Nucleic Acids Research, 1997, Vol. 25, No. 10 **1957–1964**

Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI) of endonuclease digests of RNA

Stephanie Hahner, Hans-Christian Lüdemann, Finn Kirpekar[1], Eckhard Nordhoff[1], Peter Roepstorff[1], Hans-Joachim Galla[2] and Franz Hillenkamp*

Institute for Medical Physics and Biophysics, University of Münster, Robert-Koch-Strasse 31, D-48149 Münster, Germany, [1]Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark and [2]Department of Biochemistry, University of Münster, Wilhelm-Klemm-Str. 2, D-48149 Münster, Germany

Recieved January 24, 1997; Revised and Accepted April 3, 1997

ABSTRACT

The determination of RNA sequences using base-specific enzymatic cleavages is a well established method. Different synthetic RNA molecules were analyzed for uniformity of degradation by RNase T_1, U_2, A and PhyM under reaction conditions compatible with Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS), to identify the positions of G, A and pyrimidine residues. In order to get a complete set of fragments derived from cleavage at every phosphodiester bond, the samples were also subjected to a limited alkaline hydrolysis. Additionally, the 5′-terminus fragments of a 49mer RNA transcript were isolated by way of 5′-biotinylation and streptavidin-coated magnetic beads (Dynal), followed by a RNase U_2 digestion. MALDI-MS of the generated fragments is presented as an efficient technique for a direct read out of the nucleotide sequence.

INTRODUCTION

Sequencing of RNA is an important step in elucidating its structural features and function. Several fields of research such as the detection of posttranscriptional modifications in rRNA and tRNA (1,2), the determination of the genomic sequences of different viruses (3) and studies in the field of molecular evolution (4) all require rapid and reliable sequencing techniques. Classical ‘Sanger’ chain termination sequencing has been used for this purpose, following transcription to cDNA with a reverse transcriptase (5). Major limitations of this procedure are a high error rate of the reverse transcriptase (6) and the tendency for premature transcription termination caused by secondary structures of the RNA template (7,8). Although several strategies have been developed to overcome the latter problem (8,9), base-modified nucleotides—highly abundant in RNA molecules—limit correct transcription (6).

Another promising approach for a direct sequencing of RNA involves enzymatic digestion in combination with postlabeling. Several ribonucleases with a high degree of base specificitiy have been employed to yield uniform cleavage patterns. Fragments are generated both by partial as well as complete digests. The introduction of polyacrylamide gel electrophoresis for RNA sequencing improved the direct readout (10). However, sequence determination based solely on electrophoretic or chromatographic mobility still has several limitations. Due to modified nucleotides or regions of tight secondary structure, anomalous migration in the gel has been observed, evident for example, as band compressions (11,12). Furthermore, long analysis time and the requirement for labeling or staining impair the utility of chromatographic and electrophoretic separation techniques. Mass spectrometric techniques such as MALDI-(13) and ESI-MS (14) provide new means for the analysis of nucleic acids (15). Mass is a more direct and specific property for identification of fragments as compared to electrophoretic or chromatographic mobility. Because labeling or staining is not required and automation can be readily implemented, mass spectrometry offers the potential for fast and large-scale analyses. Application of ESI-MS for detection and sequence location of posttranscriptionally modified nucleotides in ribosomal RNA has been demonstrated by McCloskey and co-workers (1,16). Reverse-phase HPLC had to be applied after enzymatic hydrolysis to facilitate the subsequent analysis by ESI-mass spectrometry. In general, the analysis of oligonucleotide mixtures is difficult with this technique because of the distribution of different mass dependent charge states for each individual fragment ion, which cannot be sufficiently resolved (17), unless complex and expensive Fourier–Transform mass spectrometers are employed (18). MALDI generates predominately singly-charged ions. A separation of oligonucleotide fragments, generated by enzymatic or chemical cleavages prior to the mass spectrometric analysis, is, therefore, usually not required, unless very complex mixtures are analyzed. However, mass analysis of oligonucleotides is compromised by salts, buffers and enzymes used for the various sequencing approaches, mainly because of the formation of heterogeneous alkali salts of the oligonucleotides (19). For the application of MALDI-MS to a direct read-out of the products of RNA sequencing reactions, it is highly desirable to establish reaction conditions compatible with both the activity of the employed enzymes and the preparation requirements for MALDI.

The use of mono-specific RNases such as RNase T_1 (G specific) and RNase U_2 (A specific) has become routine (10,20–22).

*To whom correspondence should be addressed. Tel: +49 251 835 5103; Fax: +49 251 835 5121; Email: hillenk@uni-muenster.de

Another enzyme, the chicken liver ribonuclease (RNase CL3) has been reported to cleave preferentially at cytidine, but the enzyme's proclivity for this base has been reported to be affected by the reaction conditions (23). Recent reports also claim cytidine specificity for another ribonuclease, cusativin, isolated from dry seeds of *Cucumis sativus L* (24). Alternatively, the identification of pyrimidine residues by use of the RNase PhyM (A and U specific) (25) and RNase A (C and U specific) (20,21) has been demonstrated. In order to reduce ambiguities in sequence determination, additional limited alkaline hydrolysis can be performed. Since every phosphodiester bond is potentially cleaved under this conditions, information about omitted and/or unspecific cleavages can be obtained this way (10,20). Chemical cleavage is not known to be strongly affected by secondary structure, whereas base-specific enzymatic degradation turned out to be (26,27). Therefore, denaturing conditions which do not disturb the activity and specificity of the RNases are recommended for further improvement towards uniform cleavage (10,20,23,25).

In the present study an approach utilizing base-specific ribonucleases in combination with MALDI-MS is described as an alternative method for RNA sequencing. Synthetic oligonucleotides with three different sequences in the range of 20–25 nucleotides were used to examine the specificity of the ribonucleases under optimized reaction conditions. For an isolation and exclusive detection of 5′-terminal fragments streptavidin-coated magnetic beads were applied to a mixture generated by RNase U_2 digest of a 49mer *in vitro* RNA-transcript biotinylated at the 5′-end.

MATERIALS AND METHODS

Materials

Synthetic RNA (sample A: 5′-UCCGGUCUGAUGAGUCCGUGAGGAC-3′; sample B: 5′-GUCACUACAGGUGAGCUCCA-3′; sample C: 5′-CCAUGCGAGAGUAAGUAGUA-3′) were obtained from DNA technology (Aarhus, Denmark) and purified on a denaturing polyacrylamide gel (28). RNases T_1 (Eurogentec, Seraing, Belgium), U_2 (Calbiochem, Bad Soden, Germany), A (Boehringer Mannheim, Germany) and PhyM (Pharmacia, Freiburg, Germany) were used without additional purification. Streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin, Dynal) were supplied as a suspension of $6–7 \times 10^8$ beads/ml (10 mg/ml) dissolved in phosphate-buffered saline (PBS) containing 0.1% BSA and 0.02% NaN_3. 3-Hydroxypicolinic acid (3-HPA) (Aldrich) was purified by a separate desalting step before use as described in more detail elsewhere (18).

Methods

In vitro transcription reaction. The 5′-biotinylated 49 nt *in vitro* transcript (AGGC CUG CGG CAA GAC GGA AAG ACC AUG GUC CCUNAUC UGC CGC AGG AUC) was produced by transcription of the plasmid pUTMS2 (linearized with the restriction enzyme *Bam*HI) with T7 RNA polymerase (Promega). For the transcription reaction 3 µg template DNA and 50 U T7 RNA polymerase were used in a 50 µl volume containing 1 U/µl RNA guard (RNase inhibitor, Pharmacia), 0.5 mM NTPs 1.0 mM 5′-biotin-ApG dinucleotide, 40 mM Tris–HCl (pH 8.0), 6 mM $MgCl_2$, 2 mM spermidine and 10 mM DTT. Incubation was performed at 37°C for 1 h; then another aliquot of 50 U T7 RNA polymerase was added and incubation was continued for another hour. The mixture was adjusted to 2 M NH_4-acetate and the RNA was precipitated by addition of 1 vol ethanol and 1 vol isopropanol. The precipitated RNA was collected by centrifugation at 20 000 g for 90 min at 4°C, the pellet was washed with 70% ethanol, dried and redissolved in 8 M urea. Further purification was achieved by electrophoresis through a denaturing polyacrylamide gel as described elsewhere (28). The ratio of 5′-biotinylated to non-biotinylated transcripts was ~3:1.

Ribonuclease assay. For partial digestions with selected RNases different enzyme concentrations and assay conditions were employed as summarized in Table 1. The solvents for each enzyme were selected following the suppliers' instructions. The concentrations of the synthetic RNA samples and the *in vitro* transcript were adjusted to $5–10 \times 10^{-6}$ M.

The reaction was stopped at selected times by mixing 0.6 µl aliqouts of the assay with 1.5 µl of 3-HPA-solution. The solvent was subsequently evaporated in a stream of cold air for the MALDI-MS analysis.

Limited alkaline hydrolysis was performed by mixing equal volumes (2.0 µl) of 25% ammonium hydroxide and RNA sample ($5–10 \times 10^{-6}$ M) at 60°C. Aliquots of 1 µl were taken out at selected times and dried in a stream of cold air. For these samples it turned out to be important to first dry the digests in a stream of cold air, before 1.5 µl of the matrix solution and 0.7 µl of a suspension of NH_4^+-loaded cation exchange polymer beads were added.

Table 1. Overview and assay conditions of the RNases employed in this study

RNase	Source	Concentration (U RNase/µg RNA)	Conditions	Incubation time (max. no. fragments)	Refs
T_1	*Aspergillus oryzae*	0.2	20 mM Tris–HCl, pH 5.7, 37°C	5 min	10
U_2	*Ustilago sphaerogena*	0.01	20 mM DAC, pH 5.0, 37°C	15 min	10
PhyM	*Physarum polycephalum*	20	20 mM DAC, pH 5.0, 50°C	15 min	25
A	bovine pancreas	4×10^{-9}	10 mM Tris–HCl, pH 7.5, 37°C	30 min	33
CL_3	chicken liver	0.01	10 mM Tris–HCl, pH 6.5, 37°C	30 min	23
cusativin	*Cucumis sativus* L.	*0.05ng*	10 mM Tris–HCl, pH 6.5, 37°C	30 min	24

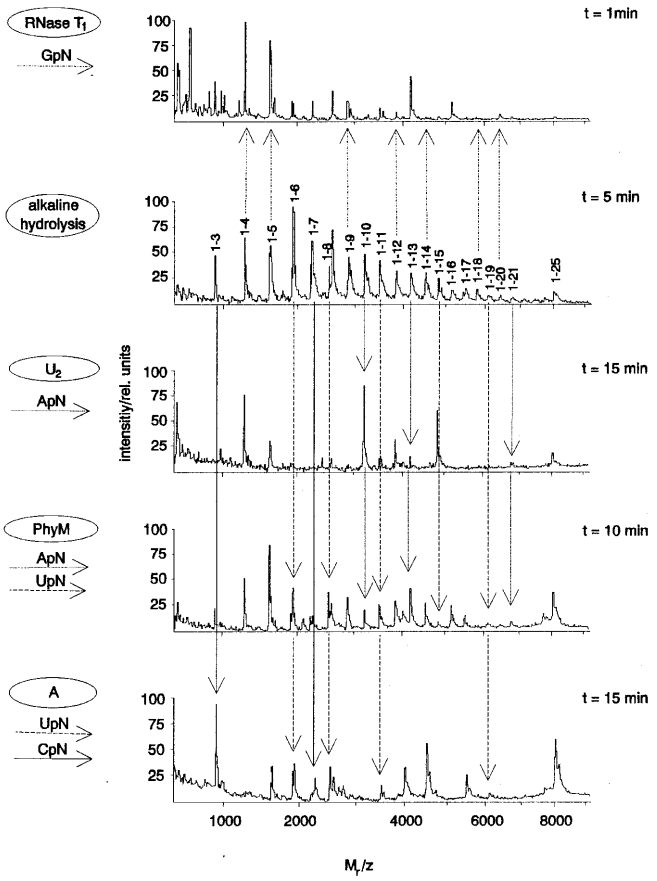

Figure 1. Positive ion UV-MALDI mass spectra of a synthetic RNA 25mer (5′-UCCGGUCUGAUGAGUCCGUGAGGAC-3′) digested with selected RNases. For each enzyme 0.6 µl aliquots of the 4.5 µl assay containing a total of ~20 pmol of the RNA were mixed with 1.5 µl matrix (3-HPA) for analysis. Fragments with retained 5′-terminus are marked by different arrows, specific for the different RNases.

Separation of 5′-biotinylated fragments. Streptavidin-coated magnetic beads were utilized to separate 5′-biotinylated fragments of the *in vitro* transcript after partial RNase degradation. The biotin moiety in this sample was introduced during the transcription reaction initiated by the 5′-biotin-pApG-dinucleotide. Prior to use the beads were washed twice with 2× binding and washing (b&w) buffer (20 mM Tris–HCl, 2 mM EDTA, 2 M NaCl, pH 8.2) and resuspended to 10 mg/ml in 2× b&w buffer. Approximately 25 pmol of the RNA *in vitro* transcript were digested by RNase U2 using the protocol described above. The digestion was stopped by adding 3 µl of 95% formamide containing 10 mM trans-1,2-diaminocyclohexane-N, N, N′, N′-tetraacetic acid (CDTA) at 90°C for 5 min, followed by cooling on ice. Subsequently, capture of the biotinylated fragments was achieved by incubation of 6 µl of the digest with 6 µl of the bead suspension and 3 µl of b&w buffer at room temperature for 15 min. Given the binding capacity of the beads of 200 pmol of biotinylated oligonucleotide per mg of beads, as specified by the manufacturer, the ~2-fold excess of oligonucleotide was used to assure a full loading of the beads. The supernatant was removed, and the beads were washed twice with 6 µl of 100 mM NH_4-acetate and then with 6 µl of H_2O. The biotinylated fragments were eluted from the beads in 6 µl of a solution containing 10 mM CDTA and 95% formamide at 90°C for 5 min. After evaporation of the solvent and the formamide the ≤2.5 pmol of fragments were resuspended in 2 µl H_2O and analyzed by MALDI-MS as described above.

Sample preparatin for MALDI-MS. 3-Hydroxypicolinic acid (3-HPA) was dissolved in ultra pure water to a concentration of ~300 mM. Metal cations were exchanged against NH_4^+ as described in detail previously (28). Aliquots of 0.6 µl of the analyte solution were mixed with 1.5 µl 3-HPA on a flat inert metal substrate. Remaining alkali cations, present in the sample solution as well as on the substrate surface, were removed by the addition of 0.7 µl of the solution of NH_4^+-loaded cation exchange polymer beads. During solvent evaporation the beads accumulate in the center of the preparation and can easily be removed with a pipette tip.

Instrument. A prototype of the Vision 2000 (ThermoBioanalysis, Hemel, Hempstead, UK) reflectron time-of-flight mass spectrometer was used for the mass spectrometry. Ions were generated by irradiation with a frequency-tripled Nd:YAG laser (355 nm, 5 ns; Spektrum GmbH, Berlin, Germany) and accelerated to 10 keV. Delayed ion extraction was used for the acquisition of the spectra shown, as it was found to substantially enhance the signal-to-noise ratio and/or signal intensity. The equivalent flight path length of the system is 1.7 m, the base pressure is 10^{-4} Pa. Ions were detected with a discrete dynode secondary-electron multiplier (R 2362, Hamamatsu Photonics), equipped with a conversion dynode for effective detection of high mass ions. The total impact energy of the ions on the conversion dynode was adjusted to values ranging from 16 to 25 keV, depending on the mass to be detected. The preamplified output signal of the SEM was digitized by a LeCroy 9450 transient recorder (LeCroy, Chestnut Ridge, NY, USA) with a sampling rate of up to 400 MHz. For storage and further evaluation, the data were transferred to a personal computer equipped with custom-made software (ULISSES). All spectra shown were taken in the positive ion mode. Between 20 and 30 single shot spectra were averaged for each of the spectra shown.

RESULTS

Specificity of RNases

Combining base-specific RNA cleavage with MALDI-MS requires reaction conditions, optimized to retain the activity and specificity of the selected enzymes on the one hand and complying with the boundary conditions for MALDI on the other. Incompatibility mainly results because the alkaline-ion buffers, commonly used in the described reaction, such as Na-phosphate, Na-citrate or Na-acetate as well as EDTA interfere with the MALDI sample preparation; presumably they disturb the matrix crystallization and/or analyte incorporation. Tris–HCl or ammonium salt buffers, in contrast, are MALDI compatible (29). Moreover, alkaline salts in the sample lead to the formation of a heterogeneous mixture of multiple salts of the analyte, a problem increasing with increasing number of phosphate groups. Such mixtures result in loss of mass resolution and accuracy as well as signal-to-noise ratio (18,28). Therefore, RNase digestions were carried out under somewhat modified conditions compared with those described in

the literature. They are summarized in Table 1. The concentration of 10–20 mM Tris–HCl and DAC, added to the enzyme reactions was found to not interfere significantly with the MALDI analysis. To examine the specificity of the selected ribonucleases under these conditions, three synthetic 20–25mer RNA molecules with different nucleotide sequences were digested. The MALDI-MS spectra of Figure 1 show five different cleavage patterns of a 25 nt RNA obtained after partial digestion with RNases T_1, U_2, PhyM, A and alkaline hydrolysis. These spectra were taken from aliquots which were removed from the assay after empirically determined incubation times, chosen to get an optimum coverage of the sequence. As the resulting samples were not fractionated prior to mass spectrometric analysis, they contain all fragments generated at that time by the respective RNases. In practice, uniformity of the cleavages can be affected by a preferential attack on specific phosphodiester bonds (10,21,25). However, as indicated in Table 2, the majority of the expected fragments are indeed observed in the spectra. [Table 2 as printed in this publication contains only the observed 5′-fragments and their masses. For a survey of the complete table of all possible fragments the reader is referred to an extended table, accessible via internet (http://medweb.uni.muenster.de/institute/impb/research/hillenkamp/publicat/tabrna.html)].

Table 2. List of the expected vis-à-vis observed 5′-fragments for digests of the synthetic RNA 25mer with the RNases T_1, U_2, PhyM, and A (compare also spectra of Fig. 1)

Enzyme	Sequence 5′UCCGGUCUGAUGAGUCCGUGAGGAC-3′	M_r calculated	M_r measured
RNase T_1	5′UCCGGUCUGAUGAGUCCGUGAGG	7458.5	7456.3
	5′UCCGGUCUGAUGAGUCCGUGAG	7113.3	7112.0
	5′UCCGGUCUGAUGAGUCCGUG	6438.8	6437.9
	5′UCCGGUCUGAUGAGUCCG	5787.5	5788.5
	5′UCCGGUCUGAUGAG	4525.7	4524.1
	5′UCCGGUCUGAUG	3851.3	3851.0
	5′UCCGGUCUG	2870.7	2870.2
	5′UCCGG	1608.0	1608.0
	5′UCCG	1262.8	1262.4
RNase U_2	5′UCCGGUCUGAUGAGUCCGUGAGGA	7787.7	–
	5′UCCGGUCUGAUGAGUCCGUGA	6768.1	6768.7
	5′UCCGGUCUGAUGA	4180.5	4179.7
	5′UCCGGUCUGA	3199.9	3198.8
RNase PhyM	5′UCCGGUCUGAUGAGUCCGUGAGGA	7787.7	–
	5′UCCGGUCUGAUGAGUCCGU	6093.7	6093.1
	5′UCCGGUCUGAUGAGUCCGUGA	6768.1	6769.2
	5′UCCGGUCUGAUGAGU	4831.9	4830.4
	5′UCCGGUCUGAUGA	4180.5	4179.8
	5′UCCGGUCUGAU	3506.1	3505.9
	5′UCCGGUCUGA	3199.9	3199.5
	5′UCCGGUCU	2525.5	2525.5
	5′UCCGGU	1914.2	1914.0
RNase A	*5′UCCGGUCUGAUGAGUCCGU*	6093.7	6094.7
	5′UCCGGUCUGAUGAGUCC	5442.3	–
	5′UCCGGUCUGAUGAGUC	5137.1	–
	5′UCCGGUCUGAUGAGU	4831.9	–
	5′UCCGGUCUGAU	3506.1	3506.2
	5′UCCGGUCU	2525.5	2525.8
	5′UCCGGUC	2219.4	2219.7
	5′UCCGGU	1914.2	1913.6
	5′UCC	917.6	916.8

It has to be kept in mind, however, that Table 2 was compiled by extracting signals from spectra of samples, removed at 2–3 different times from the digest. It is also worth noting that for the reaction protocols as used, correct assignment of all fragment masses is only possible if a 2′, 3′-cyclic phosphate group is assumed. It is well known that such cyclic phosphates are intermediates in the cleavage reaction and get hydrolyzed in a second, independent and slower reaction step involving the enzyme (30–33). In a few cases different fragments have equal mass or differ by as little as 1 Dalton. In these cases mass peaks cannot unambiguously be assigned to one or the other fragments. Digestion of two additional different 20 nt RNA samples was, therefore, performed (34) in order to sort out these ambiguities. For all samples tested, the selected ribonucleases appear to cleave exclusively at the specified nucleotides leading to fragments arising from single as well as multiple cleavages. In Figure 1, peaks, indicating fragments containing the original 5′-terminus, are marked by arrows. All non-marked peaks can be assigned to internal sequences or those with retained 3′-terminus. For a complete sequence all possible fragments bearing exclusively either the 5′- or the 3′-terminus of the original RNA would suffice. In practice, the 5′-fragments are better suited for this purpose, because the spectra obtained after incubation of all three synthetic RNA samples contain the nearly complete set of signals of 5′-ions for all different RNases. Internal fragments are somewhat less abundant and fragments containing the original 3′-terminus appear suppressed in the spectra. In agreement with observations reported in the literature (21) cleavages close to the 3′-terminus were somewhat suppressed in partial digests of the RNA 25mer by RNase T_1 and U_2 (even if they are internal or contain the original 5′-terminus). Fragments from such cleavages appear, if at all, as weak and poorly resolved signals in the mass spectra.

For larger RNA molecules secondary structure is known to influence the uniformity of the enzymatic cleavages (10,20,23,25). This can, in principle, be overcome by adding denaturing reagents such as 5–7 M urea, which are known to not interfere with the activity of the RNases used (10,23,25). However, UV-MALDI-analysis with 3-HPA as matrix is not possible under such high concentrations of urea in the sample. Up to a maximum concentration of 2 M urea in the reaction buffer, MALDI analysis of the RNA 20mer (sample B) and its digests was still possible although significant changes in matrix crystallization were observed (data not shown).

RNases CL_3 and cusativin are enzymes reported to cleave at cytidylic acid residues. Upon limited RNase CL_3 and cusativin digestion of the RNA–20mer (sample B) under non-denaturing conditions, fragments corresponding to cleavages at cytidylic residues were indeed observed (Fig. 2). Similar to the data reported so far (23,24) the degradation pattern, however, reveals that not every cytidine residue is recognized, especially for neighbouring C residues. For a RNA 20mer secondary structure effects should be negligible and the unrecognized cleavage sites must, in this case, be attributed to a lack of specificity of this enzyme. To confirm these data a further RNase CL_3-digestion was performed with the RNA 20mer (sample C). For this analyte, all three linkages containing cytidylic acid were readily hydrolyzed, but additional cleavages at uridylic acid residues were detected as well (data not shown). Since altered reaction conditions such as increased temperature (90°C), various enzyme to substrate ratios, and addition of 2 M urea did not result in a digestion of the expected specificity (data not shown), application of this enzyme to sequencing was not pursued further. Introduction of a new cytidine-specific ribonuclease, cusativin, isolated from dry seeds of *Cucumis sativus* L. looked promising for RNA sequencing (24). Not every cytidine residue was hydrolyzed by this enzyme and additional cleavages occurred at uridylic acid residues for the recommended concentration of the enzyme. RNases CL_3 and cusativin will, therefore, not yield the desired sequence information for mapping of cytidine residues and their use was not further pursued. The distinction of pyrimidine residues can be achieved, however, by use of RNases with multiple specificities, such as *Physarum polycephalum* RNase PhyM (cleaves ApN, UpN) and pancreatic RNase A (cleaves UpN, CpN) (see Fig. 1). All 5′-terminus fragments, generated by the mono-specific RNase U_2 and apparent in the spectrum of Figure 1c, are also evident in the spectrum of the RNase PhyM digest (Fig. 1d). Five of the six uridylic cleavage sites could, this way, be uniquely identified by this indirect method. In a next step the knowledge of the uridine cleavage sites was used to identify sites of cleavage of cytidylic acid residues in the spectrum recorded after incubation with RNase A (Fig. 1e), again using exclusively ions containing the original 5′-terminus. Two of the four expected cleavages were identified this way. A few limitations are apparent from these spectra, if only the fragments containing the original 5′-terminus are used for the sequence determination. The first two nucleotides usually escape the analysis, because their signals get lost in the low mass matrix background. Large fragments with cleavage sites close to the 3′-terminus are often difficult to identify, particularly in digests with RNases T_1 and U_2 because of their low yield (vide supra) and the often strong nearby signal of the non-digested transcript. Accordingly the cleavages in position 22 and 23 do not show up in the spectrum of the G-specific RNase T_1 (Fig. 1a) and the cleavage site 24 cannot be identified from the spectra of the U_2 and PhyM digests (Fig. 1c and d). Also site 16 and 17 with two neighbouring cytidylic acids cannot be identified in the RNase A spectrum of Figure 1e. These observations demonstrate that a determination of exclusively 5′-terminus fragments may not always suffice and the information contained in the internal fragments may be needed for a full sequence analysis.

Finally, limited alkaline hydrolysis provides a continuum of fragments (Fig. 1b), which can be used to complete the sequence data. Again, the spectrum is dominated by ions of fragments containing the 5′-terminus, although the hydrolysis should be equal for all phosphodiester bonds. As was true for the enzymatic digests, correct mass assignments requires the assumption that all fragments have a 2′,3′-cyclic phosphate. The distribution of peaks, therefore, resembles that obtained after a 3′-exonuclease digest (35–37). In principle, the alkaline hydrolysis alone could, therefore, be used for a complete sequencing. This is, however, only possible for quite small oligoribonucleotides, because larger fragment ions, differing in mass by only a few mass units will not be resolved in the spectra and the mass of larger ions cannot be determined with the necessary accuracy of better than 1 Da, even if peaks are partially or fully resolved.

The interpretation of the spectra particularly from digests of unknown RNA samples can in specific cases be substantially simplified, if only the fragments containing the original 5′-terminus are separated out prior to the mass spectrometric analysis. A procedure for this approach is described in the following section.

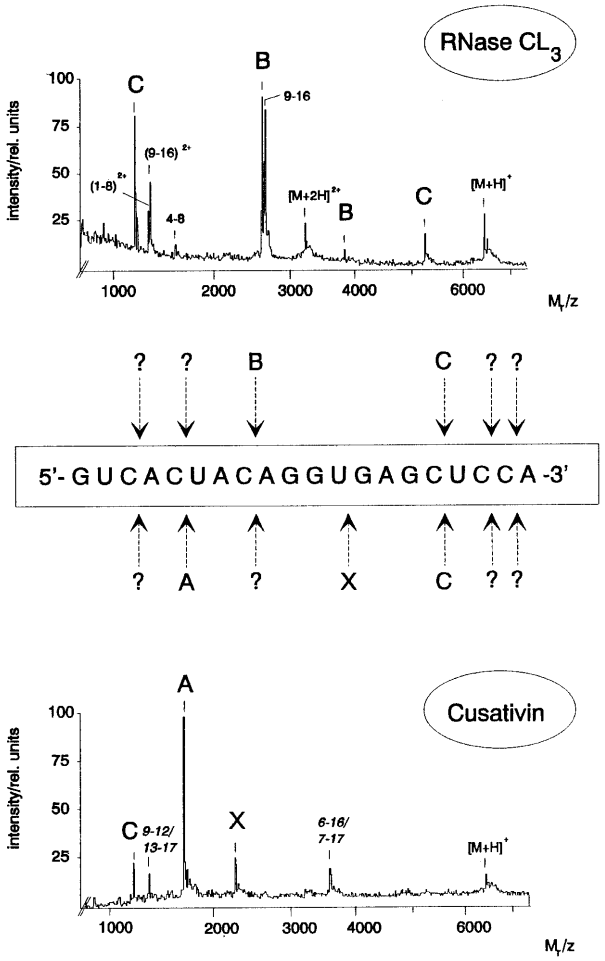

Figure 2. Investigation of the specificity of the RNases CL_3 and cusativin by positive ion UV-MALDI mass spectra of a synthetic RNA 20mer. Expected and/or observed cleavage sites are indicated by arrows. A, B, C indicate correct cleavage sites and the corresponding singly cleaved fragments. Missing cleavages are designated by a question mark (?), unspecific cleavages by an X.

Separation of 5′-biotinylated fragments

Streptavidin-coated magnetic beads (Dynal) were tested for the extraction of fragments containing the original 5′-terminus from the digests. The 5′-biotinylation of the transcript was found to not interfere measurably with the enzymatic activity of the RNases if the digestion was conducted in solution under reaction conditions, identical to those described above for the non-biotinylated synthetic oligomers. Attempts to first bind the transcripts to the beads and digest them under the same conditions on the solid phase failed because of strongly prolonged reaction times; they were not pursued further.

Surprisingly, MALDI analysis of oligomers, bound to the beads by simply incubating them in the 3-HPA matrix solution failed, even though it is well known that the biotin–streptavidin complex dissociates quantitatively when samples are prepared in solution and incubated with the matrix (38). Differerent strategies had, therefore, to be developed for an efficient binding and elution of the fragments. The specificity of binding was tested with a sample containing a mixture of two 5′-biotinylated DNA oligomers (19 nt and 27 nt) and two unlabeled DNA sequences (12 nt and 22 nt). The sample was incubated and the beads were extracted and carefully washed. Elution was then achieved by heating for 5 min at 90°C in the presence of 95% formamide. This procedure is expected to denature the streptavidin, thereby breaking the streptavidin–biotin complex (39). Clear signals of the biotinylated species and no signal of the non-biotinylated oligomers were observed in the spectra (results not shown). A slight loss of signal-to-noise ratio and mass resolution for the biotinylated species was observed nonetheless even under optimized elution conditions. Complete removal of the formamide after the elution and prior to the mass spectrometric analysis was, however, found to be important, otherwise crystallization of the matrix is disturbed. Increased suppression of nonspecific binding was reported through an addition of the detergent Tween-20 to the binding buffer (39). Although this effect could be confirmed in this study, peak broadening affected the quality of the spectra due to remaining amounts of the detergent.

For practical applications of this solid phase method to sequencing, a maximum efficiency of binding and elution of biotinylated species is of prime importance. Among a variety of conditions investigated so far, addition of salts such as EDTA gave best results in the case of DNA sequencing by providing ionic strength to the buffer (39). To examine such an effect on the solid-phase method, several salt additives were tested for the binding and elution of the 5′-biotinylated RNA *in vitro* transcript (49 nt). Judging from the relative intensity, signal-to-noise ratio, and resolution of the respective signals a 95% formamide solution containing 10 mM CDTA (adjusted to pH 8 with 25% ammonium hydroxide) turned out to be most efficient for the binding/elution (data not shown). Since CDTA acts as a chelating agent for divalent cations, formation of proper secondary and tertiary structure of the RNA is prevented. An improved sensitivity and spectral resolution has also been demonstrated under such conditions for the analysis of RNA samples by electrospray mass spectrometry (40). The improvement in the MALDI analysis is actually not very significant compared with using formamide alone but the reproducibility for spectra of good quality was substantially improved for the CDTA/formamide solution. Thus, in addition to the improved binding/elution, this additive may also improve the incorporation of the analyte into the matrix crystals. Unfortunately, a striking signal broadening on the high mass side was observed in case of formamide solutions containing EDTA, CDTA or 25% ammonium hydroxide. Since this effect is most prominent in case of 25% ammonium hydroxide and this agent was also used for adjusting EDTA and CDTA to their optimum pH, a pronounced NH_3 adduct ion formation can be assumed.

The applicability of streptavidin-coated magnetic beads for a separation in RNA sequencing was demonstrated for the RNase U_2 digest of the 5′-biotinylated RNA *in vitro* transcript (49 nt) (Fig. 3). The entire fragment pattern obtained after incubation with RNase U_2 is shown in spectrum 3a. Separation of the biotinylated 5′-fragments reduces the complexity of the spectrum (Fig. 3b). However, the signals in this spectrum are broadened and the increased number of signals in the low mass range indicate that even after stringent washing of the beads, some amounts of buffer and detergent used for the binding and elution remained. Further improvements, particularly in the elution and/or desorption of bound species are, therefore, needed.

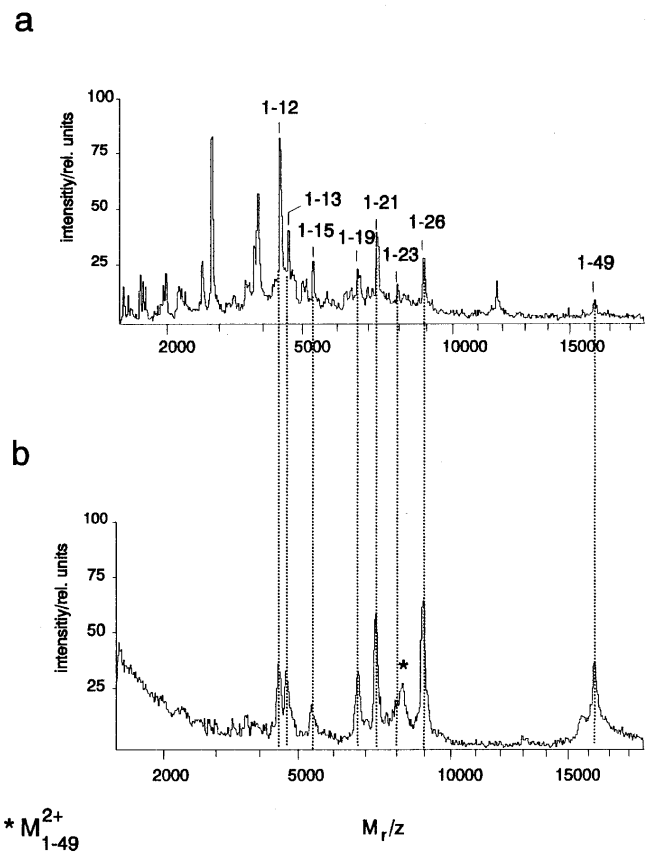

Figure 3. Positive UV-MALDI mass spectra of the 5′-biotinylated 49 nt *in vitro* transcript after RNase U_2 digest for 15 min. (**a**) Spectrum of the 25 µl assay containing ~100 pmol of the target RNA before separation; (**b**) spectrum after isolation of the 5′-biotinylated fragments with magnetic beads. Captured fragments were released by a solution of 95% formamid containing 10 mM CDTA. 1 µl aliquots of the samples were mixed with 1.5 µl matrix (3-HPA) in both cases.

DISCUSSION

RNA was chosen as the substrate for enzymatic cleavage because it exhibits higher sensitivity and better stability in MALDI-MS compared to DNA (18,28,37). Nonetheless the method can be used to obtain sequence information on DNA as well, because DNA can be transcribed into RNA prior to the analysis, as demonstrated in this study. The results also demonstrate that MALDI-MS provides an efficient technique for the direct read-out of RNA sequencing reactions based on uniform base-specific endonucleolytic cleavage by four different ribonucleases. Compared with conventional electrophoretic methods, a significantly accelerated analysis of such mixtures of RNA fragments can be achieved with MALDI-MS. Furthermore, with the present mass accuracy demonstrated in Table 2, this method allows one to unambiguously determine nearly all fragments generated in RNase digests. Limitations are incurred for fragments, which differ in mass by only a few Daltons. An additional separation of the fragment mixture may therefore be advisable in specific cases. The current limit of molecular size amenable to this method is estimated to be above 100 nucleotides. Spectra of RNA in the range of 461 nt have actually been reported (37), but the mass resolution would have to be improved considerably in this mass range, before oligoribonucleotides of this size can be subjected to a sequencing analysis as described in this paper. In addition to sequence analysis, the molecular masses of the fragments can be used for the determination of the position and possibly the identity of modified nucleotides common in rRNA and tRNA (1,2,16,37). Most probably even more importantly, endonuclease assays lend themselves to mutation analysis where the sequence is known in principle and the desired information can be obtained from but a few properly chosen fragments.

The application of mass spectrometric techniques to such sequencing analyses is likely to suffer from effects of the RNA secondary structure which affects the specificity of several ribonucleases. Preferentially RNA digestion is performed under denaturing reaction conditions (10,20,23,25), but the sample preparation of nucleic acids required for MALDI is not compatible with a direct analysis of sequencing reactions containing most of the denaturing agents.

Urea is a promising denaturing agent, because it has been demonstrated to act as a fairly efficient matrix for IR-MALDI-MS of oligonucleotides (18). A direct analysis of the reaction mixture containing 7 M urea (2 M f.c.) would, in principle, be conceivable. However, the usual concentration of urea when used as a matrix is 0.3 M. This matrix also exhibits a tendency to form multiply charged ions of the analyte, with the number of charges increasing with increasing size of the analyte (41). This will certainly complicate the interpretation of fragment spectra. A first attempt to obtain IR-MALDI spectra of fragments mixtures in 5–7 M urea were not successful so far. Further work along this or other lines to overcome the influence of RNA secondary structure on the digest is obviously desirable. If UV-MALDI is used for the analysis, additional purification is necessary, if denaturing agents of the necessary concentrations have been used for the digestion. This, however, can be achieved via a solid-phase separation, as demonstrated for DNA sequencing reactions (39,42). In addition, the use of streptavidin-coated magnetic beads in RNA sequencing can significantly simplify the complex fragment pattern generated by nuclease digestion by way of an efficient separation of biotinylated species. Introduction of a biotin tag at the 5′-terminus can easily be achieved by initiating the transcription reaction with a biotinylated dinucleotide. In contrast to mass spectrometry, detection of 5′-terminal fragments by electrophoretic separation techniques is only achievable by labeling. All non-labeled fragments (internal and those bearing the 3′-terminus) as well as the buffer and the enzyme are still present in the assay and removal is usually required before analysis (42). The solid-phase method represents a simple procedure for the efficient purification of such reactions. Since the streptavidin–biotin interaction persists under the conditions of MALDI, the strategy for the detection of the biotinylated species , presented in this study, is based on the release of the bound analytes from the support by denaturing the streptavidin with formamide. Another promising strategy for separation of selected DNA fragments, termed positional sequencing by hybridization (PSBH), has been presented by Broude and coworkers (43). In their approach duplex probes with single-stranded five base 3′ overhangs are used, immobilzed to streptavidin-coated magnetic beads. Detection of the immobilized target is achieved without elution, because double stranded species are usually denatured under MALDI conditions (44). Another promising alternative is the use of acid-labile biotin derivatives attached to the 5′-terminus. This approach is currently being pursued by the authors. Recently an UV-photocleavable

biotin (PCB) has been reported in the literature (45). Preliminary experiments by the authors have shown that a small oligonucleotide, bound to streptavidin-coated beads with this linker was released and desorbed upon irradiation with the UV-laser pulse used for the MALDI analysis (46).

The prevalence of fragments, containing the original 5′-terminus in all spectra of Figure 1, most pronounced for that after alkaline hydrolysis, is another interesting observation. This selection can hardly result from the digestion as such which will always generate both fragments in equal amounts. There must, therefore, be a selection mechanism associated with the MALDI analysis. The only feature common to all fragments containing the original 5′-terminus as well as the internal fragments for all different samples (synthetic as well as transcripts) and digests is the 2′, 3′-cyclic phosphate group. It cannot be excluded that this leads to a preferential inclusion of the fragments into the MALDI matrix. More probable, however, is a preferential charging of such fragments during desorption, even though there does not seem to be any obvious structural feature supporting such a hypothesis. It should also be mentioned in this context that signals of the intact synthetic oligonucleotides can easily be obtained at laser fluences equal to those used for the fragment analysis, even though these molecules have a 3′-hydroxyl group.

For a routine application of MALDI-MS to RNA sequencing including the solid-phase method, further improvement of the purification is required in order to remove buffer and detergent used for binding and elution of the biotinylated species entirely. Furthermore, secondary structure effects of the RNA molecules have to be minimized to ensure uniform cleavage of the RNases. Most importantly, future work will have to be directed towards a minimization of procedural steps such as washing, in order to make the method more suitable for automated DNA/RNA analysis.

ACKNOWLEDGEMENTS

The authors are grateful to Professor T. Girbes for providing the Cusativin and for the assistance concerning the enzymatic assay. Helpful discussion with Prof. H. Köster, Prof. D. Cech and Prof. K. Rothschild are also acknowledged. This work was done in partial fulfillment of the requirements of S. Hahner for the degree of Dr rer. nat. at the Westfälische-Wilhelms-Universität, Münster. Financial support of this work by the Commission of the European Communities (grant No.: GENE-CT 93-0082) is greatly appreciated.

REFERENCES

1 Kowalak, J. A., Pomerantz, S. C., Crain, P. F. and McCloskey, J. A. (1993) *Nucleic Acids Res.*, **21**, 4577–4585.
2 Lankat-Buttgereit, B., Gross, H. J. and Krupp, G. (1987) *Nucleic Acids Res.*, **15**, 7649.
3 Fichot, O. and Girard, M. (1990) *Nucleic Acids Res.*, **18**, 6162.
4 Bachellerie, J.-P. and Qu, L.-H. (1993) *Methods Enzymol.* **224**, 349–357.
5 Hahn, C.-S., Strauss, E. G. and Strauss, J. H. (1989) *Methods Enzymol.*, **180**, 121–130.
6 Lane, D. J., Pace, B., Olsen, G. J., Stahl, D. A., Sogin, M. L. and Pace, N. R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6955–6959.
7 Kramer, F. R. and Mills, D. R. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 5334–5338.
8 Bauer, J. G. (1990) *Nucleic Acids Res.*, **18**, 879–884.
9 Parvin, J. D., Smith, F. I. and Palese, P. (1986) *DNA*, **5**, 167–171.
10 Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) *Nucleic Acids Res.*, **4**, 2527–2537.
11 Frank, R. and Köster, H. (1979) *Nucleic Acids Res.*, **6**, 2069–2087.
12 Li, S., Haces, A., Stupar, L., Gebeyehu, G. and Pless, R. C. (1993) *Nucleic Acids Res.*, **21**, 2709–2714.
13 Hillenkamp, F., Karas, M., Beavis, R. C. and Chait, B. T. (1991) *Anal. Chem.*, **63**, 1193A–1202A.
14 Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1983) *Science*, **246**, 64–70.
15 Nordhoff, E., Kirpekar, F. and Roepstorff, P. (1997) *Mass Spectrom. Rev.*, **15**, 67–138.
16 Kowalak, J. A., Bruenger, E. and McCloskey, J. A. (1995) *J. Biol. Chem.*, **270**, 17758–17764.
17 Doktycz, M. J., Hurst, G. B., Habibi-Goudarzi, S., McLuckey, S. A., Tang, K., Chen, C. H., Uziel, M., Jacobson, K. B., Woychik, R. P. and Buchanan, M. V. (1995) *Anal. Biochem.*, **230**, 205–214.
18 Little, D. P., Thannhauser, T. W. and McLafferty, F. W. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 2318–2322.
19 Nordhoff, E., Ingendoh, A., Cramer, C., Overberg, A., Stahl, B., Karas, M., Hillenkamp, F. and Crain, P. (1992) *Rapid Commun. Mass Spectrom.*, **6**, 771–776.
20 Simoncsits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. and Guilley, H. (1977) *Nature*, **269**, 833–836.
21 Gupta, R. C. and Randerath, K. (1977) *Nucleic Acids Res.*, **4**, 1957–1978.
22 Kuchino, Y. and Nishimura, S. (1989) *Methods Enzymol.*, **180**, 154–163.
23 Boguski, M. S., Hieter, P. A. and Levy, C. C. (1980) *J. Biol. Chem.*, **255**, 2160–2163.
24 Rojo, M. A., Arias, F. J., Iglesias, R., Ferreras, J. M., Muñoz, R., Escarmis, C., Soriano, F., López-Fando, J., Méndez, E. and Girbés, T. (1994) *Planta*, **194**, 328–338.
25 Donis-Keller, H. (1980) *Nucleic Acids Res.*, **8**, 3133–3142.
26 Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1760–1764.
27 Waldmann, R., Gross, H. J. and Krupp, G. (1987) *Nucleic Acids Res.*, **15**, 7209.
28 Nordhoff, E., Cramer, R., Karas, M., Hillenkamp, F., Kirpekar, F., Kristiansen, K. and Roepstorff, P. (1993) *Nucleic Acids Res.*, **21**, 3347–3357.
29 Shaler, T. A., Wickham, J. N., Sannes, K. A., Wu, K. J. and Becker, C. H. (1996) *Anal. Chem.*, **68**, 576–579.
30 Richards, F. M. and Wyckoff, H. W. (1971) In Boyer, P. D.(ed.) *The Enzymes* Vol. 4, 3rd Ed. Academic Press, New York, pp. 746–806.
31 Heinemann, U. and Saenger, W. (1985) *Pure Appl. Chem.*, **57**, 417–422.
32 Ikehara, M., Satoshi, N., Morioka, H., Kim, H. J., Fuchimura, K., Kimura, T., Adiwinata, J., Tanaka, Ohtsuka, E. and Uesugi, S. (1987) *Pure Appl. Chem.*, **59**, 965–968.
33 Breslow, R. and Xu, R. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 1201–1207.
34 Hahner, S., Kirpekar, F., Nordhoff, E., Kristiansen, K., Roepstorff, P. and Hillenkamp, F. (1996) *Proceedings of th 44th ASMS Conference on Mass Spectrometry*, Portland, Oregon, p. 938.
35 Pieles, U., Zürcher, W., Schär, M. and Moser, H. E. (1993) *Nucleic Acids Res.*, **21**, 3191–3196.
36 Nordhoff, E., Hahner, S., Karas, M., Hillenkamp, F., Kirpekar, F., Kristiansen, K. and Roepstorff, P. (1993) Book of Abstracts, *13th Internat. Mass Spectrom. Conf.*, Budapest, p. 218.
37 Kirpekar, F., Nordhoff, E., Kristiansen, K., Roepstorff, P., Lezius, A., Hahner, S., Karas, M. and Hillenkamp, F. (1994) *Nucleic Acids Res.*, **22**, 3866–3870.
38 Hillenkamp, F. (1997) *New Methods in the Study of Molecular Aggregates*, Standing, K. G. (ed.), NATO Advanced Research Workshop Series, Kluwer Academic Pubs., in press.
39 Tong, X. and Smith, L. M. (1992) *Anal. Chem.*, **64**, 2672–2677.
40 Limbach, P. A., Crain, P. F. and McCloskey, J. A. (1995) *J. Am. Soc. Mass Spectrom.*, **6**, 27–39.
41 Overberg, A., Karas, M. and Hillenkamp, F. (1990) *Rapid Comm. Mass Spectrom.*, **5**, 128.
42 Hultman, T., Stahl, S. and Uhlen, M. (1989) *Nucleic Acids Res.*, **17**, 4937–4946.
43 Broude, N. E., Sano, T., Smith, C. L. and Cantor, C. R. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3072–3076.
44 Tang, K., Fu, D., Kötter, S., Cotter, R. J., Cantor, C. R. and Köster, H. (1995) *Nucleic Acids Res.*, **23**, 3126–3131.
45 Olejnik, J., Krzymanska-Olejnik and Rothschild, K. J. (1996) *Nucleic Acids Res.*, **24**, 361–366.
46 Hahner, S., Lüdemann, H.-C., Olejnik, J., Krzymanska-Olejnik, E., Rothschild, K. J. and Hillenkamp, F. (1997) 45th Annual Conference on Mass Spectrometry amd Allied Topics, Palm Springs, CA, June 1–5. Proceedings of extended abstracts, to be published.