Occurrence and location of 7-methylguanine residues in small-subunit ribosomal RNAs from eubacteria, archaebacteria and eukaryotes

Vera S. Zueva, Alexander S. Mankin, Alexei A. Bogdanov, David L. Thurlow* and Robert A Zimmermann*

Belozersky Laboratory of Molecular Biology and Bioorganic Chemistry, Moscow State University, Moscow 119899, USSR and **Department of Biochemistry, University of Massachusetts, Amherst, MA 01003, USA*

Received 25 June 1985

Specific cleavage with aniline provides a rapid and convenient method for establishing the presence and approximate location of 7-methyl-guanine (m^7G) residues in ribosomal RNA (rRNA) molecules. Using this approach, we have shown that a single m^7G occurs roughly 465 bases from the 5′-end of 16 S rRNA from the archaebacterium, *Thermoplasma acidophilum*, but that this modified base is absent from several other archaebacterial 16 S rRNAs. We have also demonstrated that a unique m^7G is found some 220–230 bases from the 3′-terminus of a number of eukaryotic 18 S rRNAs. In both cases, m^7G is present within evolutionarily conserved structural features, suggesting that this base may optimize the activity of functionally important regions of rRNAs in a kingdom-specific fashion.

7-Methylguanine *Ribosomal RNA* *Modified base* *Aniline cleavage* *Archaebacteria* *Protein synthesis*

1. INTRODUCTION

The modified base 7-methylguanine (m^7G) occurs rarely, but at very specific locations, in mature ribosomal RNA (rRNA) molecules [1]. In eubacterial 16 S rRNAs, for instance, a unique m^7G is invariably found within the 5′-proximal domain at a position corresponding to residue 527 in 16 S rRNA from *Escherichia coli* [2]. A single m^7G has also been identified in eukaryotic 18 S rRNAs, but in this case it has been located within a conserved oligonucleotide that occurs in the 3′-proximal portion of the molecule [3,4]. Although the function of the m^7G residues is not known, the evolutionary persistence of these modified sites suggests that they are important for the biological activity of the ribosome.

Specific cleavage of tRNA molecules at m^7G residues has been achieved by incubating them first with sodium borohydride and then with aniline [5]. Unexpectedly, treatment of *E. coli* 16 S rRNA in a similar fashion did not result in significant cleavage of the RNA [6]. It was shown, however, that addition of methylated carrier RNA during the reduction step was sufficient to promote quantitative cleavage at m^7G_{527} when the reduced 16 S rRNA was subsequently incubated with aniline [6].

In *E. coli* 16 S rRNA, m^7G occurs within the RNase-T_1 oligonucleotide CCm^7GCG. While this oligonucleotide is almost universally conserved in eubacterial 16 S rRNAs, the unmethylated sequence CCGCG occupies the same relative position in eukaryotic 18 S rRNAs [2]. The presence and distribution of m^7G in 16 S rRNAs from archaebacteria remains unclear, however, as CCm^7GCG has been found in only 2 of the many T_1 oligonucleotide catalogs available [2]. Accordingly, we have analyzed a number of archaebacterial 16 S rRNAs for m^7G by determining their sensitivity to aniline cleavage after reduction in the presence of methylated carrier RNA. Moreover, because the 3′-proxi-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

mal m^7G residue in eukaryotic 18 S rRNA has been located precisely in only one instance [4], we have used the aniline technique to establish whether this modified base occurs at an analogous position in other 18 S rRNAs.

2. MATERIALS AND METHODS

E. coli 16 S rRNA was isolated from 30 S subunits by phenol extraction and purified by sucrose-gradient centrifugation [7]. All other small-subunit rRNAs (S-rRNAs) used in this study were obtained from total phenol-extracted cellular RNA by fractionation on sucrose gradients [7]. Methylated carrier RNA was prepared by treating yeast soluble RNA (Calbiochem-Behring) with dimethyl sulfate (Merck or Eastman Kodak) as described by Peattie [8], with minor modifications to increase the extent of methylation [6]. Cleavage of S-rRNA at m^7G was accomplished by reduction with sodium borohydride in the presence of methylated carrier RNA followed by treatment with aniline as described by Zueva et al. [6]. Selected rRNA fragments were labeled with ^{32}P at their 5′- or 3′-termini and subjected to nucleotide sequence analysis by partial enzymatic or chemical degradation [8,9].

3. RESULTS

3.1. *m^7G in archaebacterial 16 S rRNAs*

We have used the susceptibility of S-rRNA to aniline cleavage in the presence of methylated carrier RNA to investigate the presence and location of m^7G in 16 S rRNAs from 4 species of archaebacteria that represent major phylogenetic

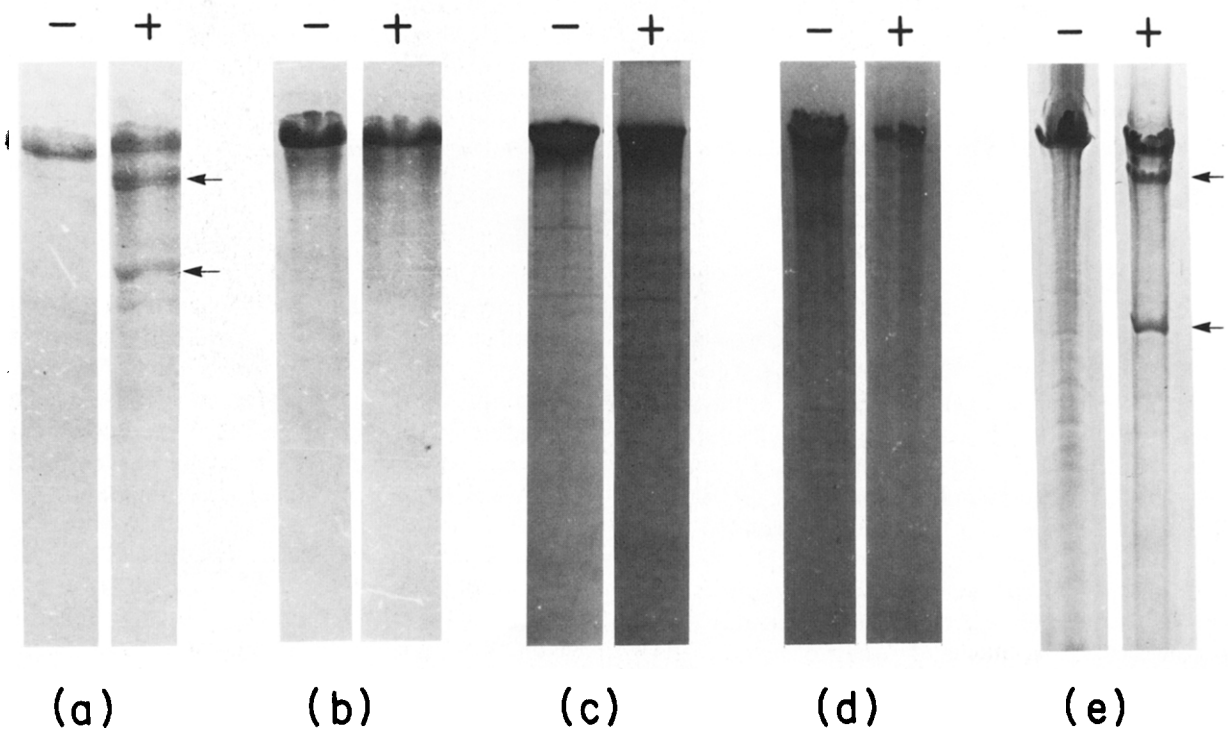

Fig.1. Susceptibility of archaebacterial 16 S rRNAs to cleavage by aniline. 16 S rRNAs from (a) *E. coli* (control), (b) *H. halobium*, (c) *M. thermoautotrophicum*, (d) *S. acidocaldarius* and (e) *T. acidophilum* were mixed with a 2–3-fold excess of methylated carrier RNA and reduced with sodium borohydride for 5 min at 20°C [6]. Each sample was divided into 2 equal portions and further incubated in the absence (−) or presence (+) of aniline for 30–120 min at 20°C. In all cases, the RNA was heated to 90°C for 30 s, loaded onto polyacrylamide gels in 50 mM Tris-50 mM boric acid, pH 8.3, 2.5 mM EDTA, 8 M urea, and electrophoresed for 3–4 h at constant current. The gels were stained with methylene blue or Stains-all.

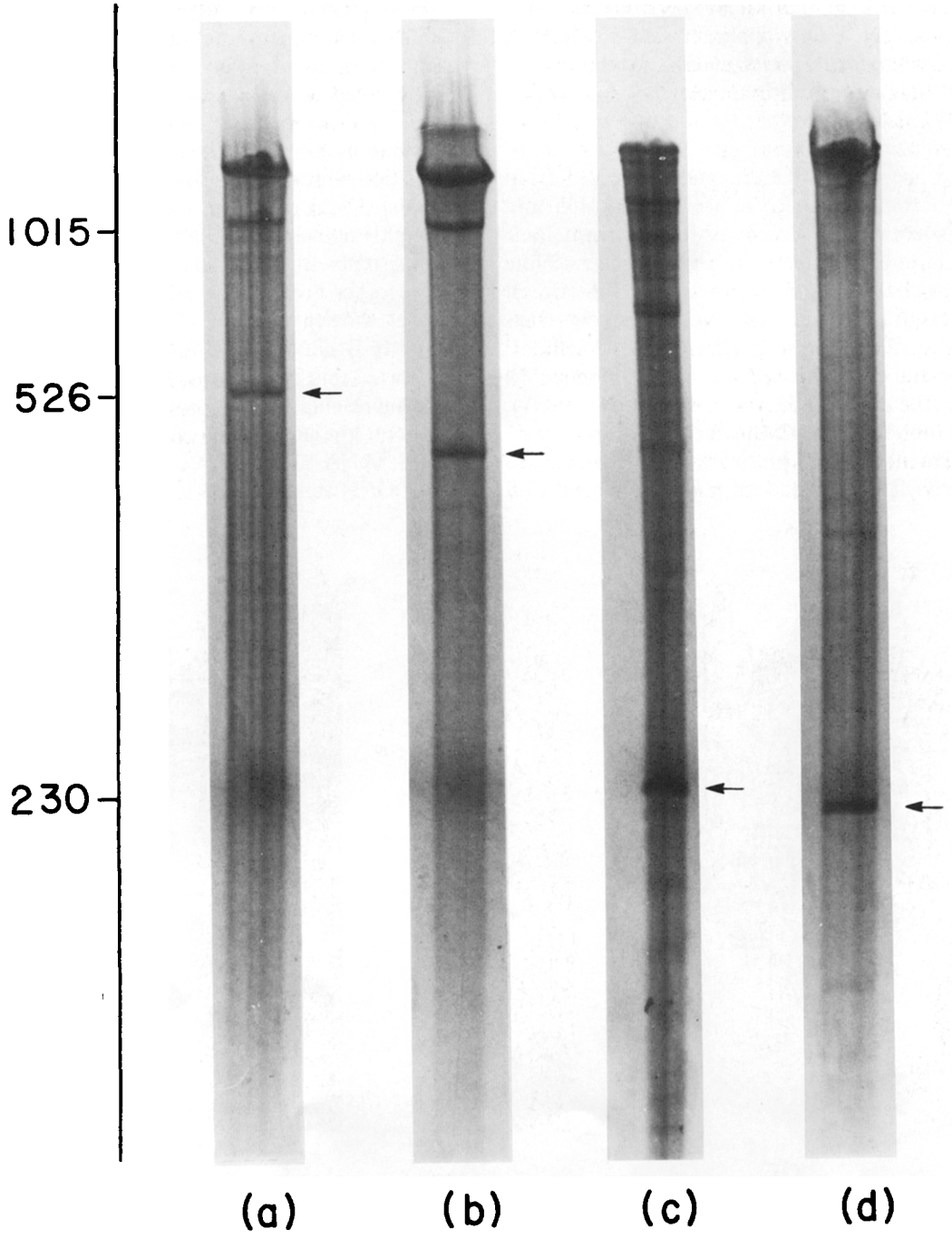

Fig.2. Aniline-induced cleavage of eubacterial, archaebacterial and eukaryotic S-rRNAs. (a) *E. coli* 16 S rRNA, (b) *T. acidophilum* 16 S rRNA, (c) *S. cerevisiae* 18 S rRNA and (d) *D. discoideum* 18 S rRNA were reduced, treated with aniline and subjected to polyacrylamide gel electrophoresis as indicated in fig.1. The scale on the left shows the lengths of the small and larger fragments from *E. coli* 16 S rRNA and of the small fragment from *S. cerevisiae* 18 S rRNA. The latter value was estimated in a separate experiment from a comparison with RNA size standards generated by digestion of *E. coli* 30 S ribosomal subunits with the double-strand specific RNase from *Naja oxiana* [20].

subdivisions within this kingdom. From the data presented in fig.1, it is apparent that 16 S rRNA from *Thermoplasma acidophilum* is cleaved by aniline under our conditions whereas 16 S rRNAs from *Methanobacterium thermoautotrophicum, Halobacterium halobium* and *Sulfolobus acidocaldarius* are not. The 2 fragments released from *Thermoplasma* 16 S rRNA are roughly 465 and 1000 nucleotides in length as judged from their electrophoretic mobilities relative to the aniline cleavage products from *E. coli* 16 S rRNA (fig.1).

To establish the identity of the small rRNA fragment from *Thermoplasma,* it was labeled with ^{32}P at the 5′-end and subjected to nucleotide sequence analysis by partial digestion with base-specific RNases (not shown). The 5′-terminal sequence of the fragment agreed precisely with that found previously at the 5′-end of *Thermoplasma* 16 S rRNA ([10]; K.-M. Cao, D.L. Thurlow and R.A. Zimmermann, unpublished). As noted above, the small fragment was estimated to be only 465 bases long, whereas the corresponding fragment from *E. coli* 16 S rRNA is 526 bases long. Thus, the m^7G residue in *Thermoplasma* 16 S rRNA can be in a position similar to that in *E. coli* 16 S rRNA only if the 5′-third of the former molecule contains roughly 60 bases less than the latter. This is in fact consistent with the observation that the 5′-domains of 4 other archaebacterial 16 S rRNAs are 60–65 bases shorter than the comparable portion of *E. coli* 16 S rRNA [11–14]. Moreover, it has been reported that the oligonucleotide CCG*CG, where G* represents an unknown modification of G, occurs in the RNase-T_1 catalog of *Thermoplasma* 16 S rRNA [10]. Our findings therefore suggest that G* is actually m^7G and indicate that this base

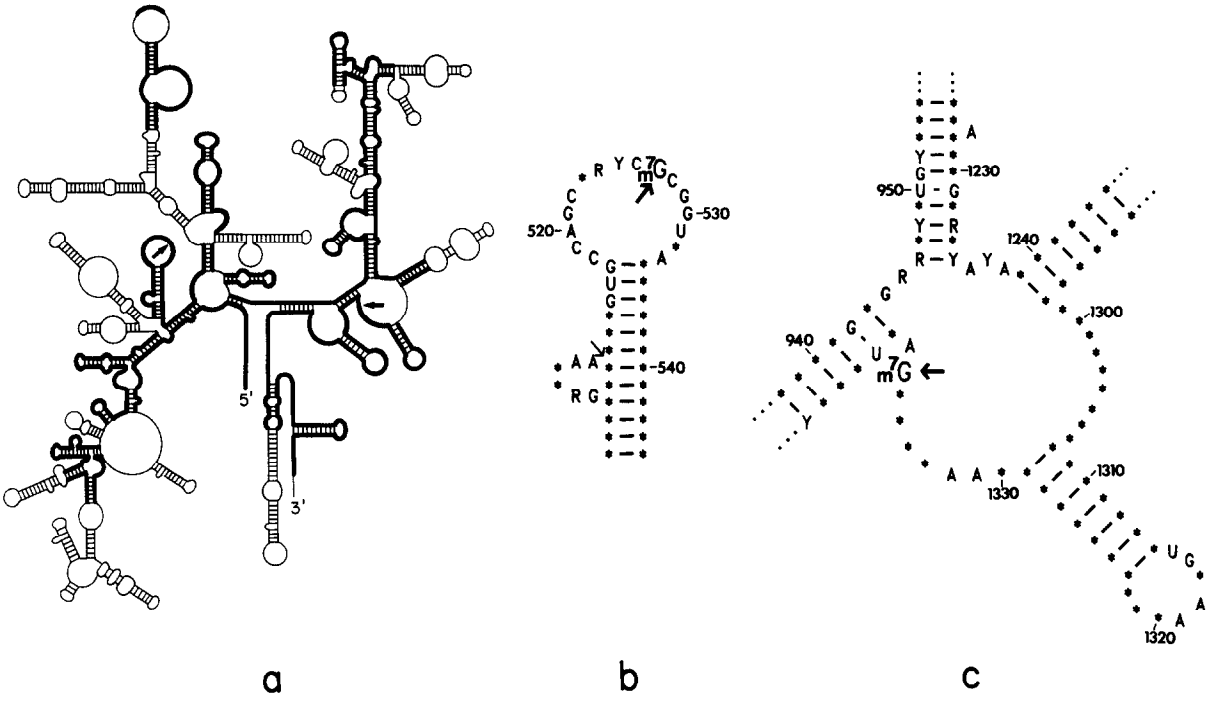

Fig.3. Location of m′G in the S-rRNA secondary structure. (a) Secondary-structure model of *E. coli* 16 S rRNA in outline form, adapted from [2]. The core structure, which reflects features common to all S-rRNAs, is indicated by thick lines. The positions of the 5′-proximal m^7G of eubacterial 16 S rRNA and of the 3′-proximal m^7G of eukaryotic 18 S rRNAs are marked by thick arrows. (b) Detail of the conserved stem-loop structure that contains the 5′-proximal m^7G in eubacterial 16 S rRNAs. An additional nucleotide is present in eukaryotic S-rRNAs that is not found in prokaryotic S-rRNAs; its position is indicated by the thin arrow. (c) Detail of the conserved secondary structure that surrounds the 3′-proximal m^7G in eukaryotic 18 S rRNAs. In (b) and (c), bases present in all S-rRNAs analyzed to date are designated by the appropriate symbol; Y, pyrimidine; R, purine.

occupies a position in the archaebacterial rRNA which corresponds to its position in *E. coli* 16 S rRNA. From an evolutionary perspective, the presence of m^7G at analogous sites in 16 S rRNAs from the archaebacterium, *Thermoplasma,* and the eubacterium, *E. coli,* is quite remarkable in the light of the extensive differences in overall primary structure between the 2 molecules. Moreover, the absence of m^7G from organisms representing 3 other branches of the archaebacterial lineage suggests that *Thermoplasma* may enjoy a unique evolutionary position in this group.

3.2. *m^7G in eukaryotic 18 S rRNAs*

It has been recognized for several years that certain eukaryotic S-rRNAs contain m^7G within the oligonucleotide sequence m^7GAAU [3]. Recently, the methylated G in 18 S rRNA from *Xenopus laevis* has been located at position 1596 by hybridization mapping and by sequence analysis of the 18 S rRNA gene [4,15]. Position 1596 occurs within the oligonucleotide segment $AAC^G_CAGGAAU$ which has been found some 220–230 bases from the 3′-end of nearly all eukaryotic 18 S rRNA molecules whose primary structures are known [4,16–19]. Owing to the evolutionary conservation of this segment, it is possible that the G residue analogous to base 1596 in *Xenopus* 18 S rRNA is methylated in many, if not all, eukaryotic 18 S rRNAs. Although the sequence GAAU also occurs at the corresponding site of most eubacterial and archaebacterial 16 S rRNAs, the relevant G residue is apparently never methylated [2].

Fig.2 provides direct experimental evidence that m^7G occurs close to one end of *Saccharomyces cerevisiae* 18 S rRNA. Following cleavage of the rRNA molecule with aniline, the length of the smaller rRNA fragment was determined to be 230 ± 5 bases by electrophoresis on a denaturing gel alongside RNA fragments of known size. In addition, the small yeast rRNA fragment was excised from the gel, labeled with ^{32}P at the 3′-end and sequenced by partial chemical degradation (not shown). The structure determined coincides exactly with the 3′-terminal sequence of yeast 18 S rRNA derived from the corresponding rDNA [16]. Since m^7G is generally found in the oligonucleotide m^7GAAU [3], and because the sequence GAAU occupies positions 1573–1576 in yeast 18 S rRNA [16], our data identify the location of m^7G as base 1573, 225 residues from the 3′-terminus.

Fragments 220–230 bases in length were also derived by aniline treatment of a number of other eukaryotic 18 S rRNAs, including those from loach, rat, rabbit and human cells (not shown). As shown in fig.2, the small rRNA fragment produced from the 18 S rRNA of *Dictyostelium discoideum* was slightly shorter than that from yeast and was demonstrated to arise from the 3′-end of the molecule by nucleotide sequence analysis (not shown). We suggest that the irregularity in the sequence of *Dictyostelium* 18 S rRNA which was reported to be an unidentified modification of A_{1744} [17] actually stems from an m^7G residue at position 1743, since we observe cleavage by aniline at approximately that site.

4. DISCUSSION

The highly conserved sequence GCCGCGGU, which encompasses residues 524–531 in *E. coli* 16 S rRNA, occurs at the same relative location in almost all S-rRNAs investigated to date [2]. More significantly, this sequence occupies precisely the same position within the core secondary structure of these molecules, forming part of the apex loop of a stem-loop feature close to the border between the 5′- and central domains (fig.3). Several bases within the GCCGCGGU segment are accessible to chemical reagents in the ribosome and, in *E. coli* 30 S subunits, the m^7G residue at position 527 is available for interaction with specific antibody [2,21]. The latter property has been exploited for immune electron microscopy which showed that m^7G maps between the 'head' and 'body' of the 30 S particle on the surface facing the 50 S subunit [21].

While its evolutionary persistence, exposed position and probable location on the subunit interface all indicate that GCCGCGGU is important for ribosome function, the pattern of post-transcriptional modification within this segment is even more persuasive. In all eubacterial, and at least one archaebacterial, S-rRNAs, the residue corresponding to G_{527} in *E. coli* is converted to m^7G whereas in all eukaryotic, and most archaebacterial, S-rRNAs, the G remains unmodified. However, in eukaryotic 18 S rRNAs, the nucleotide equivalent to U_{531}, only 4 bases away, is generally modified to 2′-*O*-methyluridine (2′*O*MeU) [2]. The presence

or absence of particular modifications may reflect the manner in which the ribosomes of each major evolutionary lineage have become adapted for optimal activity in different cellular environments. If GCCGCGGU occurs at a ligand binding site, for instance, the interaction may be most appropriately modulated by m^7G_{527} in eubacterial ribosomes, by $2'OMeU_{531}$ in eukaryotes and by the unmodified sequence in archaebacterial ribosomes.

In a similar vein, the consensus sequence GGAAU, situated at positions 1337–1342 in *E. coli* 16 S rRNA, is also subject to differential modification. This sequence occurs within a conserved feature of the core secondary structure, some 200–230 bases from the 3′-end of all S-rRNAs (fig.3). Here the residue corresponding to G_{1338} is unmodified in eubacterial and archaebacterial 16 S rRNAs, but is altered to m^7G in most, if not all, eukaryotic 18 S rRNAs. In this case, modification could either 'fine-tune' the 18 S rRNA for a function shared by all ribosomes or mediate an activity peculiar to eukaryotic ribosomes. Kingdom-specific post-transcriptional modifications in highly conserved regions of S-rRNA molecules may therefore be indicative of subtle changes in the molecular mechanism of translation that occurred within each of the major lines of descent at an early stage of evolution.

ACKNOWLEDGEMENTS

We are indebted to Drs V.K. Kagramanova and P.M. Rubtsov for helpful discussions. A.A.B. was a participant in the Fulbright-Hayes Exchange Program when the project was initiated. This work was supported in part by Research Grants PCM81-10885 and PCM84-10832 to R.A.Z. from the US National Science Foundation.

REFERENCES

[1] Fellner, P. and Sanger, F. (1968) Nature 219, 236–238.
[2] Woese, C.R., Gutell, R., Gupta, R. and Noller, H.F. (1983) Microbiol. Rev. 47, 621–669.
[3] Kahn, M.S.N., Salim, M. and Maden, B.E.H. (1978) Biochem. J. 169, 531–542.
[4] Salim, M. and Maden, B.E.H. (1981) Nature 291, 205–208.
[5] Wintermeyer, W. and Zachau, H.G. (1975) FEBS Lett. 58, 306–309.
[6] Zueva, V.S., Mankin, A.S., Bogdanov, A.A. and Baratova, L.A. (1985) Eur. J. Biochem. 146, 679–687.
[7] Zimmermann, R.A. (1979) Methods Enzymol. 59, 551–583.
[8] Peattie, D.A. (1979) Proc. Natl. Acad. Sci. USA 76, 1760–1764.
[9] Donis-Keller, H. (1979) Nucleic Acids Res. 7, 179–192.
[10] Woese, C.R., Maniloff, J. and Zablen, L.B. (1980) Proc. Natl. Acad. Sci. USA 77, 494–498.
[11] Gupta, R., Lanter, J.M. and Woese, C.R. (1983) Science 221, 656–659.
[12] Leffers, H. and Garrett, R.A. (1984) EMBO J. 3, 1613–1619.
[13] Hui, I. and Dennis, P.P. (1985) J. Biol. Chem. 260, 899–960.
[14] Kagramanova, V.K., Mankin, A.S., Teterina, N.L., Rubtsov, P.M., Kopylov, A.M., Baratova, L.A. and Bogdanov, A.A. (1985) Bioorg. Khim. 11, 516–522.
[15] Maden, B.E.H. (1980) Nature 288, 293–296.
[16] Rubtsov, P.M., Musakhanov, M.M., Zakharyev, V.M., Krayev, A.S., Skryabin, K.G. and Bayev, A.A. (1980) Nucleic Acids Res. 8, 5779–5794.
[17] McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R. and Sogin, M.L. (1983) Biochemistry 22, 5858–5868.
[18] Connaughton, J.F., Rairkar, A., Lockard, R.E. and Kumar, A. (1984) Nucleic Acids Res. 12, 4731–4745.
[19] Nelles, L., Fang, B.-L., Volckaert, G., Vandenberghe, A. and DeWachter, R. (1984) Nucleic Acids Res. 12, 8749–8768.
[20] Vassilenko, S.K., Carbon, P., Ebel, J.P. and Ehresmann, C. (1981) J. Mol. Biol. 152, 699–721.
[21] Trempe, M.R., Ohai, K. and Glitz, D.G. (1982) J. Biol. Chem. 257, 9822–9829.