RNA (2000), *6*:296–306. Cambridge University Press. Printed in the USA.

METHOD

Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry

FINN KIRPEKAR, STEPHEN DOUTHWAITE, and PETER ROEPSTORFF
Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark

ABSTRACT

We present a method to screen RNA for posttranscriptional modifications based on Matrix Assisted Laser Desorption/Ionization mass spectrometry (MALDI-MS). After the RNA is digested to completion with a nucleotide-specific RNase, the fragments are analyzed by mass spectrometry. A comparison of the observed mass data with the data predicted from the gene sequence identifies fragments harboring modified nucleotides. Fragments larger than dinucleotides were valuable for the identification of posttranscriptional modifications. A more refined mapping of RNA modifications can be obtained by using two RNases in parallel combined with further fragmentation by Post Source Decay (PSD). This approach allows fast and sensitive screening of a purified RNA for posttranscriptional modification, and has been applied on 5S rRNA from two thermophilic microorganisms, the bacterium *Bacillus stearothermophilus* and the archaeon *Sulfolobus acidocaldarius*, as well as the halophile archaea *Halobacterium halobium* and *Haloarcula marismortui*. One *S. acidocaldarius* posttranscriptional modification was identified and was further characterized by PSD as a methylation of cytidine$_{32}$. The modified C is located in a region that is clearly conserved with respect to both sequence and position in *B. stearothermophilus* and *H. halobium* and to some degree also in *H. marismortui*. However, no analogous modification was identified in the latter three organisms. We further find that the 5′ end of *H. halobium* 5S rRNA is dephosphorylated, in contrast to the other 5S rRNA species investigated. The method additionally gives an immediate indication of whether the expected RNA sequence is in agreement with the observed fragment masses. Discrepancies with two of the published 5S rRNA sequences were identified and are reported here.

Keywords: *Archaea*; *Bacillus*; *Halobacteriacae*; rRNA; sequence; *Sulfolobus*

INTRODUCTION

All types of stable RNA in the cell can be modified from the appearance of the primary transcript to the formation of the final functional entity. Ribosomal RNA (rRNA) in bacteria and archaea is normally transcribed as one long precursor molecule, which is then processed by a series of ribonucleases to generate the three functional rRNAs, 5S, 16S, and 23S rRNA, and additional tRNAs. All tRNAs are subjected to numerous nucleotide modifications (hereafter denoted posttranscriptional modifications) with more than 10% of the nucleotides being modified. In contrast, rRNA only contains relatively few such modifications. For example, there are 22 known posttranscriptionally modified nucleotides in the 23S rRNA, 11 in the 16S rRNA, and none in the 5S rRNA of *Escherichia coli* (Rozenski et al., 1999). The posttranscriptional modifications are important for the function of the ribosome. Reconstitution of active 50S ribosomal subunits with in vitro-transcribed 23S rRNA is not possible in *E. coli* whereas reconstitution can be performed efficiently using natural (modified) 23S rRNA (Green & Noller, 1996), implying that the posttranscriptional modifications are important in the assembly and/or activity of the ribosome. 50S ribosomal subunits from the thermophile bacteria *Thermus aquaticus* and *Bacillus stearothermophilus*, are only 20–30% as active when reconstituted with in vitro-transcribed 23S rRNA compared to reconstitution with the natural 23S rRNA (Khaitovich et al., 1999b, Green & Noller, 1999). Comparable results are obtained with *E. coli* in 30S ribosomal subunit reconstitution experiments with 16S rRNA (Cunningham et al., 1991). Furthermore, experiments would suggest that the natural 23S rRNA possesses peptidyl transferase activity (Noller et al., 1992), whereas it has not been possible to identify this enzyme activity in the in vitro-transcribed analog (Khaitovich et al., 1999a). Genetic examples of the importance of post-

Reprint requests to: Finn Kirpekar, Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark; e-mail: f.kir@PR-group.ou.dk.

transcriptional modifications include the requirement of a 23S rRNA-specific pseudouridine synthetase for *E. coli* growth (Raychaudhuri et al., 1998), and of a 21S rRNA-specific 2′-*O*-guanosine methylase for the assembly of yeast mitochondrial 54S ribosomal subunits (Sirum-Connolly & Manson, 1993).

Modifying rRNA posttranscriptionally is also a means by which microorganisms establish resistance towards antibiotics. Examples of 16S and 23S rRNA nucleotide methylations that confer antibiotic resistance can be found in a review by Cundliffe (1987). The most thoroughly investigated example is the N^6 mono- or dimethylation on adenine 2058 (*E. coli* numbering) of the 23S rRNA, which confers resistance against macrolides, lincomycin, and streptogramin B antibiotics (Skinner et al., 1983). The modification is catalyzed by structurally related, site-specific methyltransferases, which have been discovered in a wide range of microorganisms (Weisblum, 1995). A second example is thiostrepton resistance induced by methylation of the 23S rRNA adenine 1067 (Thompson et al., 1982). In both cases, the affinity of ribosomes for the antibiotics is greatly reduced as a result of the posttranscriptional modifications.

The contemporary method for localizing posttranscriptional modifications is based on electrospray ionization mass spectrometry (ESI-MS) of RNase digests (Kowalak et al., 1993; Felden et al., 1998). The RNA molecule in question is digested to completion with RNase T1 and the fragment masses are determined by ESI-MS. Divergences between the obtained mass spectrum and the one expected from the gene sequence suggest the presence of RNA fragments harboring posttranscriptional modifications. Characterization of the modified nucleotide(s) is performed by total hydrolysis into mono-nucleosides of the relevant RNA fragment followed by reversed phase HPLC and mass spectrometry. The retention time in HPLC combined with the mass data (including fragment masses) give strong indications of the identity of the nucleoside. This approach has been applied to characterize various RNAs with respect to posttranscriptional modifications. These include the central loop of domain V in 23 S rRNA and tmRNA from *E. coli* (Kowalak et al., 1995; Felden et al., 1998) and 5S rRNA in archaea (Bruenger et al., 1993).

The purpose of the present study was to develop a method to screen for posttranscriptional modifications based on MALDI time-of-flight mass spectrometry (Karas & Hillenkamp, 1988). In this mass spectrometric method, the analyte is cocrystallized with an excess of a suitable matrix, frequently an organic acid. The analyte/matrix mixture is desorbed/ionized by a short laser pulse; the generated ions are all accelerated to the same potential in an electrical field, and are subsequently separated in a field-free drift region according to their mass-over-charge ratio. The advantages of MALDI-MS are its high sensitivity, its tolerance against impurities, and its ability to handle complex mixtures, the latter two properties allowing analysis of the digests without prior separation or purification. Furthermore, MALDI-MS with an electrostatic ion reflector (ion mirror) and a second field-free drift region opens the possibility of mass spectrometric sequencing of putatively modified fragments by PSD (Spengler et al., 1992). The analysis requires prior specific digestion of the RNA. The concept behind the RNase digestion is similar to the one introduced by McCloskey and co-workers (Kowalak et al., 1993), except that we apply parallel digestion with two different RNases. The MALDI-MS/RNase digestion method was applied to 5S rRNA from four different microorganisms, identifying a single posttranscriptional modification, which was analyzed further by PSD.

RESULTS

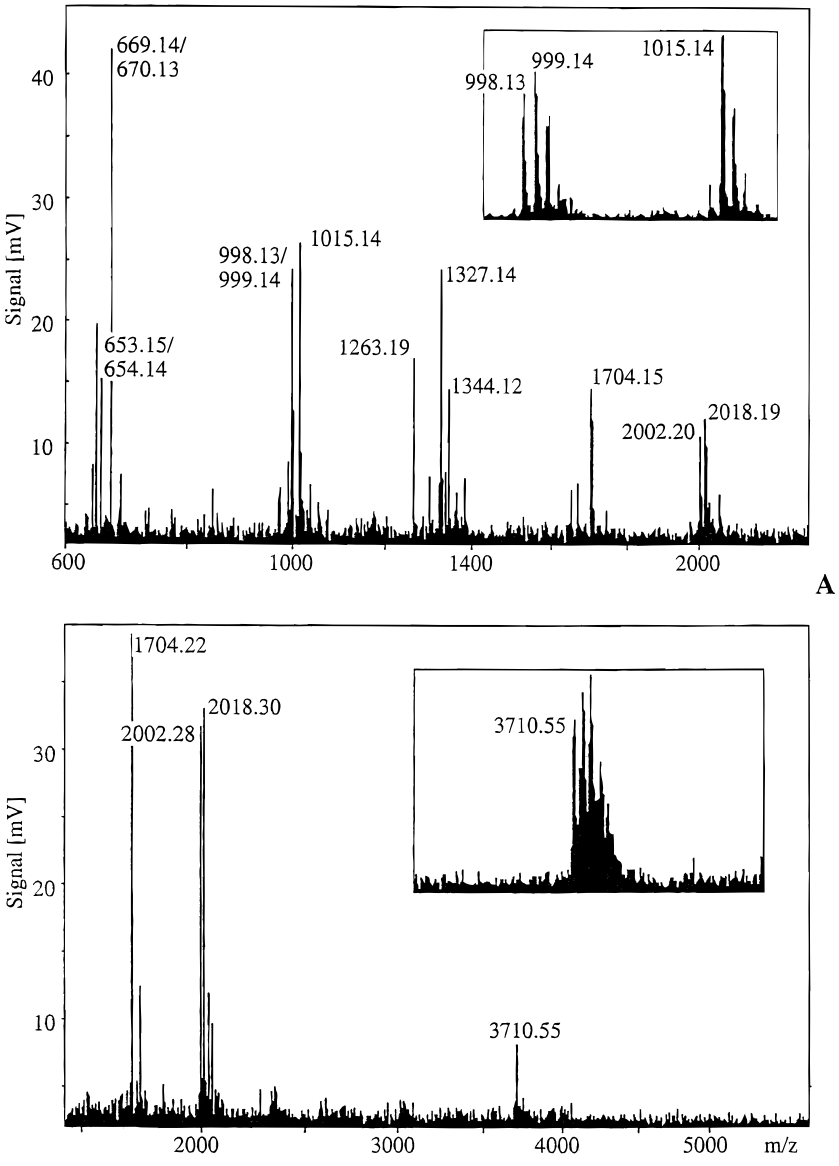
Development of the method

To develop RNase digestion conditions that were directly compatible with MALDI mass spectrometric analysis, we used 5S rRNA from *B. stearothermophilus*, because no posttranscriptional modifications have been found in bacterial 5S rRNA. We could thus work with a substrate that should yield easily interpretable mass spectra upon RNase digestion, provided that the sequence of the 5S rRNA was known. *B. stearothermophilus*, however, contains numerous rRNA operons that might differ in sequence. We therefore sequenced the 5S rRNA by both enzymatic and chemical degradation as well as dideoxy nucleotide termination using reverse transcriptase. Our *B. stearothermophilus* 5S rRNA was apparently homogeneous and the sequence was identical to that published by Zimmermann & Erdmann (1978) except for a C-to-U substitution at position 106 (data not shown).

For the RNase digestions, we chose RNase T1 and RNase A, which have an absolute specificity for guanosine and pyrimidine nucleotides, respectively, and these enzymes are commercially available in high quality. Complete digestion is crucial for comparison of obtained versus sequence-derived mass data to identify posttranscriptionally modified fragments. The secondary and tertiary structure of RNA renders many nucleotides inaccessible to RNases, and it is therefore necessary to perform the digestion under denaturing conditions. The traditional procedures using urea or acidic sodium phosphate with or without previous heating did not completely digest the RNA and furthermore had an adverse effect on the MALDI preparation. We then tested whether 3-hydroxypicolinic acid (3-HPA), one of the most common matrices in MALDI-MS of nucleic acids, would serve as a buffer and denaturing agent for the RNase digestions. At a concentration of about 5 g/L, 3-HPA served as an adequate buffer for both RNase T1 and RNase A, and the digest could be

used directly for MALDI-MS. Though formation of alkali salts is a major problem in mass spectrometry of nucleic acids (Nordhoff et al., 1996), it proved unnecessary to purify the digest solution further before MALDI analysis. However, a tiny amount of ammonium-loaded ion exchange beads (Nordhoff et al., 1992) may be added to the digest about 1 h prior to MALDI sample preparation. This was particularly the case when RNA from the halophilic archaea was analyzed (see below).

The outcome of an RNase A digestion of the *B. stearothermophilus* 5S rRNA is displayed in Figure 1, together with a table of all expected dinucleotides or larger. All the major peaks in the spectrum are easily assignable to expected digest fragments, and all expected fragments generate distinct signals. Thus, signal suppression, that is, severely reduced intensity of certain of the expected signals, which is often observed in MALDI analysis, is not a problem here. There is, however, a tendency for fragments ending in a uridine to be less intense than those ending in a cytidine. The RNase A digestions were usually carried out for 1–4 h, because longer digestion times occasionally revealed undesired cleavages, particularly between neighboring adenosines. Though the dinucleotide signals are clearly present, it is difficult to identify unambiguously posttranscriptional modifications in this m/z range. First, other signals originating from the matrix or substances in the digest crop up in this m/z range, making it problematic to assign a signal in this region to an assumed posttranscriptionally modified fragment. Secondly, numerous identical dinucleotides originating from different positions in the RNA are frequently generated (see, e.g., Fig. 1B), whereas a modified dinucleotide will only occur in a single copy. The relative signal intensity of the modified dinucleotide may thus be so weak that it will be overlooked.



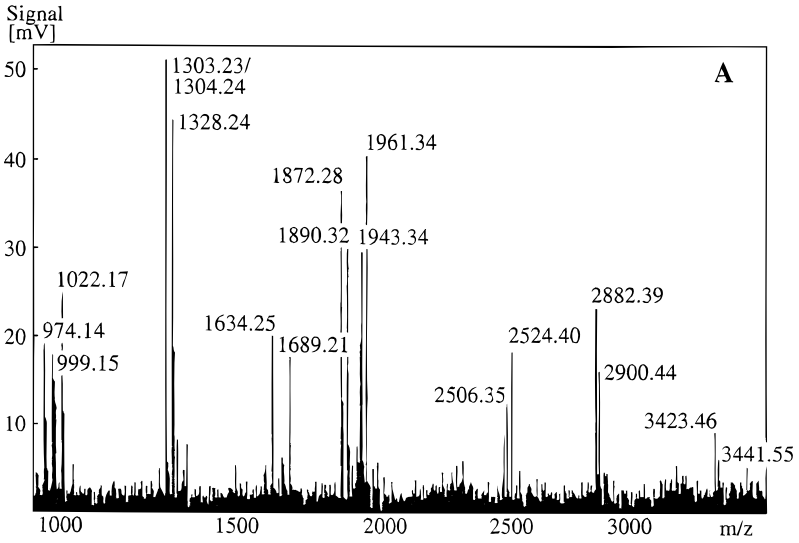
B

Position	m/z	Sequence
27- 28	653.11	AC
46- 47	653.11	AC
37- 38	654.10	AU
67- 68	669.11	GC
89- 90	669.11	GC
95- 96	669.11	GC
108-109	669.11	GC
111-112	669.11	GC
31- 32	670.10	GU
13- 15	998.16	AGC
64- 66	998.16	AGC
86- 88	998.16	AGC
4- 6	999.14	AGU
10- 12	999.14	GAU
70- 72	999.14	GAU
76- 78	999.14	AGU
7- 9	1015.14	GGU
73- 75	1015.14	GGU
114-117	1263.24	AGGC
42- 45	1327.21	GAAC
55- 58	1327.21	AAGC
103-106	1344.19	AGGU
80- 84	1704.25	GGGGC
97-102	2002.30	AAGAGU
48- 53	2018.29	GGAAGU
16- 26	3710.55	GGAGGGGAAAC

FIGURE 1. *Bacillus stearothermophilus* 5S rRNA digested with RNase A. **A**: MALDI time-of-flight mass spectrum of the digest. The spectrum was recorded as two separate segments because the digitizer of the instrument could not accommodate all of the desired m/z range in one spectrum. Insert around m/z 1000 shows the peak pattern for two species separated by one mass unit (m/z 998.1 and m/z 999.1) compared to the isotope pattern of a single species (m/z 1015.1). Insert around m/z 3710 demonstrates the unit resolution in the upper m/z range. **B**: Table of the expected fragments above mononucleotides with position, m/z (singly protonated), and sequence.

The mass difference between uridine and cytidine is 1 Da, demanding unit resolution in the m/z range investigated. The use of this resolution is demonstrated in the expanded m/z 1000 segment of Figure 1A, showing two peaks of nearly equal intensity at m/z 998.13 and 999.14, revealing the presence of two groups of fragments with the compositions AGC and AGU, respectively. In contrast, the isotope distribution of the signal at m/z 1015.15 reveals that only fragments with the composition GGU are present in the digest. The largest RNA fragment encountered in this study was at m/z 3710.55, and expansion of this region (Fig. 1A) illustrates that unit resolution was also obtained here.

A digest of the *B. stearothermophilus* 5S rRNA with RNase T1 is shown together with a table of the expected fragments of 3 nt or more (Fig. 2). All expected fragments are again represented, and no major unexplainable peaks arise. It should, however, be noted that some of the peaks appear as doublets with an 18-Da difference. This is because an RNase T1 digestion occur via a 2′-3′ cyclic phosphate intermediate, which is then converted into a 2′ (or 3′) "linear" phosphate. The mass difference between these two phosphate forms is 18 Da. The conversion into the linear phosphate occurs more readily when the neighbor to the last 3′ nucleotide is an adenosine as opposed to a pyrimidine. The conversion of the cyclic phosphate may be enhanced by increasing the enzyme/substrate ratio or by prolonged digestion time to simplify spectrum interpretation. Alternatively, one can aim at a high degree of cyclic phosphate products using the reverse means. We have not observed undesired cleavages with RNase T1; hence 16-h digestions were often performed. In contrast to the RNase A digest, signal suppression should be taken into account when using RNase T1. Fragments containing many uridines, par-



B

Position	m/z	Sequence
68- 70	974.15	CCG
106-108	975.13	UCG
109-111	975.13	CUG
78- 80	976.12	UUG
71- 73	999.14	AUG
75- 77	999.14	UAG
102-104	999.14	UAG
50- 52	1022.17	AAG
84- 87	1303.19	CCAG
112-115	1304.18	CUAG
96- 99	1327.21	CAAG
11- 14	1328.20	AUAG
53- 57	1634.22	UUAAG
1- 5	1689.19	CCUAG
90- 95	1890.26	CCCCUG
43- 48	1961.31	AACACG
58- 65	2524.35	CCCUCCAG
23- 31	2900.44	AAACACCCG
32- 42	3441.44	UUCCCAUCCCG

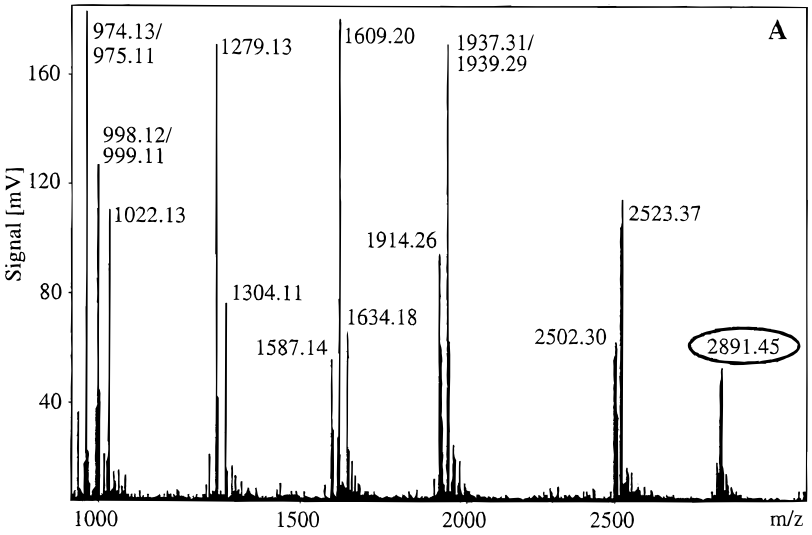
FIGURE 2. *Bacillus stearothermophilus* 5S rRNA digested with RNase T1. **A**: MALDI time-of-flight mass spectrum of the digest. The mixture of cyclic and linear phosphates that result from the digestion is evident in the 18 m/z difference between certain pairs of peaks (e.g., 1872.28/1890.32 and 3423.46/3441.55). **B**: Table of the expected fragments (3′ linear phosphate) above dinucleotides.

ticularly at the 5′ end of the fragment, do not yield very intense signals in the positive ion mode (more clearly seen in Fig. 3).

Analysis of *Sulfolobus acidocaldarius* 5S rRNA

Having optimized the digestion conditions, we applied the method to 5S rRNA purified from the thermophile archaeon *Sulfolobus acidocaldarius*. Its single 5S rRNA gene has been sequenced (Durovic et al., 1994), and our result of sequencing a 0.3-kbp PCR fragment containing the 5S rRNA gene was in complete agreement with that mentioned above (data not shown). The outcome of an RNase T1 digestion of the 5S rRNA together with a table of the expected masses are shown in Figure 3. The expected nonanucleotide signal at m/z 2877.42 (corresponding to nt 25–32) is replaced by a signal at m/z 2891.50. The simplest explanation for this 14-Da mass increment is a posttranscriptional modification in the form of a methylation. After digesting the 5S rRNA with RNase A, an unexpected signal at m/z 1662.24 appears (Fig. 4A). The only possible site of methylation, which could give rise to the extra signal at m/z 1662.24, is on C_{32}, rendering this nucleotide insensitive to RNase A digestion (Fig. 4B). Thus, the extra signal results from a pentanucleotide with the sequence 5′-HO-C_{Me}GGACp-3′ (positions 32–36). Although the fragment from positions 33–36 is absent, a signal at m/z 1343.12 remains, because two other RNase A fragments (positions 18–21 and 110–113) have the same nucleotide composition.

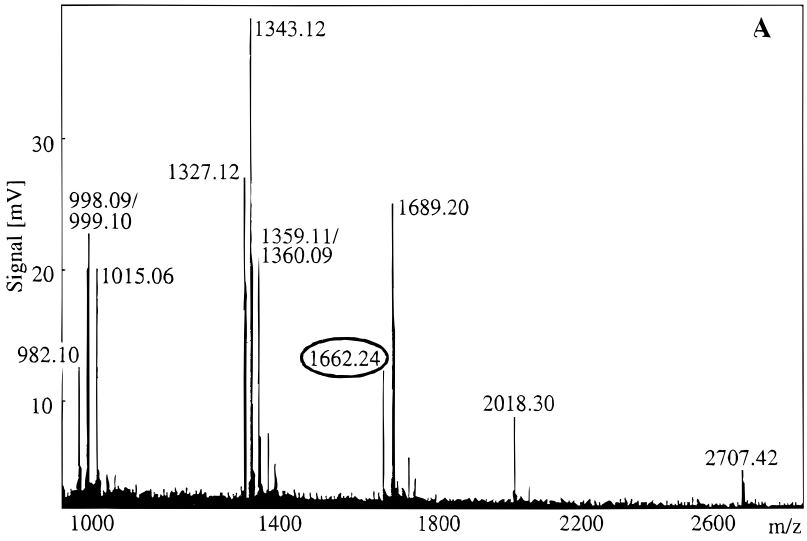
We used PSD fragmentation (Spengler et al., 1992) on the m/z 1662.24 species to verify the presence of a methylated C_{32}. The concept of PSD is to expose the sample briefly to excess laser energy to make the ions more energetic; all ions but the one of interest (here m/z 1662.27, calculated) are then removed by electro-



B

Position	m/z	Sequence
60- 62	974.15	CCG
78- 80	974.15	CCG
113-115	975.13	CUG
100-102	998.16	CAG
118-120	999.14	AUG
52- 54	1022.17	AAG
96- 99	1279.19	CCCG
69- 72	1304.19	UCAG
40- 44	1587.18	UUUCG
35- 39	1609.23	ACUCG
85- 89	1609.23	AUCCG
55- 59	1634.22	UUAAG
63- 68	1914.27	CUCACG
45- 50	1937.29	AACCCG
11- 16	1939.25	UCAUAG
103-110	2502.31	CCUCUCUG
2- 9	2523.37	CCCACCCG
25- 33	2877.41	UAACACCCG

FIGURE 3. *Sulfolobus acidocaldarius* 5S rRNA digested with RNase T1. **A**: MALDI time-of-flight mass spectrum of the digest. **B**: Table of the expected fragments (3′ linear phosphate) above dinucleotides. The circled numbers denote the observed 14 m/z increment.



B

Position	m/z	Sequence
22- 25	1360.19	GGGU
26- 28	982.16	AAC
29- 30	653.11	AC
31- 31	324.06	C
32- 32	338.07	C_{Me} } = 1662.27
33- 36	1343.21	GGAC }

FIGURE 4. *Sulfolobus acidocaldarius* 5S rRNA digested with RNase A. **A**: MALDI time-of-flight mass spectrum of the digest. The circled m/z denotes a signal not anticipated from the gene sequence. **B:** Table of the expected RNase A fragments around the methylated RNase T1 nonanucleotide fragment observed in Figure 3.

static means. The isolated ion fragments substantially in the first field-free drift region of the instrument because of its increased internal energy. By step-wise adjustment of the electrostatic ion reflector, all fragmented ions can be focused onto a detector located after the second field-free drift region. A previous precise calibration of the instrument with fragmented ions of known m/z is used to convert the flight time of the unknown fragmented ions into m/z values. The PSD spectrum is depicted in Figure 5. The fragmentation makes it possible to read a sequence from the 5′ end based on z-ions and y-ions (see Fig. 5 legend for fragmentation nomenclature) in that the successive losses of C_{Me}p, Gp, and Gp are observed. This is in perfect agreement with the methylation of C_{32}, leading to an RNase A fragment with the sequence 5′-HO-C_{Me}GGACp-3′. The z_4 and y_4 ion signals are not prominent in Figure 5, which was obtained with a mixture of 2,3,4- and 2,4,6-trihydroxy-acetophenone (THAP) as matrix. Using 3-HPA instead yields intense z_4 and y_4

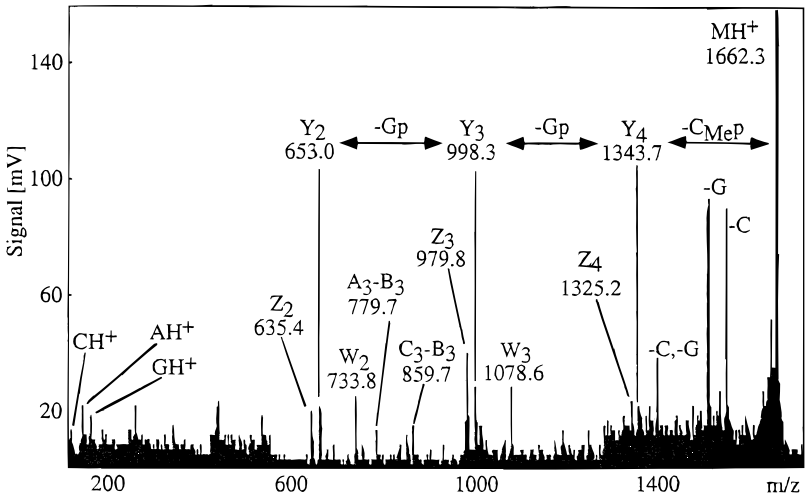


FIGURE 5. MALDI post source decay analysis of the m/z 1662.24 fragment observed in Figure 4. Only fragments arising by a single cleavage of the phosphodiester backbone and fragments resulting from nucleobase loss are indicated. a, b, c, and d ions contain the 5′ end of the original fragment and arise through cleavage of the 3′ proximal phosphodiester backbone between, from the 5′ side, C-O, O-P, P-O, and O-C, respectively. w, x, y, and z ions contain the 3′ end of the original fragment and arise through cleavage of the 5′ proximal phosphodiester backbone between, from the 5′ side, C-O, O-P, P-O, and O-C, respectively. The number associated with the fragment indicates its length in nucleosides. Loss of a nucleobase from a fragment is marked by $-B_n$, where *n* denotes the nucleobase position read from the 5′ end. The nomenclature of the fragment ions is according to McLuckey et al. (1992).

ions, and we are therefore certain of the assignment of these signals. However, the smaller fragment ions are not so readily generated with 3-HPA (data not shown). Our data located the methyl group to C_{32}, but do not reveal where this nucleotide is modified. Investigations of the 5S rRNA from the related species *Sulfolobus solfataricus* suggested that the modification is located on the 2′ position of the ribose (Bruenger et al., 1993). This modification position would be in accordance with the observed RNase A-resistance of the C_{32}, but other cytidine-methylations may also affect RNase A cleavage similarly.

Analysis of 5S rRNA from halophile archaea

In addition to the identified posttranscriptional modification in the 5S rRNA of *S. acidocaldarius*, modifications have previously been identified in analogous positions in 5S rRNA of the thermophile archaea *S. solfataricus* and *Pyrodictium occultum*. We therefore investigated the two halophile archaea *Halobacterium halobium* and *Haloarcula marismortui* to see if this was a general phenomenon for archaea. The gene sequence of the respective 5S rRNA genes isolated by PCR was determined. In *H. halobium*, the 5S rRNA gene was identical to previously published data (Mankin & Kagramanova, 1986). *H. marismortui* has two rDNA operons, but separate isolation of the two 5S rRNA genes by PCR is easily accomplished because they have different 3′ flanking sequences (Dennis et al., 1998). The 5S rDNA of the *rrnA* operon was in agreement with the data of Dennis and colleagues, whereas the 5S rDNA of the *rrnB* operon contained a cytidine-to-guanosine substitution at position 107 in our *H. marismortui* strain. In our strain, *rrnA* and *rrnB* 5S rRNAs differ at positions 73 (A/G), 106 (U/C), and 107 (C/G).

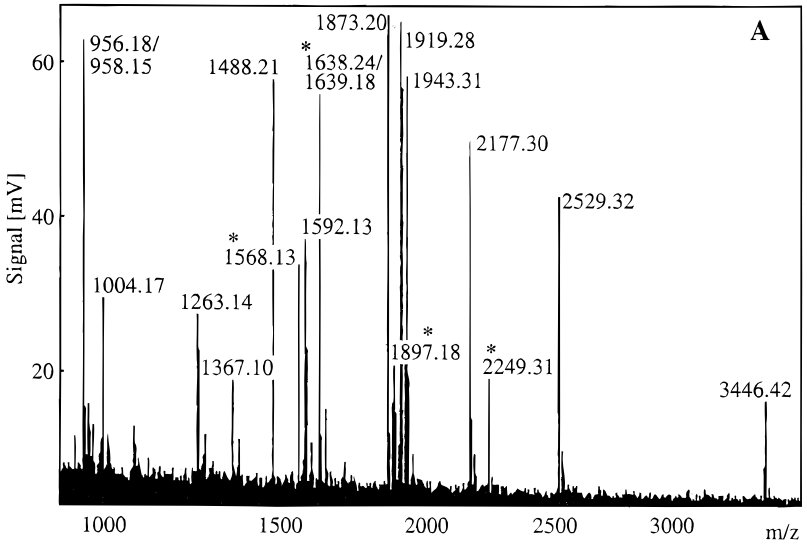
RNase T1 digestion of the *H. halobium* 5S rRNA revealed that the expected signal for nt 1–5 at m/z 1696.18 is replaced by a signal at 1616.22 (data not shown), corresponding to the loss (hydrolysis) of a phosphate. We ascribe this to loss of the original 5′ phosphate of the 5S rRNA because all the other RNase T1 fragments have retained their 3′ phosphates. In an RNase A digest of the *H. halobium* 5S rRNA, the 5′ terminal uridine is released as a mononucleotide and is therefore not observable. We cannot tell when this phosphate loss occurs, that is, whether it is induced during the purification of the 5S rRNA or it is a cellular event. However, the 5′ phosphate was absent in two independent batches of *H. halobium* 5S rRNA, whereas it was present in *H. marismortui* 5S rRNA (Fig. 6), which was purified under identical conditions.

The outcome of an RNase T1 digestion of the mixture of the two *H. marismortui* 5S rRNAs is shown in Figure 6 and did not reveal any posttranscriptional modifications (neither did employing RNase A; data not shown). However, these data corroborate the sensitivity and dynamic range of the approach in that fragments originating from just one of the two 5S rRNAs could be unambiguously detected.

DISCUSSION

The method described here is a fast and sensitive approach to screen a given RNA for posttranscriptional modifications based on MALDI-MS. All spectra shown were obtained with no more than 2.5 pmol of 5S rRNA, corresponding to less than approximately 100 ng of RNA. The typical procedure involves digestion of 1 pmol of RNA in a 1–2 μL volume followed by direct MALDI sample preparation of the entire digest mixture. The only purification step applied was a prior incubation with ammonium-loaded cation exchange beads to remove alkali ions. The use of 3-HPA as digestion buffer eliminated any need for removal of adverse buffer components. Quite complex mixtures can be analyzed directly because MALDI-MS predominantly generates singly charged ions, which results in straightforward interpretable spectra. The examples shown contain typically around 15 signals originating from trinucleotides or larger. Even though these fragments may be separated by only 1 Da in mass, all fragment signals could be identified because the unit resolution was obtained up to approximately m/z 4000. Suppression of signals from certain fragments would be a major obstacle to the described approach, but severe signal suppression is fortunately not observed. This contrasts strongly with application of MALDI-MS to peptide mapping, where full sequence coverage is essentially never achieved. This can be ascribed to the relatively similar chemical properties of the nucleotides, resulting in similar mass spectrometric behavior of oligonucleotides and contrasting with the widely varying characteristics of amino acids and peptides. The main important exception is the lack of a basic functionality (low proton affinity) for uridine, and oligonucleotides with a high uridine content accordingly yielded less intense signals in the positive ion mode. Fortunately, reversing the polarity of ion detection can alleviate the problem because the nucleotides differ to a lesser extent in their acidities.

Only RNase fragments from the trinucleotide level and upwards are truly useful, inasmuch as matrix signals obscure the mono- and dinucleotide regions. With the use of RNase T1 and RNase A, not every nucleotide can therefore be expected to be observed in a fragment within the diagnostic m/z range. In the different 5S rRNAs investigated in this work, between 1 and 6 nt out of approximately 120 could not be covered in the screen for posttranscriptional modifications. Supplementing with an adenosine nucleotide-specific RNase would be required if any nucleotide, regardless of sequence context, should be present in a trinucleotide or larger.



B

Position	m/z	Sequence
115-117 (AB)	956.14	CCG
23- 25 (AB)	958.11	UUG
51- 53 (AB)	1004.16	AAG
111-114 (AB)	1263.15	UUCG
1- 4 (AB)	1367.12	UUAG
118-122 (AB)	1488.26	CCACC
69- 73 (B)	1568.19	UUCCG
79- 83 (AB)	1592.19	UACUG
103-107 (B)	1638.25	AAACG
54- 58 (AB)	1639.25	AUAAG
95-100 (AB)	1873.22	CCUCU
69- 74 (A)	1897.23	UUCCAG
9- 14 (AB)	1919.28	CCACAG
44- 49 (AB)	1943.29	AACACG
26- 32 (AB)	2177.28	CCUCCCG
103-109 (A)	2249.31	AAAUCCG
59- 66 (AB)	2529.36	CCCACCAG
33- 43 (AB)	3446.45	UACCCAUCCCG

FIGURE 6. *Haloarcula marismortui* 5S rRNA digested with RNase T1. **A**: MALDI time-of-flight mass spectrum of the digest. The peaks marked with asterisks are unique for a single of the two 5S rRNA species. **B**: Table of the expected fragments (3′ cyclic phosphates) above dinucleotides. Fragments originating from the A, the B, or both alleles of the 5S rRNA are marked accordingly in the column labeled Position.

Our approach will evidently not recognize mass-neutral posttranscriptional modifications. This means that, for example, pseudouridines are systematically overlooked because RNase A is not inhibited by pseudouridine (B. Porse & F. Kirpekar, unpubl.). We are presently working on combining the carbodiimide modification method of pseudouridine (Bakin & Ofengand, 1993) with MALDI-MS.

We have not investigated the upper size limit of the sample amenable to digestion and mass spectrometric analysis; the largest RNA tested so far was 160 nt and could be analyzed satisfactorily. Digestion of, for example, 1 pmol of a 23S rRNA (around 3,000 nucleotides) would generate at least 1 pmol of any fragment appearing at a given m/z, which is well within the sensitivity of the method. However, the risk that a posttranscriptional modification would alter the mass of one fragment so that it coincided with the mass of another fragment increases with the size of the analyte. This modification may go undetected if a different fragment also represented the sequence-derived (unmodified) mass. Hence, such a mass shift would only affect the signal intensities, which does not generally reflect the abundance of a species, because the peak pattern is qualitative rather than quantitative. An off-line HPLC separation of the digest mixture would be a conceivable, but time consuming, solution. However, this shortcoming should be viewed in relation to the sequence coverage that is possible with this method.

It is important to know the gene sequence of the RNA to be analyzed because unexpected signals could be interpreted as a posttranscriptionally modified fragment. The outcome of our sequencing, which matched previously published data in only three out of five cases, stresses this point. In the case of *B. stearothermophilus*, at least five different sequences, varying in both length and sequence, can be found in databases and in literature, but none of them is identical to the one we analyzed. Thus, the possibility of clonal variations and perhaps sequencing errors has to be taken into ac-

count. The RNase fragment pattern may, however, also be indicative of sequence variations. The parallel use of two RNases allows one to distinguish between sequence variations and posttranscriptional modifications, making the gene sequencing nonessential.

RNA fragmentation by mass spectrometry is a scarcely investigated phenomenon, and the use of PSD on oligoribonucleotides has not previously been reported. DNA fragmentation is initiated by loss of a nucleobase, which is normally followed by cleavage of the 3′ C-O bond (reviewed by Nordhoff et al., 1996). Whether RNA would follow the same fragmentation pathway in MALDI-PSD analysis was not known, but fast atom bombardment mass spectrometry (Wolter et al., 1995) and ESI-MS data (Wolter & Engels, 1995) suggested that this might not be the case, because nucleobase loss was generally not observed in these experiments. We tested the two most versatile matrices for MALDI-MS of nucleic acids, 3-HPA and THAP, for PSD performance. The two matrices yielded qualitatively similar PSD spectra, but 3-HPA favored the production of fragments in the higher m/z range whereas THAP gave more ions in the mid/low m/z range. The RNase A fragment 5′-HO-C_{Me}GGACp-3′ (Fig. 5) exhibited neutral loss of the nucleobases cytosine and guanine but not adenosine, a trend also observed with other oligoribonucleotides (F. Kirpekar, unpubl.). In the other end of the m/z scale, the protonated nucleobases adenosine, guanosine, and, to a lesser extent, cytosine are seen. The fragment ions useful for sequence determination belonged to the 3′ series (nomenclature according to McLuckey et al., 1992): y_{2-4}, z_{2-4}, and w_{2-3} ions were observed. Isolated 5′ ions (a and c type) also occurred, but sequence information could not be derived from these. Though we can only draw conclusions from a limited set of MALDI-PSD data, we found that base loss *is* a prominent fragmentation reaction for RNA, but ions of the type a_n-B_n, in contrast to what is observed for DNA, do not frequently occur. It should be noted that the PSD data presented here were acquired in positive ion mode, whereas essentially all previous studies on the fragmentation behavior of RNA were done with negative ions. This should be taken into account when comparing our results with earlier findings. While PSD has distinct advantages as an additional analytical tool, the mass accuracy is clearly below that of the normal reflector time-of-flight mode. Mass accuracy drops from around 0.1 Da (up to 4,000 Da) to around 0.5 Da with the PSD signals. This means that distinction between C and U is not possible, and, hence, that *de novo* sequencing is not feasible.

We have identified a methylcytidine in position 32 of *S. acidocaldarius* 5S rRNA, the last cytidine in the RNase T1 fragment 5′-UAACACCCG-3′. The analogous cytidine has been reported to be 2′-*O*-methylated in *S. solfataricus* and N_4-acetylated in *P. occultum*, two other thermophile archaea (Bruenger et al., 1993). The corresponding RNase T1 fragment in the thermophile bacterium *B. stearothermophilus* has the sequence 5′-AAACACCCG-3′, and is (as in the other microorganisms mentioned above) preceded by three guanosine nucleotides, but the fragment is clearly unmodified. In *H. halobium*, the corresponding sequence motif is 5′-GGGUUACUCCCG-3′, and also unmodified. The same holds for *H. marismortui*, where the sequence similarity is lower. Our data thus suggest that modifications at this position are specific to thermophile archaea. To our knowledge, there are no reports that this particular part of the 5S rRNA interacts with other molecules, making speculations about the function of these posttranscriptional modifications difficult, but an increased level of rRNA 2′-*O*-methylations in response to increased growth temperature has recently been reported (Noon et al., 1998). We have also observed that the 5′ end of *H. halobium* 5S rRNA does not harbor a phosphate group. The biological significance of this observation is, however, not clear, and further investigations are required before definite conclusions can be drawn.

We feel that the approach presented here would be very useful for researchers working on RNA modifications of various classes. Its experimental simplicity and high sensitivity combined with the increasing acceptance of mass spectrometry as an important tool in the biological sciences would, at this point, allow implementation of our results in numerous laboratories.

MATERIALS AND METHODS

rRNA purification

The purification of the *B. stearothermophilus* (strain NCA 1503) 5S rRNA has been described (Douthwaite et al., 1982). Growth and harvesting of *S. acidocaldarius* (strain DSM 639) can be found in Brock et al. (1972). The pellet from 500 mL of *S. acidocaldarius* culture was resuspended in 2 mL of 20 mM HEPES-KOH (pH 7.6), 10 mM $MgCl_2$, 50 mM NH_4Cl, and 3 mM DTT and lysed by vortexing 10 times for 30 s with ~2 mL of acid-washed glass beads (∅ 212–300 μm, Sigma-Aldrich Chemie GmbH). The lysate was incubated with 3 μg of RQ1 DNase (Promega) at room temperature for 30 min, and glass beads/cell debris were pelleted by centrifugation (Beckman JA20/8,000 rpm/20 min/4 °C). Crude ribosomes were collected by centrifugation (Beckman Ti50/40,000 rpm/150 min/4 °C). The pellet was dissolved in 0.3 mL TE buffer, extracted three times with phenol/chloroform, and the RNA was precipitated with ethanol. The 5S rRNA band was excised from a 10% polyacrylamide/8 M urea gel and eluted as described (Kirpekar et al., 1994). Growth and harvesting of *H. halobium* (strain R_1) and *H. marismortui* was performed according to Rodriguez-Valera (1995). Cell pellets from 400 mL of culture were dissolved in 3 mL of 3 M KCl, 70 mM HEPES-KOH (pH 7.8), 60 mM Mg-acetate, and 3 mM DTT. Cell disruption and the entire 5S rRNA purification were performed identically as described for *S. acidocaldarius*.

DNA and RNA sequencing

The pellet from a 250-mL culture of *S. acidocaldarius* was resuspended in 2 mL TE buffer/50 mM NaCl and vortexed with 2 mL acid-washed glass beads as described above, and an additional 4 mL TE buffer/50 mM NaCl was added. The pellet of 250 mL *H. halobium* or *H. marismortui* culture was lysed by briefly mixing with 6 mL TE buffer/50 mM NaCl. 0.6 mL 10% SDS was added to the lysates and incubated for 30 min on ice. Then 2.65 mL of 3.4 M KCl was added followed by a further 60 min incubation on ice. The cell debris was pelleted by centrifugation (JA20/10,000 rpm/10 min/4 °C), and the DNA was precipitated by the addition of 2 vol of ethanol followed by immediate centrifugation (JA20/7,000 rpm/15 min/20 °C). The pellet was washed with 70% ethanol and dissolved in 4 mL TE buffer. The DNA was treated with 200 μg RNase A and 200 U RNase T1 for 4 h at room temperature, and then with 200 units proteinase K for 4 h at 37 °C. The DNA was extracted twice with phenol/chloroform, ethanol precipitated, and redissolved in TE-buffer. About 1 μg of genomic DNA was used for PCR. The primer positions were 415–437 and 744–721 (Durovic et al., 1994) for *S. acidocaldarius*, 5859–5880 and 6131–6110 (Mankin & Kagramanova, 1986) for *H. halobium*, 5416–5436 and 6185–6166 (Dennis et al., 1998) for the *rrnA* allele of *H. Marismortui*, and 5493–5513 and 5951–5930 (Dennis et al., 1998) for the *rrnB* allele of *H. marismortui*. The PCR fragments were purified using a "High pure PCR product purification kit" (Boehringer Mannheim GmbH), sequenced from both ends with an "ABI Prism BigDye terminator cycle sequencing kit" (PE Applied Biosystems), and analyzed on an ABI Prism 310 sequencer (PE Applied Biosystems).

The *B. stearothermophilus* 5S rRNA was sequenced chemically according to Peattie (1979) and enzymatically as described by Donis-Keller et al. (1977) and Donis-Keller (1980). Sequencing by dideoxy nucleotide termination, using a 5′-[^{32}P]-end-labeled DNA primer complementary to position 117–100 of the 5S rRNA and AMV reverse transcriptase (Life Sciences) has been described (Stern et al., 1988).

RNase digestions and mass spectrometry

One to two picomoles of purified 5S rRNA were digested in 1–2 μL containing 5 g/L 3-HPA and 2% acetonitrile (MeCN). RNase A (Sigma-Aldrich Chemie GmbH) digestions were performed for 1–4 h at 37 °C with 1–2 μg of enzyme. RNase T1 digestions were done for 16 h at 37 °C with 5–10 U of enzyme (Boehringer Mannheim GmbH). The RNase stock solutions had been heat-treated (85 °C/20 min) before use. The digest solutions were frequently incubated with ~0.1 μL of ammonium-loaded cation beads (Nordhoff et al., 1992) for 30–60 min prior to MALDI sample preparation. The RNase digest was mixed on the MALDI sample support with 0.7 μL 3-HPA in 50% MeCN and ~0.1 μL 0.3 M diammonium citrate, and the preparation was air dried. MALDI mass spectra were recorded on a Bruker Reflex II instrument (Bruker-Franzen Analytik GmbH) in reflector mode using delayed ion extraction. The spectra were normally recorded in the positive ion mode. Post source decay analysis was performed directly from the 3-HPA preparation or from a separate preparation with 1 μL of a 2:1:1 mixture of 0.2 M 2,4,6-trihydroxyacetophenone in 50% MeCN:2,3,4-trihydroxyacetophenone in 50% MeCN:0.3 M diammonium citrate as matrix. All matrix compounds were obtained from Sigma-Aldrich Chemie GmbH.

ACKNOWLEDGMENTS

We thank Bo Porse, University of Copenhagen, Denmark, for supplying the *S. acidocaldarius* cells; Birte Vester, University of Copenhagen, Denmark, for helpful discussions on the halophile archaea; and Pia N. Nielsen for help with the DNA sequencing. The *H. marismortui* strain was obtained from JPG Ballesta, Madrid, Spain. This work was made possible by a grant from the Carlsberg Foundation to F.K.

Received September 28, 1999; returned for revision October 18, 1999; revised manuscript received November 8, 1999

REFERENCES

Bakin A, Ofengand J. 1993. 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. *Biochemistry 32*:9754–9762.

Brock TD, Brock KM, Belly RT, Weiss RL. 1972. Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Microbiol 84*:54–68.

Bruenger E, Kowalak JA, Kuchino Y, McCloskey JA, Mizushima H, Stetter KO, Crain PF. 1993. 5S rRNA modification in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Pyrodictium occultum*. *FASEB J 7*:196–200.

Cundliffe E. 1987. On the nature of antibiotic binding sites in ribosomes. *Biochimie 69*:863–869.

Cunningham PR, Richard RB, Weitzmann CJ, Nurse K, Ofengand J. 1991. The absence of modified nucleotides affects both in vitro assembly and in vivo function of the 30S ribosomal subunit of *Escherichia coli*. *Biochimie 73*:789–796.

Dennis PP, Ziesche S, Mylvaganam S. 1998. Transcriptional analysis of two disparate rRNA operons in the halphilic archaeon *Haloarcula marismortui*. *J Bacteriol 180*:4804–4813.

Donis-Keller H. 1980. Phy M: An RNase activity specific for U and A residues useful in RNA sequence analysis. *Nucleic Acids Res 8*:3133–3142.

Donis-Keller H, Maxam AM, Gilbert W. 1977. Mapping adenines, guanines, and pyrimidines in RNA. *Nucleic Acids Res 4*:2527–2538.

Douthwaite S, Christensen A, Garrett RA. 1982. Binding site of ribosomal proteins on prokaryotic 5S ribonucleic acids: A study with ribonucleases. *Biochemistry 21*:2313–2320.

Durovic P, Kutay U, Schleper C, Dennis PP. 1994. Strain identification and 5S rRNA gene characterization of the hyperthermophilic archaebacterium *Sulfolobus acidocaldarius*. *J Bacteriol 176*:514–517.

Felden B, Hanawa K, Atkins JF, Himeno H, Muto A, Gesteland RF, McCloskey JA, Crain PF. 1998. Presence and location of modified nucleotides in *Escherichia coli* tmRNA: Structural mimicry with tRNA acceptor branches. *EMBO J 17*:3188–3196.

Green R, Noller HF. 1996. In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for *Escherichia coli* 50S ribosomal subunit assembly and function. *RNA 2*:1011–1021.

Green R, Noller HF. 1999. Reconstitution of functional 50S ribosomes from in vitro transcripts of *Bacillus stearothermophilus* 23S rRNA. *Biochemistry 38*:1772–1779.

Karas M, Hillenkamp F. 1988. Laser desorption ionisation of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem 60*:2299–2301.

Khaitovich P, Tenson T, Mankin AS, Green R. 1999a. Peptidyl transferase activity catalyzed by protein-free 23S ribosomal RNA remains elusive. *RNA 5*:605–608.

Khaitovich P, Tenson T, Kloss P, Mankin AS. 1999b. Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with in vitro-transcribed RNA. *Biochemistry 38*:1780–1788.

Kirpekar F, Nordhoff E, Kristiansen K, Roepstorff P, Lezius A, Hahner S, Karas M, Hillenkamp F. 1994. Matrix assisted laser desorption/ionization mass spectrometry of enzymatically synthesized RNA up to 150 kDa. *Nucleic Acids Res 22*:3866–3870.

Kowalak JA, Bruenger E, McCloskey JA. 1995. Posttranscriptional modification of the central loop of domain V in *Escherichia coli* 23 S ribosomal RNA. *J Biol Chem 270*:17758–17764.

Kowalak JA, Pomerantz SC, Crain PF, McCloskey JA. 1993. A novel method for the determination of posttranscriptional modification in RNA by mass spectrometry. *Nucleic Acids Res 21*:4577–4585.

Mankin AS, Kagramanova VK. 1986. Complete nucleotide sequence of the single RNA operon of *Halobacterium halobium*: Secondary structure of the archaebacterial 23 S rRNA. *Mol Gen Genet 202*: 152–161.

McLuckey SA, Van Berkel GJ, Glish GL. 1992. Tandem mass spectrometry of small multiply charged oligonucleotides. *J Am Soc Mass Spectrom 3*:60–70.

Noller HF, Hoffarth, Zimniak L. 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science 256*:1416–1419.

Noon KR, Bruenger E, McCloskey JA. 1998. Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus solfataricus*. *J Bacteriol 180*:2883–2888.

Nordhoff E, Ingendoh A, Cramer R, Overberg A, Stahl B, Karas M, Hillenkamp F, Crain PF. 1992. Matrix-assisted laser desorption/ionization mass spectrometry of nucleic acids with wavelengths in the ultraviolet and infrared. *Rapid Commun Mass Spectrom 6*:771–776.

Nordhoff E, Kirpekar F, Roepstorff P. 1996. Mass spectrometry of nucleic acids. *Mass Spectrom Rev 15*:67–138.

Peattie DA. 1979. Direct chemical method for sequencing RNA. *Proc Natl Acad Sci USA 76*:1760–1764.

Raychaudhuri S, Conrad J, Hall BG, Ofengand J. 1998. A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of *Escherichia coli*. *RNA 4*:1407–1417.

Rodriguez-Valera F. 1995. Cultivation of halophile archaea. In: DasSarma S, Fleischmann EM, eds. *Archaea: A laboratory manual. Halophiles*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 13–16.

Rozenski J, Crain PF, McCloskey JA. 1999. The RNA modification database: 1999 update. *Nucleic Acids Res 27*:196–197.

Sirum-Connolly K, Manson TL. 1993. Functional requirement of a site-specific ribose methylation in ribosomal RNA. *Science 262*: 1886–1889.

Skinner R, Cundliffe E, Schmidt FJ. 1983. Site of action of a ribosomal RNA methylase responsible for resistance to erythromycin and other antibiotics. *J Biol Chem 258*:12702–12706.

Spengler B, Kirsch D, Kaufmann R, Jaeger E. 1992. Peptide sequencing by matrix-assisted laser-desorption mass spectrometry. *Rapid Commun Mass Spectrom 6*:105–108.

Stern S, Moazed D, Noller HF. 1988. Structural analysis of RNA using chemical and enzymatic probing by primer extension. *Methods Enzymol 164*:481–489.

Thompson J, Schmidt F, Cundliffe E. 1982. Site of action of a ribosomal RNA methylase conferring resistance to thiostrepton. *J Biol Chem* 257:7915–7917.

Weisblum B. 1995. Erythromycin resistance by ribosome modification. *Antimicrob Agents Chemother 39*:577–585.

Wolter MA, Engels JW. 1995. Nanospray mass spectrometry/mass spectrometry for the analysis of modified oligoribonucleotides. *Eur Mass Spectrom 1*:583–590.

Wolter MA, Engels JW, Montanarella L, Tilio R, Facchetti S. 1995. Influence of the matrix on the analysis of small oligoribonucleotides by fast atom bombardment mass spectrometry. *J Mass Spectrom 30*:485–491.

Zimmermann J, Erdmann VA. 1978. Binding sites of *E. coli* and *B. stearothermophilus* ribosomal proteins on *B. stearothermophilus* 5S RNA. *Nucleic Acids Res 5*:2267–2288.