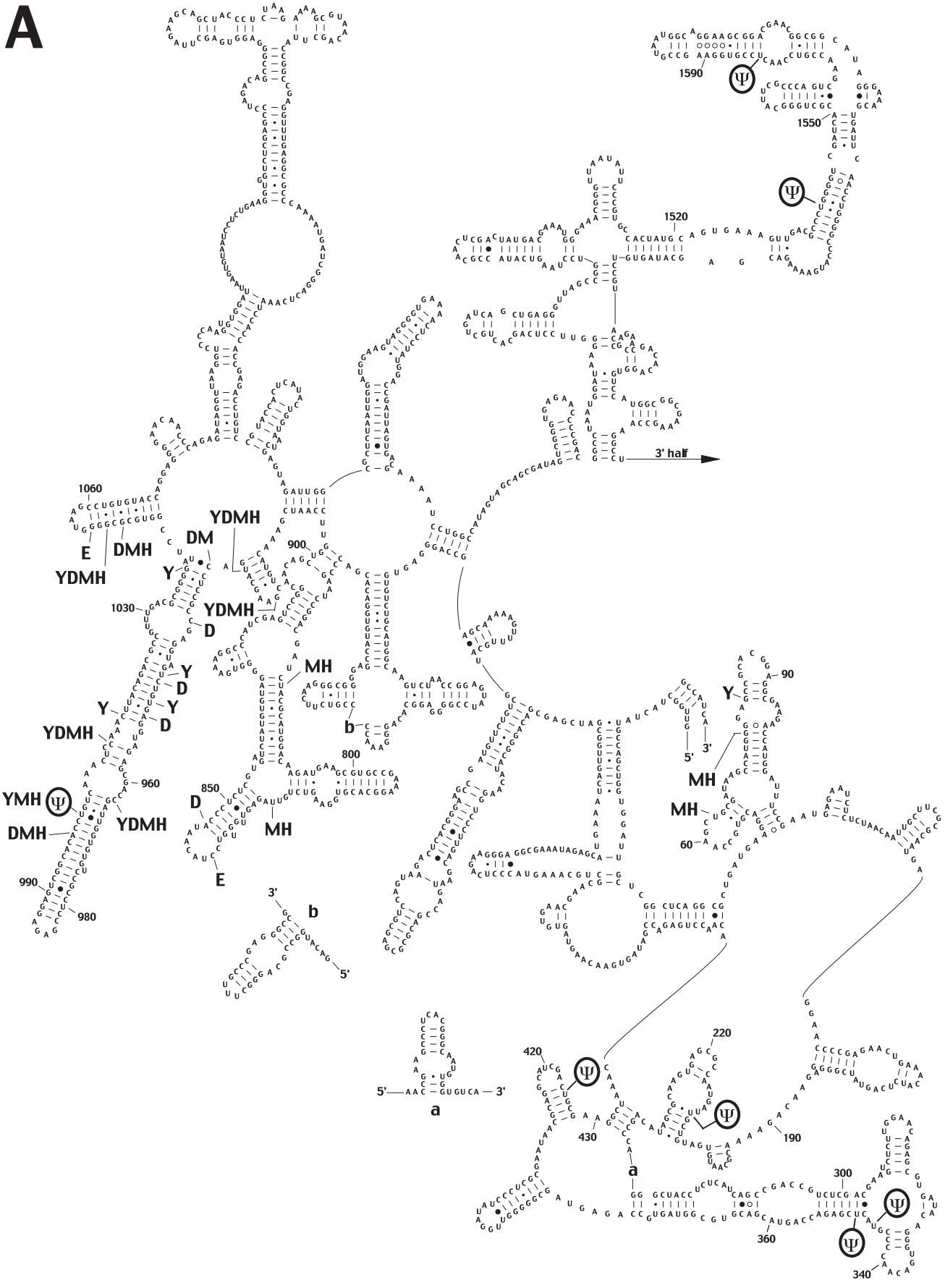
A
3' half
YDMH
DM
DMH
E
Y
D
MH
b
a
Ψ
1060
1030
1590
1550
1520
900
800
850
960
990
980
60
90
190
220
300
340
360
420
430
5'
3'

Fig. 1.

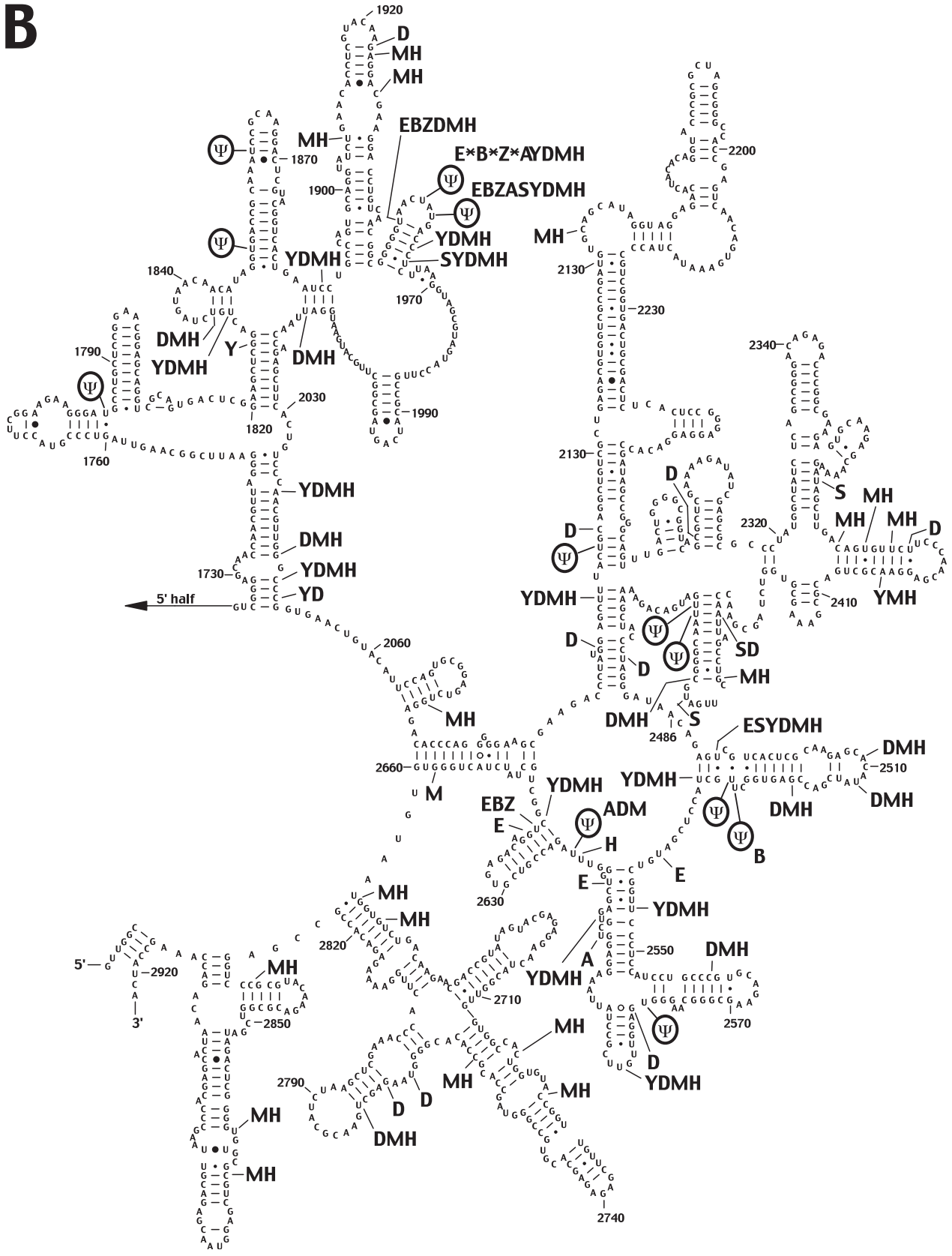


Fig. 1. Ψ locations in large subunit RNAs. A: 5′-half. B: 3′-half. The secondary structure of *H. marismortui* rrnB 23S RNA (AF034620) from the Gutell web site (www.rna.icmb.utexas.edu) is shown using *H. marismortui* numbering. The sites for Ψ in *H. marismortui* (Ψ), *E. coli* (E), *B. subtilis* (B), *Z. mays* (Z), *H. halobium* (A), *S. acidocaldarius* (S), *S. cerevisiae* (Y), *D. melanogaster* (D), *M. musculus* (M), and *H. sapiens* (H) were obtained from the sources indicated in Table 1. E* is $m^3\Psi$ [23] and B* and Z* are probably the same [16]. Seven Ψ in *D. melanogaster* are not shown because they occur in a large expansion segment absent in *H. marismortui*.

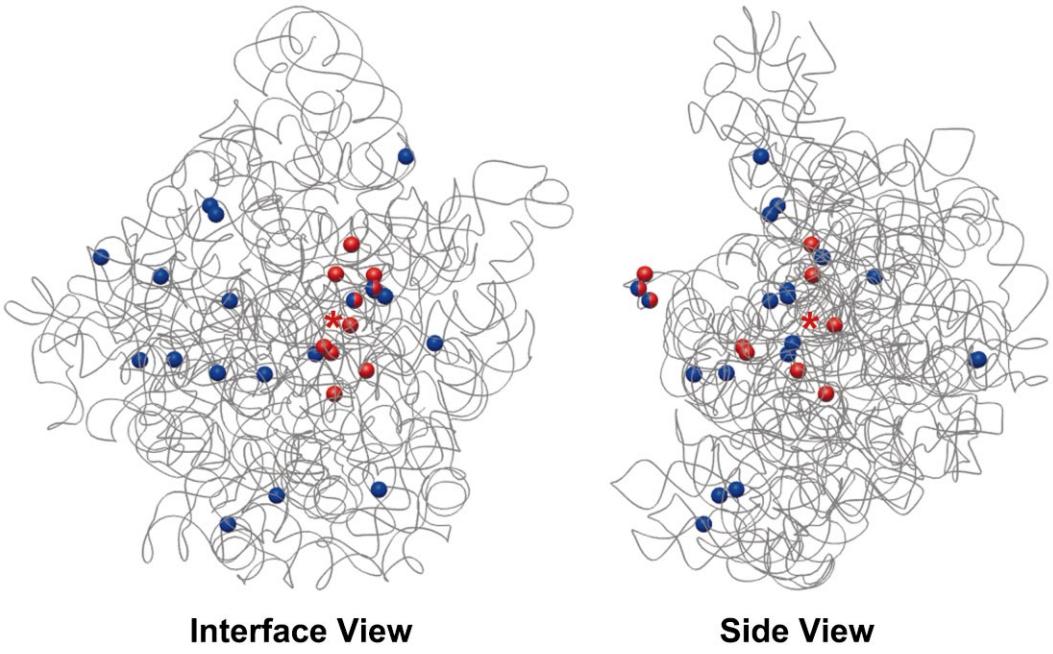


Fig. 2. Three-dimensional structure of the RNA in the LSU showing the location of the Ψ residues in *H. marismortui* and *E. coli*. The *H. marismortui* 50S subunit [25] with helix 69 [24] added is shown with all proteins removed and the phosphodiester backbone represented by a continuous line. The Ψ locations were marked by a 6 Å diameter sphere whose center was located at the glycosyl N of the base in question except for bases 1952, 1956, 1958, and 1962 derived from helix 69 since only the phosphate coordinates were available. The interface view and an approximately 90° side view are shown. *H. marismortui* Ψ are shown as blue spheres, *E. coli* as red spheres, and the two sites which coincide as half-blue, half-red spheres. The peptidyl transferase center is marked by a red star at A2486.

## 5. Pseudouridine biosynthesis in archaea

It is not known experimentally whether archaea make their Ψ like bacteria or like eukaryotes. However, analysis of the 12 complete genome sequences available (Table 3) reveals a paucity of recognizable synthase genes. The three Crenarchaeota have but one gene which is homologous to the yeast Cbf5 gene, the one that works with guide RNAs. Since *S. solfataricus* has eight or nine rRNA Ψ (Table 1), we suppose that a guide RNA system is used. The nine Euryarchaeota also have a Cbf5-like synthase gene and since *H. halobium* with four Ψ is closely related to *Halobacterium* sp. NRC-1 [47], these organisms likely use guide RNAs also. The 19 Ψ found in *H.*

Table 2
Properties of *E. coli* pseudouridine synthases[a]

| Name | Swiss-Prot accession number | RNA substrate | Ψ site | Essential aspartate and mutants | Aspartate mutant activity (% of WT) | | References |
|---|---|---|---|---|---|---|---|
| | | | | | In vivo | In vitro | |
| *RsuA family* | | | | | | | |
| RsuA | P33918 | 16S RNA | 516 | D102T/N | < 1 | n.t.[b] | [26,27] |
| RluB | P37765 | 23S RNA | 2605 | D110T/N | < 1 | n.t. | [28] |
| RluE | P75966 | 23S RNA | 2457 | D69T/N | < 1 | n.t. | [28] |
| RluF | P32684 | 23S RNA | 2604 | D107T/N | < 1 | n.t. | [28] |
| *RluA family* | | | | | | | |
| RluA | P39219 | 23S RNA<br>tRNA | 746<br>32 | D64T/N/A/C | < 1 | < 2 | [29–31] |
| RluC | P23851 | 23S RNA | 955<br>2504<br>2580 | D144T/N | < 1 | n.t. | [32,33], unpub.[d] |
| RluD | P33643 | 23S RNA | 1911<br>1915<br>1917 | D139T/N | n.t. | < 5 | [33–35] |
| TruC | Q46918 | tRNA | 65 | D54T | < 5 | n.t. | [28] |
| *TruB family* | | | | | | | |
| TruB | P09171 | tRNA | 55 | D48C/A | < 1 | < 0.1 | [31,36,37] |
| *TruA family* | | | | | | | |
| TruA | P07649 | tRNA | 38–40 | D60X[c] | n.t. | < 0.01 | [4,38,39] |

[a]Modified from [28].
[b]n.t., not tested.
[c]X = A, E, K, N, S.
[d]unpub., S. Jean-Charles and J. Ofengand, unpublished results.

Table 3
Family distribution of pseudouridine synthases in completely sequenced genomes

| Organism | Family | | | |
|---|---|---|---|---|
| | RsuA | RluA | TruA | TruB |
| Bacteria | | | | |
| *Agrobacterium tumefaciens* | 1 | 3 | 1 | 1 |
| *Aquifex aeolicus* | 2 | 1 | 1 | 1 |
| *Bacillus halodurans* | 2 | 2 | 1 | 1 |
| *Bacillus subtilis* | 2 | 3 | 1 | 1 |
| *Borrelia burgdorferi* | 1 | 2 | 1 | 1 |
| *Buchnera sp.* | 1 | 2 | 1 | 1 |
| *Campylobacter jejuni* | 2 | 3 | 1 | 1 |
| *Caulobacter crescentus* | 3 | 3 | 1 | 1 |
| *Chlamydia muridarum* | 1 | 2 | 1 | 1 |
| *Chlamydia pneumoniae* | 1 | 1 | 1 | 1 |
| *Chlamydia trachomatis* | 1 | 2 | 1 | 1 |
| *Chlamydophila pneumoniae* AR39 | 1 | 2 | 1 | 1 |
| *Clostridium acetobutylicum* | 3 | 3 | 1 | 1 |
| *Corynebacterium glutamicum* | 2 | 4 | 1 | 1 |
| *Deinococcus radiodurans* | 1 | 3 | 1 | 1 |
| *Enterococcus faecalis* | 3 | 3 | 1 | 1 |
| *Escherichia coli* | 4 | 4 | 1 | 1 |
| *Haemophilus influenzae* | 3 | 5 | 1 | 1 |
| *Helicobacter pylori* | 1 | 3 | 1 | 0 |
| *Lactococcus lactis* | 3 | 3 | 1 | 1 |
| *Listeria innocua* | 2 | 3 | 1 | 1 |
| *Listeria monocytogenes* EGD-e | 2 | 3 | 1 | 1 |
| *Mesorhizobium loti* | 2 | 2 | 1 | 1 |
| *Mycobacterium leprae* | 1 | 1 | 1 | 1 |
| *Mycobacterium tuberculosis* | 1 | 2 | 1 | 1 |
| *Mycoplasma genitalium*[+] | 0 | 2 | 1 | 0 |
| *Mycoplasma pneumoniae*[+] | 0 | 2 | 1 | 0 |
| *Mycoplasma pulmonis* | 1 | 2 | 0 | 1 |
| *Neisseria meningitidis* MC58 | 3 | 4 | 1 | 1 |
| *Pasteurella multocida* PM70 | 3 | 5 | 1 | 1 |
| *Porphyromonas gingivalis* | 1 | 3 | 1 | 1 |
| *Pseudomonas aeruginosa* | 4 | 4 | 1 | 1 |
| *Rickettsia conorii* | 1 | 2 | 1 | 1 |
| *Rickettsia prowazekii* | 1 | 2 | 1 | 1 |
| *Sinorhizobium meliloti* | 2 | 2 | 1 | 1 |
| *Staphylococcus aureus* N315/Mu50 | 2 | 3 | 1 | 1 |
| *Streptococcus pneumoniae* | 3 | 3 | 1 | 1 |
| *Streptococcus pyogenes* | 3 | 3 | 1 | 1 |
| *Synechocystis* sp. | 2 | 2 | 1 | 1 |
| *Thermotoga maritima* | 1 | 2 | 1 | 1 |
| *Treponema pallidum* | 1 | 3 | 1 | 1 |
| *Ureaplasma urealyticum* | 0 | 2 | 1 | 1 |
| *Vibrio cholerae* | 4 | 6 | 1 | 1 |
| *Xylella fastidiosa* | 2 | 2 | 1 | 1 |
| *Yersinia pestis* | 3 | 4 | 1 | 1 |
| Crenarchaeota | | | | |
| *Aeropyrum pernix* | 0 | 0 | 0 | 1* |
| *Sulfolobus solfataricus* P2 | 0 | 0 | 0 | 1* |
| *Sulfolobus tokodaii* | 0 | 0 | 0 | 1* |
| Euryarchaeota | | | | |
| *Archaeoglobus fulgidus* | 0 | 0 | 1 | 1* |
| *Halobacterium* sp. NRC-1 | 0 | 0 | 1 | 1* |
| *Methanococcus jannaschii* | 0 | 0 | 1 | 1* |
| *Methanothermobacter thermoautotrophicus* | 0 | 0 | 1 | 1* |
| *Pyrococcus abyssii* | 0 | 0 | 1 | 1* |
| *Pyrococcus furiosus* | 0 | 0 | 1 | 1* |
| *Pyrococcus horikoshii* | 0 | 0 | 1 | 1* |
| *Thermoplasma acidophilum* | 0 | 0 | 1 | 1* |
| *Thermoplasma volcanium* | 0 | 0 | 1 | 1* |
| Eukaryotes | | | | |
| *Arabidopsis thaliana* | 1 | 6 | 6 | 1, 1* |
| *Caenorhabditis elegans* | 0 | 1 | 2 | 0, 1* |
| *Drosophila melanogaster* | 0 | 2 | 3 | 1, 1* |
| *Guillardia theta* | 0 | 0 | 0 | 0, 1* |
| *Saccharomyces cerevisiae* | 0 | 4 | 3 | 1, 1* |

Genome sequences were obtained from public databases or as described previously and analyzed for predicted Ψ synthase genes [42]. Superscript * in the TruB column indicates the number of ORFs with homology to yeast Cbf5 and which are presumed to function with guide RNAs. [+]Other *Mycoplasma* strains, *capricolum* and *mycoides*, have Ψ55, the TruB product.

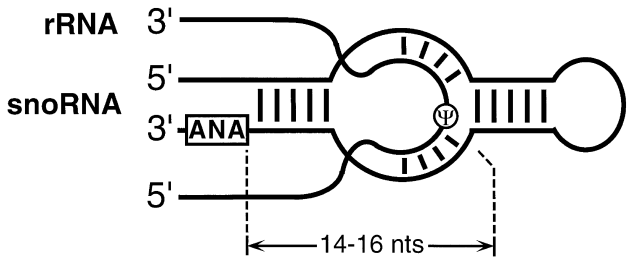


Fig. 3. Base-pairing of a guide snoRNA with complementary rRNA sequences which flank the uridine to be modified. The lengths of the guide sequences vary, but the target uridine sits in an unpaired pocket that is quite constant in size. Pairing on the 5′ side of the uridine uses 4–10 base pairs and that on the 3′ side 3–10. The distance between the target uridine and the helix on the 5′ side of the rRNA is mostly 0 but occasionally 1 residue and that on the 3′ side is mostly 1 and rarely 2. As shown, the distance to the H box is a near constant 14–16 nt. Reprinted from [7] with permission.

*marismortui* also imply a guide RNA system since the number of synthases required would otherwise be unwieldy. When that genome is sequenced, it will be possible to examine this question. The existence of homologs to Nhp2p, Gar1p, and Nop10p in archaeal organisms further supports the concept of guide RNAs in archaea [48,49].

It is also worthy of note that while the Euryarchaeota all have a TruA-like synthase gene, it is absent in the Crenarchaeota, and neither class has a TruB gene. The absence of TruA and TruB genes has interesting implications for the presence of Ψ in the tRNAs from archaea, some of which are known to contain Ψ (see discussion in [37]).

## 6. Function of pseudouridines in rRNA

Based on the fact that Nature has seen fit to maintain two distinct systems for Ψ synthesis, and on the observation that Ψ formation in eukaryotes is restricted to the functionally important core region of LSU and SSU RNAs [1], it is a reasonable expectation that Ψ contributes to the proper functioning of the mature ribosome. Nevertheless, it has been very difficult to show experimentally that this is so.

In *E. coli*, the approach taken has been to delete the synthase genes which then result in the absence of the Ψ dependent on those synthases. In yeast, the analogous approach involves deleting a particular snoRNA guide. In both systems, growth has been used as a preliminary screen for an effect. Deletion of six of the seven rRNA Ψ synthases in *E. coli* (Table 2) had no effect on growth over a range of temperatures [27,28,30,32]. Only the loss of RluD resulted in a major growth rate defect [34,35]. However, in this case, the growth defect could be fully complemented by a plasmid-borne RluD gene mutated at its essential aspartate (Table 2) so as to block Ψ formation in the cell [35]. Thus, while the RluD synthase protein was essential for normal cell growth, its enzyme products, Ψ1911, Ψ1915, and Ψ1917, were not. Subsequent analysis showed that the growth defect was directly correlated with a failure to correctly assemble 50S ribosomes [22], suggesting that RluD is directly involved in 50S subunit assembly, possibly as an RNA chaperone. The failure to show a need for these Ψ is surprising in view of their close association with A and P site-bound tRNAs (see above) but perhaps could have been anticipated by the ease with which second site revertants able to grow normally without these Ψ could be obtained [34].

With regard to Ψ in the peptidyl transferase center (PTC) vicinity, a quadruple mutant lacking the genes for RluB, RluC, RluE, and RluF has been constructed. This mutant lacks six Ψ, five of which are in and around the PTC. Although the cells grow, preliminary results indicate that they have some metabolic defects (M. Del Campo and J. Ofengand, unpublished results). It may be, therefore, that in most cases Ψ will not have unique effects, but rather will each contribute partially to some cooperative effect, for example in stabilizing a particular RNA conformation.

A similar situation exists in yeast. Single or even multiple deletions of a series of snoRNAs had no effect on growth, and snoRNA deletions which individually removed all six of the Ψ around the yeast PTC had no effect [7]. However, when all six of the Ψ are removed at the same time, the cells display a reduced growth rate and protein synthetic rate, a hypersensitivity to antibiotics that act on the LSU, and subtle disturbances in RNA structure in the LSU (T.H. King and M.J. Fournier, personal communication).

## 7. Coda

Although Ψ are found in virtually all ribosomes, their function is still unknown. In fact, it may be simplistic to search for a unitary function for all Ψ. Perhaps Nature, having learned how to make Ψ without any energy expenditure, has taken advantage of its properties for a variety of purposes. Given the fact that Ψ is only found in RNA molecules whose tertiary structure is important to their function, and that the very act of making Ψ generates a potent H-bond donor, it is tempting to suggest that the main function of Ψ is to act as a molecular glue, tightening up and/or reinforcing necessary RNA conformations that of themselves would be insufficiently rigid. This might explain (a) the large variation in number of Ψ among different species, (b) the failure to find closer correspondence in the sites for Ψ among different species, (c) the difficulty in detecting a functional effect upon single Ψ removal, and (d) the apparent better success in measuring a functional effect upon multiple Ψ deletions.

*Acknowledgements:* I thank Mark Del Campo for creating Figures 1 and 2, Serge Jean-Charles for unpublished information, and Prof. Kenneth Rudd for performing the genome analysis of *A. thaliana* and *D. melanogaster* in Table 3. Supported in part by NIH Grant GM58879.

## References

[1] Maden, B.E.H. (1990) Progr. Nucleic Acid Res. Mol. Biol. 39, 241–300.
[2] Johnson, L. and Söll, D. (1970) Proc. Natl. Acad. Sci. USA 67, 943–950.
[3] Ciampi, M.S., Arena, F., Cortese, R. and Daniel, V. (1977) FEBS Lett. 77, 75–82.
[4] Huang, L., Pookanjanatavip, M., Gu, X. and Santi, D.V. (1998) Biochemistry 37, 344–351.
[5] Gu, X., Liu, Y. and Santi, D.V. (1999) Proc. Natl. Acad. Sci. USA 96, 14270–14275.
[6] Foster, P.G., Huang, L., Santi, D.V. and Stroud, R.M. (2000) Nature Struct. Biol. 7, 23–27.
[7] Ofengand, J. and Fournier, M. (1998) in: Modification and Editing of RNA (Grosjean, H. and Benne, R., Eds.), pp. 229–253, ASM Press, Washington, DC.

[8] Bachellerie, J.-P., Cavaille, J. and Qu, L.H. (2000) in: The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions (Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B. and Noller, H.F., Eds.), pp. 191–203, ASM Press, Washington, DC.
[9] Kiss, T. (2001) EMBO J. 20, 3617–3622.
[10] Bakin, A., Kowalak, J.A., McCloskey, J.A. and Ofengand, J. (1994) Nucleic Acids Res. 22, 3681–3684.
[11] Niu, L. and Ofengand, J. (1999) Biochemistry 38, 629–635.
[12] Bakin, A. and Ofengand, J. (1995) Nucleic Acids Res. 23, 3290–3294.
[13] Kowalak, J.A., Bruenger, E., Crane, P. and McCloskey, J.A. (2000) J. Biol. Chem. 275, 24484–24489.
[14] Noon, K.R., Bruenger, E. and McCloskey, J.A. (1998) J. Bacteriol. 180, 2883–2888.
[15] Bakin, A. and Ofengand, J. (1993) Biochemistry 32, 9754–9762.
[16] Ofengand, J. and Bakin, A. (1997) J. Mol. Biol. 266, 246–268.
[17] Massenet, S., Ansmant, I., Motorin, Y. and Branlant, C. (1999) FEBS Lett. 462, 94–100.
[18] Bakin, A., Lane, B.G. and Ofengand, J. (1994) Biochemistry 33, 13475–13483.
[19] Rubin, G.M. (1973) J. Biol. Chem. 248, 3860–3875.
[20] Nazar, R.N., Sitz, T.O. and Busch, H. (1976) Biochemistry 15, 505–508.
[21] Khan, M.S.N. and Maden, B.E.H. (1977) Nucleic Acids Res. 4, 2495–2505.
[22] Ofengand, J., Malhotra, A., Remme, J., Gutgsell, N., Del Campo, M., Jean-Charles, S., Peil, L. and Kaya, Y. (2001) Cold Spring Harbor Symp. Quant. Biol. 66, in press.
[23] Kowalak, J.A., Bruenger, E., Hashizume, T., Peltier, J.M., Ofengand, J. and McCloskey, J.A. (1996) Nucleic Acids Res. 24, 688–693.
[24] Yusupov, M.M., Yusupova, G.Z., Baucom, A., Lieberman, K., Earnest, T.N., Cate, J.H. and Noller, H.F. (2001) Science 292, 883–896.
[25] Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) Science 289, 905–920.
[26] Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. and Ofengand, J. (1995) Biochemistry 34, 8904–8913.
[27] Conrad, J., Niu, L., Rudd, K., Lane, B.G. and Ofengand, J. (1999) RNA 5, 751–763.
[28] Del Campo, M., Kaya, Y. and Ofengand, J. (2001) RNA 7, in press.
[29] Wrzesinski, J., Nurse, K., Bakin, A., Lane, B.G. and Ofengand, J. (1995) RNA 1, 437–448.
[30] Raychaudhuri, S., Niu, L., Conrad, J., Lane, B.G. and Ofengand, J. (1999) J. Biol. Chem. 274, 18880–18886.
[31] Ramamurthy, V., Swann, S.L., Paulson, J.L., Spedaliere, C.J. and Mueller, E.G. (1999) J. Biol. Chem. 274, 22225–22230.
[32] Conrad, J., Sun, D., Englund, N. and Ofengand, J. (1998) J. Biol. Chem. 273, 18562–18566.
[33] Huang, L., Ku, J., Pookanjanatavip, M., Gu, X., Wang, D., Greene, P.J. and Santi, D.V. (1998) Biochemistry 37, 15951–15957.
[34] Raychaudhuri, S., Conrad, J., Hall, B.G. and Ofengand, J. (1998) RNA 4, 1407–1417.
[35] Gutgsell, N.S., Del Campo, M., Raychaudhuri, S. and Ofengand, J. (2001) RNA 7, 990–998.
[36] Nurse, K., Wrzesinski, J., Bakin, A., Lane, B.G. and Ofengand, J. (1995) RNA 1, 102–112.
[37] Gutgsell, N., Englund, N., Niu, L., Kaya, Y., Lane, B.G. and Ofengand, J. (2000) RNA 6, 1870–1881.
[38] Arps, P.J., Marvel, C.C., Rubin, B.C., Tolan, D.A., Penhoet, E.E. and Winkler, M.E. (1985) Nucleic Acids Res. 13, 5297–5315.
[39] Kammen, H.O., Marvel, C.C., Hardy, L. and Penhoet, E.E. (1988) J. Biol. Chem. 263, 2255–2263.
[40] Koonin, E.V. (1996) Nucleic Acids Res. 24, 2411–2415.
[41] Gustafsson, C., Reid, R., Greene, P.J. and Santi, D.V. (1996) Nucleic Acids Res. 24, 3756–3762.
[42] Ofengand, J. and Rudd, K. (2000) in: The Ribosome: Structure, Function, Antibiotics, and Cellular Interactions (Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B. and Noller, H.F., Eds.), pp. 175–189, ASM Press, Washington, DC.
[43] Ansmant, I., Motorin, Y., Massenet, S., Grosjean, H. and Branlant, C. (2001) J. Biol. Chem. 276, 34934–34940.
[44] Ansmant, I., Massenet, S., Grosjean, H., Motorin, Y. and Branlant, C. (2000) Nucleic Acids Res. 28, 1941–1946.
[45] Zebarjadian, Y., King, T., Fournier, M.J., Clarke, L. and Carbon, J. (1999) Mol. Cell. Biol. 19, 7461–7472.
[46] Liang, X., Liu, L. and Michaeli, S. (2001) J. Biol. Chem. 276, 40313–40316.
[47] Ng, W.A., Kennedy, S.P., Mahairas, G.G., Berquist, B. and Pan, M. et al. (2000) Proc. Natl. Acad. Sci. USA 97, 12176–12181.
[48] Henras, A., Henry, Y., Bousquet-Antonelli, C., Noaillac-Depeyre, J., Gélugne, J.-P. and Caizergues-Ferrer, M. (1998) EMBO J. 17, 7078–7090.
[49] Watanabe, Y. and Gray, M.W. (2000) Nucleic Acids Res. 28, 2342–2352.