Nucleic Acids Research, 1993, Vol. 21, No. 15 3347–3357

Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry

E.Nordhoff, R.Cramer, M.Karas, F.Hillenkamp, F.Kirpekar[1], K.Kristiansen[1] and P.Roepstorff[1]
Institute for Medical Physics and Biophysics, University of Münster, Robert-Koch-Strasse 31, 4400 Münster, Germany and [1]Department of Molecular Biology, University of Odense, Campusvej 55, 5230 Odense M., Denmark

Received May 28, 1993; Accepted June 21, 1993

ABSTRACT

Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) with infrared laser light of a wavelength of 2.94 μm has been used for the analysis of nucleic acids. Spectra of oligodeoxynucleotides up to 26 nucleotides, oligothymidylic acids up to 100 nucleotides as well as different synthetic RNA oligomers and RNA transcripts up to 104 nucleotides are presented. A main problem in the analysis of oligodeoxynucleotides was found to be related to the loss of bases. The stability of oligothymidylic acids as opposed to oligodeoxynucleotides containing all four bases indicates that the loss of bases is correlated with A, C and G protonation which decreases the stability of the N-glycosidic bond. Experiments indicate that the breakage of the N-glycosidic bond probably occurs during the desorption process due to proton transfer from the phosphodiester groups to the ionizable bases. RNA displayed a significantly higher stability in MALDI-MS due to the presence of a 2'-OH group. Consequently, signals of RNA transcripts with a length of up to 142 nucleotides could be detected by MALDI-MS. Technical details of the method, including the distribution of positive counterions on the phosphodiester backbone, the upper mass limit and mass accuracy are discussed along with a number of potential analytical applications.

INTRODUCTION

Whereas the matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) has found broad applications for the analysis of oligopeptides and proteins[1–6], investigations on nucleic acids are limited so far. Since its introduction in 1988 by Karas and Hillenkamp[1], MALDI-MS has become an established method for desorbing ions of proteins with molecular weights up to 500 kDa. The determination of molecular masses with an accuracy of up to 0.01% is possible. Typical sample amounts required are 1 pmol, but a sensitivity in the low femtomol range can be achieved[7,8]. The key idea of this technique is to embed the macromolecules (the analyte) in a suitable matrix of molecules having a strong absorption at the laser wavelength and present in high molar excess over the analyte. This induces an efficient transfer of the laser-pulse energy to the analyte and results in a soft desorption process. It is suggested that the matrix molecules also play a role in the ionization of the analyte molecules.[9] Different lasers and wavelengths in the UV (mainly 266 nm, 337 nm and 355 nm) and IR (mainly 2.94 μm, 2.79 μm and 10.6μm) have been used for this technique. Complex mixtures of proteins and peptides are routinely mass analysed. Carbohydrates[10], glycoproteins, lipoproteins and lipids, as well as many other synthetic and natural polymers[11], are also amenable to MALDI-MS. In all of these cases the key to successful analysis of a new class of analyte molecules has been to find a specifically suited combination of matrix and laser wavelength.

Provided that the right conditions for desorption of nucleic acids of >100 nucleotides could be found, rapid DNA sequencing would be possible by direct mass spectrometric analysis of the Sanger reaction products[12], replacing the time consuming gel electrophoretic procedure. In many other fields of DNA research such as antisense DNA applications a quick and easy analysis of oligonucleotides of 10 to 30 bases would also be of high interest. Typical modifications such as labelling with a fluorescence marker, the addition of a hydrophobic anchor or the introduction of thiophosphate groups, etc. could be verified by a determination of the accurate molecular mass. Among the limited number of results, so far reported for MALDI-MS of oligonucleotides, are analyses of oligothymidylic acids, $pd[T]_8$ to $pd[T]_{30}$, and mixtures of them[13–21], of oligodeoxynucleotides (containing all four different bases)[13,17,19,20,22,23] and of RNA samples. Karas et al. have reported spectra of tRNA[13,17,20], and Williams et al. have been able to obtain spectra of oligodeoxynucleotide samples of up to 60 nucleotides[24]. Recently, Becker and co-workers[25] have published mass spectra of oligodeoxynucleotides up to a 67 mer.

In a recent publication[20], we have compared oligodeoxynucleotide analysis with ultraviolet (UV)-MALDI with the 337 nm wavelength of a N_2-Laser and a 2-aminobenzoic acid matrix to that of infrared (IR)-MALDI with the 2.94 μm wavelength of an Er:YAG Laser and succinic acid as matrix and found the latter combination to give superior results. In the meantime,

further tests have shown that the combination of 3-hydroxypicolinic acid as a matrix and a wavelength of 337 nm also yields very good results, in agreement with the finding of other groups[19,25]. In this paper, results on the IR-MALDI analysis of a series of synthetic oligodeoxynucleotides in the size range from 12 to 26 nucleotides as well as for some synthetic RNA oligomers up to 19 nucleotides and RNA-transcripts up to 104 nucleotides are reported. The aim of the study was to find optimal conditions for sample preparation and to prove the feasibility of IR-MALDI-MS for nucleic acid analysis.

EXPERIMENTAL

Instruments

The mass spectrometer used for the analysis of the oligonucleotides is a home built reflectron time-of-flight instrument (RTOF, Fig. 1)[20]. The system is equipped with an electro optically Q-switched Er:YAG laser (1-2-3 Schwartz Electro-Optics, Orlando, FL, USA) emitting 150 ns pulses at a wavelength of 2.94 μm. The laser beam is focused onto the sample surface under 45° to a spot diameter of ca. 100 μm. The laser irradiance varied between 10^6 and 10^7 W/cm^2. The ions are accelerated to an energy of 12 keV and detected by a secondary electron multiplier (EMI 9643) equipped with a separately mounted dynode for ion-ion conversion. An instrumental mass resolution at up to 1000 (m/Δm, Δm: full width at the half maximum (FWHM) of the peak) can be obtained with this instrument for ions below ca. 5000−6000 Da under optimal conditions. Above 10,000 Da, the mass resolution is reduced to a value of <200 due to both the ion detection process (conversion of the analyte ion to secondary ions) and the ion formation process. The analog detector signal is digitized by a transient recorder (LeCroy 9400) at time intervals of 10 or 20 ns. Further data processing was done on a PC-AT using in-house generated software. Generally, single shot spectra clearly show the molecular ion signals. To improve the signal-to-noise ratio, 10 to 20 single shot spectra were usually accumulated and averaged.

Sample preparation

Matrix. The matrix, succinic acid, was dissolved in ultra pure water to a concentration of ca. 50 g/l and then carefully desalted on an acid (0.1 M HCl (p.a.) solution) activated cation exchange column (cation exchange polymer: BioRad, 50W-X8, mesh size 100−200 μm).

Subsequently, the purified matrix solution was diluted to a concentration of 20 g/l, aliquoted in 50 μl portions and stored at −20°C. To reduce salt contaminations by repeated handling (alkali salt traces on the pipette tips, etc.), a new portions was used every day.

Oligodeoxynucleotides. The oligodeoxynucleotides were synthesized and deprotected by Dr. Otto Dahl, University of Copenhagen, Denmark.

Prior to mass spectrometry, the oligodeoxynucleotides were purified on the Pharmacia SMART™ FPLC system using a reversed phase column (C2/C18 SC2/10):

Buffer A: 5% acetonitrile/10 mM triethylammonium acetate (TEAA)
Buffer B: 30% acetonitrile/10 mM TEAA
Gradient: 0−5 min: 100% A. 5−30 min: 0−100% B. 30−35 min: 100% B. 35−40 min: 100−0% B. 40−45 min: 100% A
Flow rate: 50 μl/min

Oligoribonucleotides. Synthetic oligoribonucleotides as well as RNA oligomers from in vitro RNA polymerase transcription reactions and also from limited degradation of polyuridylic acids were used.

1. Synthetic RNA. The synthetic oligoribonucleotides were synthesized and HPLC purified by Dr. Jens Peter Fürste, Freie Universität of Berlin, Germany. These samples were used as supplied.

2. In vitro transcripts. Aliquots of the plasmid pBluescript KS$^+$ were linearized with the restriction endonucleases *Hin*dIII, *Eco*RI and *Not*I, which will generate transcripts of 55 nucleotides, 67 nucleotides, and 104 nucleotides, respectively, in an in vitro transcription system using T3 RNA polymerase. Also, the plasmid pBluescript SK$^+$ was digested with the restriction enzymes *Hin*fI, which will generate a transcript of 142 nucleotides in the above mentioned transcription system. The plasmids were subsequently extracted with phenol/chloroform and precipitated with ethanol/Na-acetate.

In vitro transcription was performed in 100 μl containing 40 mM Tris−HCl (pH 7.5), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, 100 units RNAsin (Promega Biotec), 0.5 mM each of ATP, CTP, GTP and UTP, 5 μg of linearized plasmid and 50 units of T3 RNA polymerase (Stratagene). The reaction was incubated at 37°C for one hour, another aliquot of 50 units T3 RNA polymerase was added and incubation was continued for one more hour. The reaction mixtures were extracted once with phenol/chloroform, precipitated with ethanol/NH_4-acetate and dissolved in 15 μl of 7 M urea.

The transcripts were electrophoresed through a 5% polyacrylamide sequencing gel containing 8 M urea and visualised on a fluorescent thin layer chromatographic plate under 254 nm irradiation and subsequently cut out of the gel. RNA was extracted from the gel plug by gentle shaking in 200 μl 10 mM Tris-HCl (pH 8.0), 20 mM EDTA, 0.5% SDS and 50 μg/ml proteinase K at 30°C for 16 hours; the gel plug was washed with 100 μl TE-buffer and the pooled liquid was extracted once with chloroform. RNA was precipitated twice with 2 M NH_4-acetate and 3 volumes of ethanol.

In addition to the preparative transcription reactions, analytical reactions of 1/10 volume supplemented with 8 μM of [α-^{32}P]-UTP (800 Ci/mmol) were made and purified in parallel. The purified analytical samples were tested on sequencing gels to estimate the heterogeneity of the preparative samples.

3. Oligouridylic acids. Poly-U from Sigma was partially hydrolyzed at a concentration of 10 μg/μl by incubation in 40 mM $NaHCO_3$, 60 mM Na_2CO_3 (pH 10.0) at 60°C for 1 hour. 0.25 vol. of 2M NH_4-acetate was added and the poly-U was precipitated with 3 vol. of ethanol.

Sample preparation for IR-MALDI-MS

Samples were lyophilized over night to remove excess ammonium salts and redissolved in ultra pure water to a final concentration of 0.1 to 0.01 g/l. Aliquots of 0.5 to 1 μl of analyte solution and 1 to 2 μl of matrix solution were mixed on a flat metallic

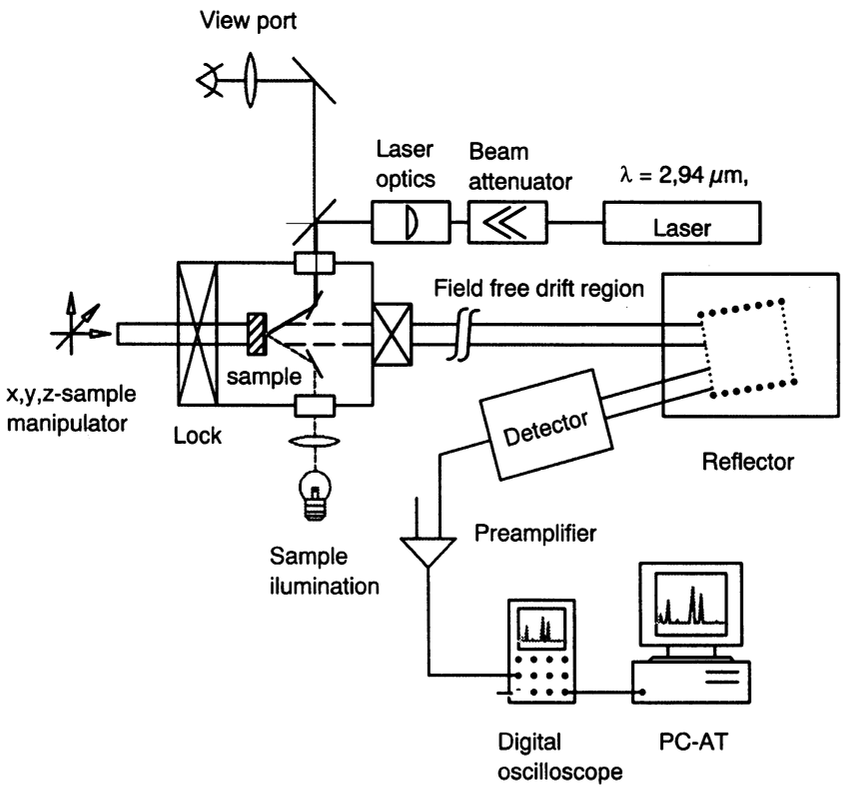

Figure 1. Schematics of the reflectron time of flight (RTOF) mass spectrometer used for IR-MALDI-MS.

sample support and dried in a stream of cold air. Alkali cations present in the sample solution and on the sample support were removed with NH_4^+- loaded cation exchange polymer beads (BioRad; 50W-X8, mesh size: 100−200 μm, activated with a saturated ammonium acetate (p.a.) solution) as follows: After the matrix solution was placed onto the sample support, 0.5 μl suspension of the cation exchange material in ultra pure water, containing ten to twenty polymer beads, were pipetted to the sample droplet before the analyte solution was added. These beads do not interfere with the mass analysis. Upon drying of the sample, the beads aggregate in the centre of the sample, leaving the outer rim of the specimen undisturbed. Accordingly, the best spectra are usually obtained from this area (Fig. 2).

Acidic hydrolysis of a oligodeoxynucleotide

10 pmoles of the oligodeoxynucleotide 41-mer 5′-CGG AAA ACT TTT CGT TTG TAT TTT ATT TGT ATT TTA GAC AT-3′ were end-labelled at the 5′-end using [γ-^{32}P]-ATP and polynucleotide kinase as described[26], precipitated and dissolved in 30 μl of ultra pure water. Two aliquots (samples A and B) of 5 μl were supplemented with 15 μl of 20 g/l succinic acid and dried down, while two aliquots (samples C and D) of 5 μl were dried down with 15 μl of water. The samples were dissolved in 200 μl of 2M NH_4-acetate and precipitated with 700 μl of ethanol and 30 μg of carrier herring sperm DNA. Samples B and D were dissolved in a formamide/dye mix. To cleave the oligodeoxynucleotide phosphodiester backbone at abasic sites, samples A and C were dissolved in 90 μl of 1M piperidine and incubated at 90°C for 30 minutes, precipitated with 1.2 ml of n-butanol, redissolved in water, lyophilized and dissolved in formamide/dye mix. The samples were electrophoresed on a 10% sequencing gel alongside a purine sequencing reaction[27] performed on 10 μl of the end-labelled oligodeoxynucleotide.

RESULTS AND DISCUSSION

General features of MALDI spectra of nucleic acids

All spectra reported here are negative ion spectra. As has been reported by other investigators as well[16,19,21−25], oligonucleotides yield stronger signals and better mass resolution in the negative ion mode as compared to the positive one for all tested combinations of matrices and wavelengths. Fig. 3a shows a typical MALDI mass spectrum of a 12 mer oligodeoxynucleotide. Besides a strong signal of the singly-charged, deprotonated molecular ion $(M-H)^-$, signals of singly-charged multimers $(nM-H)^-$ as well as momomers and multimers of higher charge states $(nM-mH)^m$ are observed in the spectrum with lower intensity. Signals of matrix ions are present in the lower mass range, normally up to 500 Da. Under optimal conditions these signals are comparable in magnitude or even lower than the analyte signals, despite the high molar excess of the matrix. Minor signals of different alkali salts of the phosphodiester groups are also observed, even though the matrix and the analyte solutions had been carefully desalted before deposition on the sample support. The cleavage of at least one N-glycosidic bond is observed as well. Under optimized sample preparation conditions (see below), a mass resolution of over 800 (FWHM) has been achieved for IR-MALDI-MS of RNA- and DNA-oligomers (see Fig. 4c and 8a). A Mass accuracy of typically 0.01−0.02% was obtained for oligonucleotides with up to 26 bases if well-defined

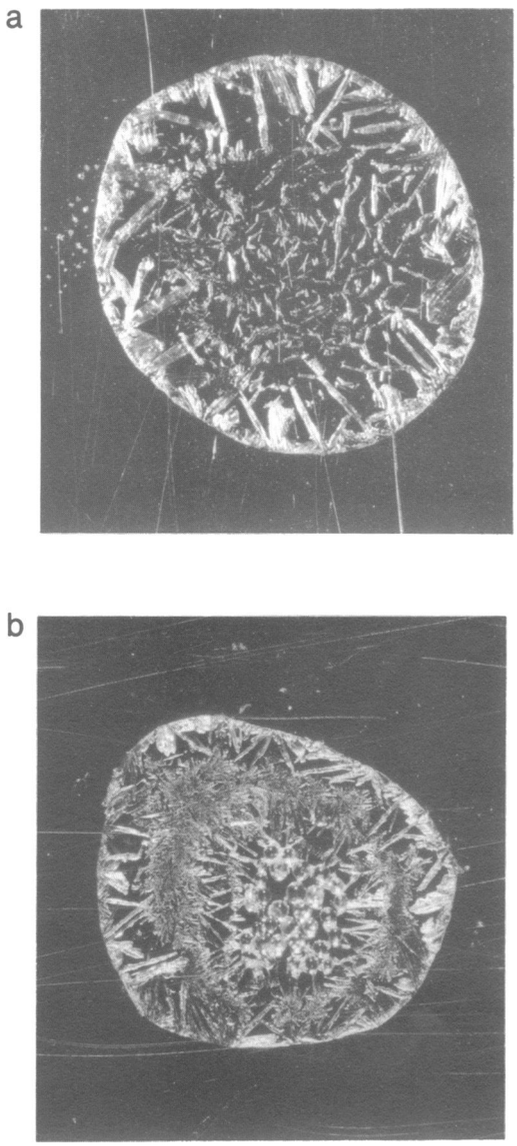

Figure 2. Micrographs of the samples for IR-MALDI-MS analysis: **(a)** without and **(b)** with added cation exchange polymer beads. The diameter of the sample spot is ca. 3–4 mm. The best mass spectra are usually obtained from the crystalline rim of the sample, free of polymer beads.

standards were used for calibration. Oligothymidylic acids are best suited as mass standards, because of their high stability and homogeneity (Fig. 3b).

Analysis of DNA

Salt formation. Figure 4 demonstrates the effect of different counter ions to the negatively-charged phosphodiester groups of nucleic acids for the oligodeoxynucleotide 12 mer 5′-CCT TTT GAA AAG-3′. Figure 4a shows the mass spectrum for a sample identical to that of Fig. 3a, i.e. carefully desalted, but with the addition of 20 mM of KBr. The major peak distribution represents signals of the potassium salts with a varying number of potassium ions up to the maximum possible number of ten (one phosphate group without a counterion is, of course, necessary to provide the negative charge of the ion). A corresponding distribution of lower intensity of sodium salts, present as an impurity, is also seen. Similar distributions have been observed for a variety of other metal counterions such as Na^+, Cs^+, Ca^{2+}, Mg^{2+}, Fe^{2+}/Fe^{3+} and Al^{3+}, where the multivalent ions balance an equivalent number of negative countercharges of phosphodiester groups. Such a chemical heterogeneity of the detected ions of a single given oligonucleotide is detrimental to the analysis for several reasons. Firstly, the different peaks of the distribution will not be resolved for higher ion masses given by the limited mass resolution of the method. This would not only result in broad peaks and a decreased accuracy in mass determination, it would also prevent the resolution of the single components in a mixture of oligonucleotides of different sizes as encountered in sequencing applications. Also, the total available charge would get distributed among a very large number of ions, thereby decreasing the signal of each of them, quickly decreasing the signal to noise ratio to values below the detection limit.

Fig. 4b shows a spectrum of the same sample, except that 20 mM of ammonium acetate had been added instead of a metal salt. No peak distribution due to a counterion distribution is observed under these conditions The same has been observed for the presence of other ammonium ions such as $HNMe_3^+$ and $HNEt_3^+$. The most likely explanation for this observation is a dissociation of all ammonium phosphate ion pairs, formed in solution, into the free acid form of the phosphodiester group and the amine via proton transfer during the desorption process. Such a neutralization reaction by dissociation is not possible for alkali phosphate ion pairs. Minor signals of Na- and K-salts are still present in the spectrum of Fig. 4b. These observations have lead to the sample preparation procedure of carefully desalting the solutions and adding a few ion exchange beads to the sample as described in the experimental section. Fig. 4c shows a spectrum of the same sample prepared under such optimized conditions. Contributions from alkali salts are no longer detectable and a mass resolution of 800 is obtained.

It is interesting to note that under identical conditions multiple salt formation is most pronounced for oligothymidylic acids as compared to any other oligodeoxynucleotides containing other bases besides thymine (see Fig. 3b and ref. 20) This is thought to be due to the fact that thymidine (and uridine) in contrast to all other nucleotides can not form zwitterions by protonation of the base. Thus, neutralisation of the negatively-charged phosphodiester groups by salt formation is more likely to occur with oligothymidylic acids than with oligodeoxynucleotides containing bases which by protonation allow zwitterion formation.

Fragmentation. The mass spectra 3a, 3b, 4b and 4c all show small signals of fragment ions that originate from the loss of one or two bases. The frequency of loss of G, A and C is observed in that sequence. No loss of T is observed. It is also worth noting that the extent of base loss is significantly higher in the presence of salt formation (Fig. 3a and 4b) than in the absence of it (Fig. 4c). In the spectrum shown in figure 4a the fragment signals are mainly buried under the broad salt distribution. Figure 5 shows mass spectra of a set of oligodeoxynucleotides with increasing sequence length (16 mer, 22 mer, and 26 mer). Besides a decrease in signal to noise ratio, indicative of a decreasing signal intensity of the corresponding analyte ions with increasing ion mass, an increasing fragmentation of the analyte molecules is observed. Both effects lead to an upper limit of ca. 26 nucleotides for a proper MALDI-analysis under the current experimental conditions; above this size the $(M-H)^-$-signal of the

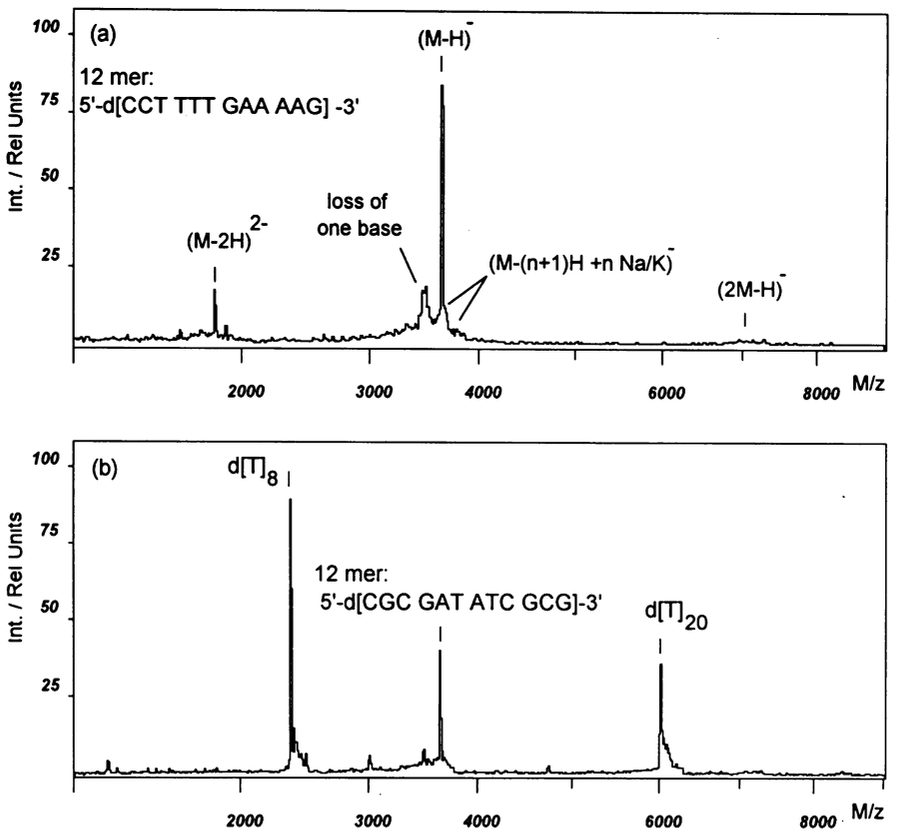

Figure 3. IR-MALDI mass spectra of **(a)** the oligodeoxynucleotide 12-mer 5′-CCT TTT GAA AAG-3′; total sample load: 5 pmoles; and **(b)** the oligodeoxynucleotide 12-mer 5′-CGC GAT ATC GCG-3′; 2 pmoles loaded; 2 pmoles of $d[T]_8$ and $d[T]_{20}$ had been added as internal mass calibration standards. Both spectra are the sum of 10 single-shot spectra.

unfragmented ion cannot be resolved any longer. Though ions of oligodeoxynucleotides of up to ca. 45 nucleotides can be detected, their broad unresolved peaks render an accurate mass determination impossible; peak maxima will systematically give a too low mass. The observed fragmentation can mostly be attributed to single and multiple losses of bases. Cleavage of the deoxyribose phosphate backbone occurs much less frequently.

Mass differences observed for the loss of a single base correspond to the molecular mass of the corresponding free base (111 Da for C, 135 Da for A, and 151 Da for G). This reaction could, in principle, result either from an acid-catalysed base elimination during sample preparation or from a laser induced fragmentation during the desorption process.

To investigate whether the succinic acid matrix causes depurination of the samples in solution, a oligodeoxynucleotide 41-mer was prepared and dried down under conditions identical to those used for MALDI samples. The phosphodiester backbone was then cleaved with piperidine at depurinated sites and the product analysed by electrophoresis on a sequencing gel (Fig. 6, lane 1). Differences between the lane 1 (succinic acid plus piperidine), lane 2 (succinic acid, no piperidine), lane 3 (piperidine without succinic acid), and lane 4 (neither succinic acid nor piperidine) are only marginal. These observations, including the fact that more than 95% of the sample remained intact regardless of the treatment, suggest that the depurination takes place during the laser desorption/ionization process rather than during the sample preparation.

The mass spectra suggest that the stability of the N- glycosidic bonds of the different nucleotides drops dramatically from thymidine, where no loss is seen, to cytidine and further on to adenine and guanine. This is in good agreement with the acid stability of the N-glycosidic bond linking the different bases to the deoxyribose ring of DNA[28]. For the laser-induced fragmentation, this similarity can be explained by a weakening of the N-glycosidic bond following base protonation by either the succinic acid matrix (pH=2 for the solution used) or by the analyte molecules themselves. The latter can occur during or after the desorption process via proton transfer from the strongly acidic phosphodiester groups to their corresponding bases (pKa=1 for $HOPO(OR)_2$; pKa=4.2, pKa=3.5, and pKa=2.1, respectively, for the protonated cytidine, adenine, and guanine moieties). Though self-protonation is rather unlikely, as long as all nucleotides of the polymer are present as ammonium salts , proton transfer could be expected following the dissociation of the ammonium phosphodiester groups into the free acids and bases upon desorption into the gas phase. Since thymidine cannot be protonated such a weakening of the N-glycosidic bond is not possible. The concept of a self-protonation rather than protonation by the matrix is supported by the observation that similar fragmentation patterns are observed if the neutral matrix, urea, is used for the desorption instead of the succinic acid.[20]

Figure 7 shows two mass spectra of oligothymidylic acids with higher molecular masses (**7a**: $(dT)_{70}$, and **7b**: $(dT)_{100}$). These spectra serve as another demonstration for the above model,

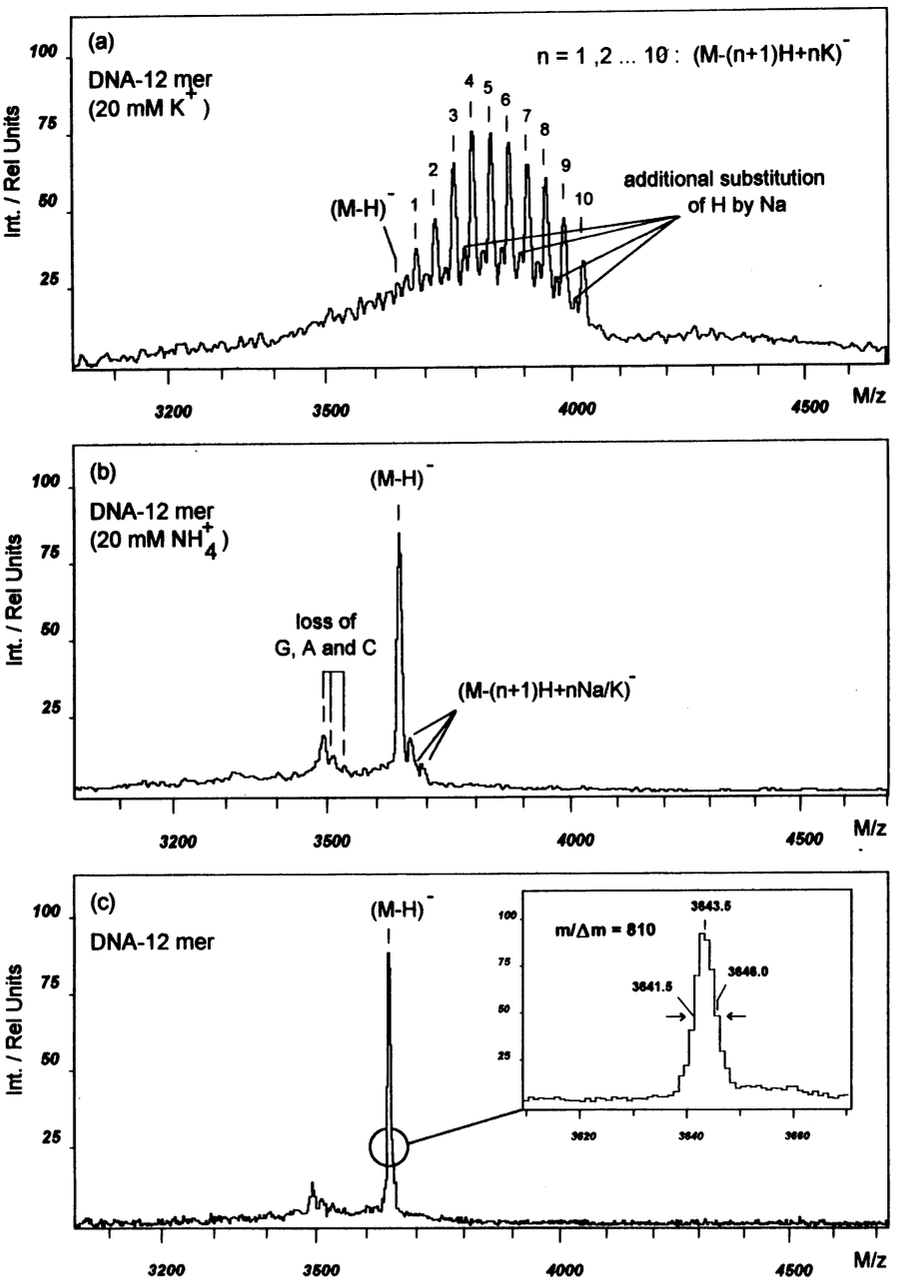

Figure 4. IR-MALDI mass spectra of the oligodeoxynucleotide 12-mer 5′-CCT TTT GAA AAG-3′ under three different conditions: in the presence of **(a)** 20 mM potassium bromide or **(b)** 20 mM ammonium acetate in the analyte solution; **(c)** 15–20 cation exchange polymer beads, but without the addition of salt. 5 pmoles oligodeoxynucleotide were loaded for each measurement. All spectra are the sums of 10 single-shot spectra.

predicting that oligothymidylic acids should be much more stable under MALDI analysis than oligodeoxynucleotides containing all four different bases. Up to the 30 mer level, the oligothymidylic acids can be measured with a mass resolution of 800 to 1000 and a mass accuracy of 0.02% can easily be obtained. Above this molecular size, mass resolution drops to values of 170–200 because of the instrument limitations discussed above. These spectra also demonstrate that MALDI analysis of oligodeoxynucleotides so far seems to be limited to small polymers of only a few tens of bases, due to the elimination of bases.

Analysis of RNA

Because base cleavage is the major problem in IR-MALDI analysis of DNA, we expect RNA to give better results, due to the much higher stability of the N-glycosidic bonds in ribonucleotides caused by the 2′-OH group in trans-position to the base of the nucleotide (the acid-catalyzed hydrolysis rate of the N-glycosidic bond for deoxyguanosine[29] and deoxyadenosine[30] is ca. 520 and ca. 1200 times higher, respectively, than for their ribonucleoside counterparts).

Figure 8 shows mass spectra of a synthetic RNA 12 mer and

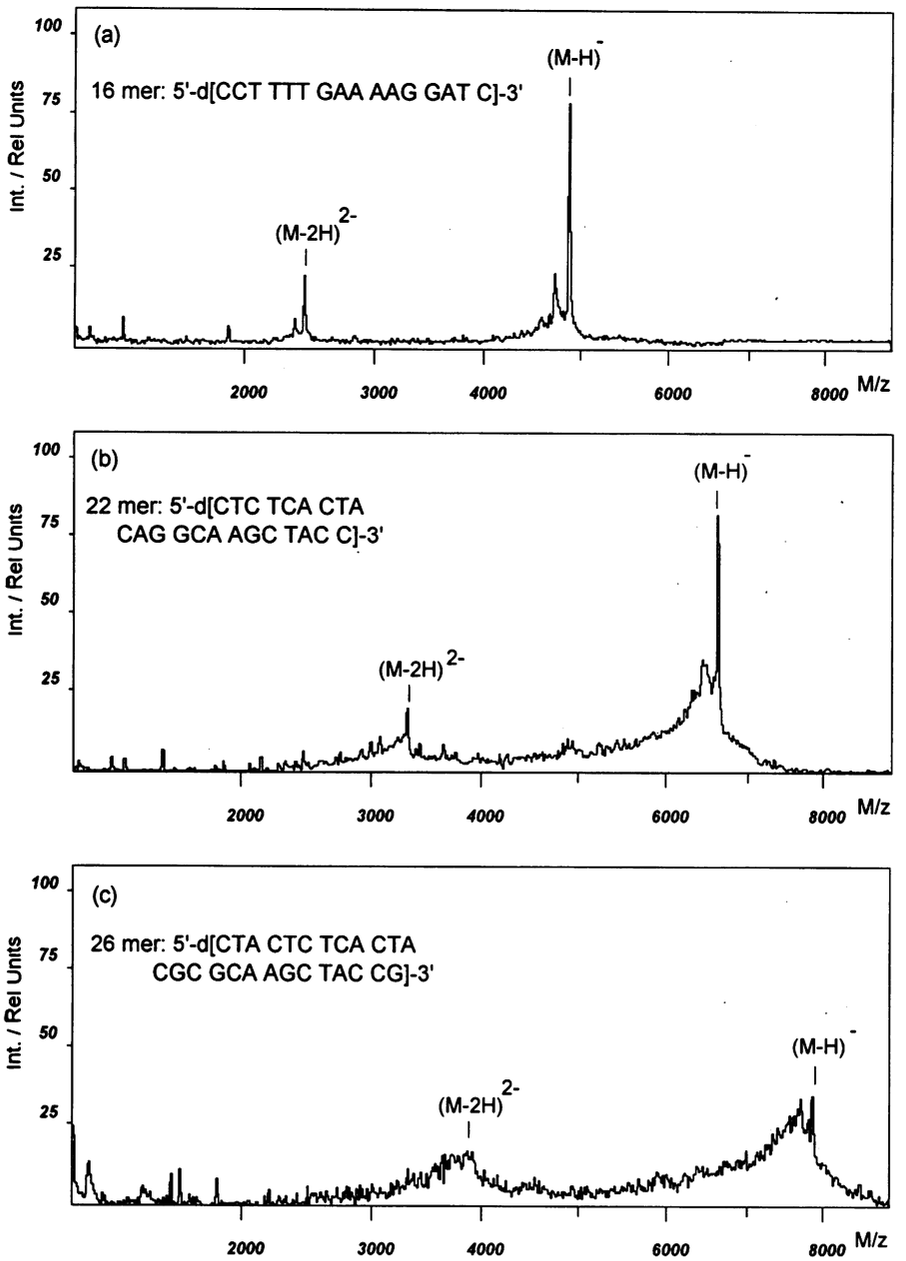

Figure 5. IR-MALDI mass spectra of oligodeoxynucleotides of increasing sequence length: **(a)** a 16-mer, **(b)** a 22-mer and **(c)** a 26-mer. 5 pmoles oligodeoxynucleotide were loaded for each measurement, all spectra are the sums of 10 single-shot spectra.

a 19 mer. The loss of bases is significantly less for these RNA molecules in comparison to that of oligodeoxynucleotides of comparable size (Fig. 5). Signal intensities and peak widths are equal or slightly better than for the oligodeoxynucleotide samples. To explore the upper mass limit for a successful desorption of RNA molecules a set of defined RNA transcripts (55, 67, 104, and 142 nucleotides long) was prepared. In Figure 9, two IR-MALDI mass spectra of such RNA-transcripts (55 mer and 104 mer) are shown. In contrast to oligodeoxynucleotides, signals from large RNA molecules are easily obtained. So far we have obtained reproducible signals for samples of up to 142 nucleotides (data not shown), but further improvements, extending the upper mass limit beyond this, can be expected.

Differences in secondary structure could also be considered to cause the different behaviour of DNA and RNA. However, both must be assumed to be denatured at the low pH of 2.0 of the succinic acid matrix. This pH is below the pH at which the Watson-Crick base paring is disrupted and DNA[31–33] as well as RNA[34] denature. Support for this notion comes from the observation that the self complementary oligodeoxynucleotide 5'-CGC GAT ATC GCG-3', which by NMR studies has been shown to be mainly double stranded at neutral pH[35], is desorbed/ionized exclusively as the single stranded monomer (Fig. 3b).

Signal intensity of the RNA ions is comparable to those of the $(dT)_{70}$-mer, but the peak widths are significantly larger (ca.

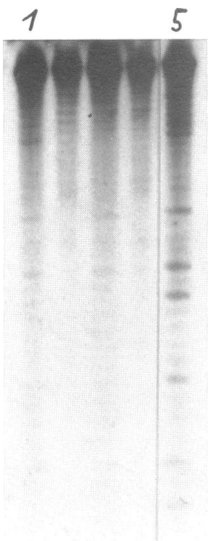

Figure 6. Electrophoretic separation of the 5'-end-labelled oligodeoxynucleotide 41-mer: 5'-CGG AAA ACT TTT CGT TTG TAT TTT ATT TGT ATT TTA GAC AT-3' after succinic acid treatment, performed as during IR-MALDI sample preparation, followed by cleavage of the phosphodiester backbone with piperidine at depurinated sites **(lane 1)**, after succinic acid treatment without further cleavage reactions **(lane 2)**, after cleavage with piperidine without prior succinic acid treatment **(lane 3)**, neither succinic acid treatment nor cleavage with piperidine **(lane 4)** and after a purine sequencing reaction **(lane 5)**.

240–270 Da for the 55 mer, ca. 800–900 Da for the 67 mer, 1100–1200 Da for the 104 mer, and above 1500 Da for the 142 mer). These peak widths correspond to apparent mass resolutions of 70 to 25 (FWHM), which is significantly short of the instrumental limits. Sample inhomogeneity is thought to be the main reason for the observed peak broadening of the higher molecular mass RNA-molecules. Analysis of the purified RNA transcripts on denaturating gels showed that all transcripts still contained considerable amounts of the corresponding n−1 mer species (ca. 30% for the 104 mer and 142 mer and ca. 5–10% for the 55 mer and 67 mer). Partial hydrolysis of the 5'-end triphosphate group would also add to the inhomogeneity by giving rise to three different molecules each differing by 80 Da. Such mass differences are not resolved above 10,000 Da leading to a substantial peak broadening. The mass difference of 161 Da between the measured molecular mass of the 55 mer transcript (17,843 Da) and the calculated value (18,004 Da) can accordingly be explained by the loss of two phosphate groups (Dm=160 Da). Also, the observed shoulder of the analyte-ion-peak towards higher masses can partly be explained by the presence of diphosphate- and triphosphate ions (Fig. 9b).

The above mentioned inhomogeneities cannot, however, explain the measured peak width of >1000 Da for the 104 mer transcript. In addition, the mass difference of 360 Da between the measured molecular mass (33,410 Da) and the one calculated from the known sequence (33,769 Da) cannot be rationalized by such inhomogeneities; this difference is even above the value which would appear if the n−1 species was the predominant

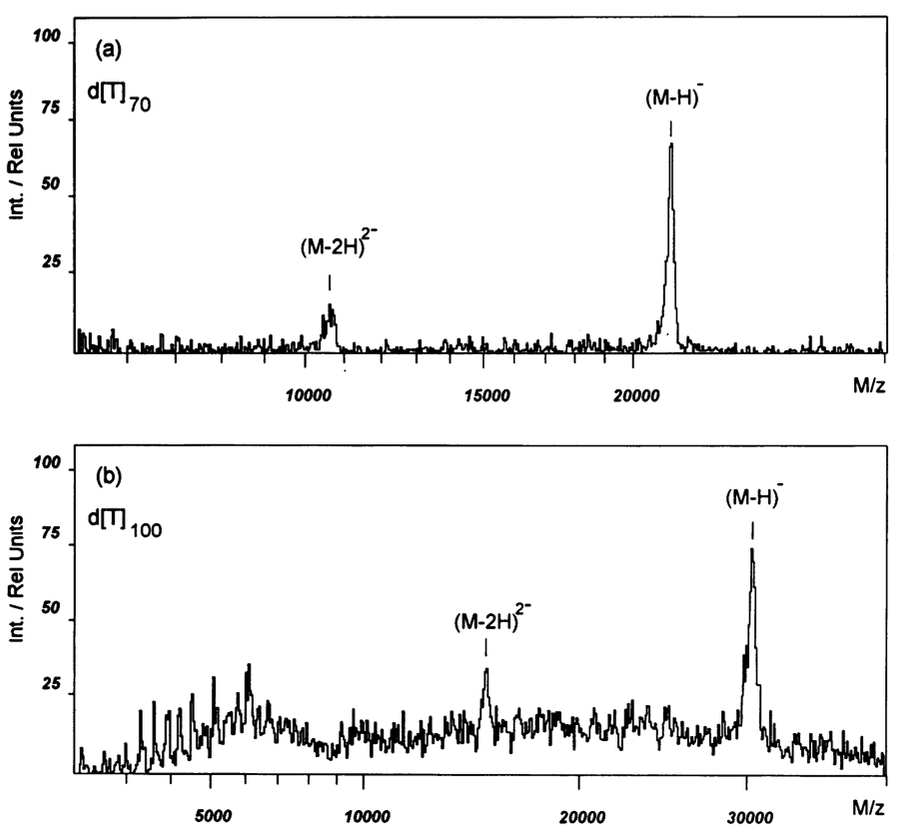

Figure 7. IR-MALDI mass spectra of the oligothymidylic acids $[T]_{70}$ **(a)** and $d[T]_{100}$ **(b)** ; total oligodeoxynucleotide load: 5 pmoles for each measurement. Both spectra are the sums of 10 single-shot spectra.

species (Dm=305 Da). A varying sequence length of 101 to 104 nucleotides, with the 103 mer as the most abundant species, combined with the loss of phosphate groups at the 5′-end could, in principle, explain the observed peak width. However, gel analysis of the generated transcripts did not support such an explanation. Limited degradation by exonucleases could be another cause for such a mass distribution and cannot be positively excluded during MALDI preparation, because the matrix solution as well as the cation exchange material and the sample support surface could in principle serve as sources for RNases. However, such a degradation has never been observed for the synthetic RNA samples. Rather extensive fragmentation could also be a major cause of the observed broad peaks, but the relatively symmetrical peak shapes contradict this hypothesis; the spectra of the larger oligodeoxynucleotide molecules, where extensive fragmentation is known to occur, all show a strong tailing of the peaks towards the lower mass end (Fig. 5c).

To improve our understanding of the fragmentation of the transcripts, two strategies will be pursued. Shorter transcripts will be produced to reduce the molecular mass to values where such inhomogeneities can be resolved by IR-MALDI-MS. The second strategy will aim at an improvement of the sample homogeneity by an optimized purification after the transcription reaction, including the release of the triphosphate group at the 5′-end of the transcripts by alkaline phosphatase treatment.

Potential applications

In the mass range of up to 7,000 Da, i.e. for oligonucleotides up to ca. 23 bases the mass can be determined to better than +/− 1 Da as shown in Figs. 3, 4, 5 and 8. In this range MALDI analysis can be used to directly determine the composition of oligonucleotides. The structures of synthetic or naturally occurring products, containing modified bases, could be verified in short time and with very little sample consumption. So far ca. 20 different, modified oligodeoxynucleotides for antisense-DNA strategies against the AIDS disease and more than 40 derivatized oligoribonucleotides from the ribozyme research field have been successfully mass analysed. These molecules were modified by e.g. the introduction of thiophosphodiester groups, by the attachment of cholesterol as a hydrophobic anchor molecule to increase target cell incorporation, by the introduction of derivatized bases such as 5-bromouracil, and the substitution of the 2′-OH-group by 2′-OCH_3. The results of this work will be published elsewhere.

Fig. 10 shows a mass spectrum of a mixture of oligouridylic acids, generated by limited acidic hydrolysis of commercially available Poly-U. The spectrum demonstrates that mixture analysis of oligonucleic acids by IR-MALDI-MS in principle is feasible as long as the single species are resolved in the spectrum. The stability of the oligouridylic acids in MALDI-MS, as apparent

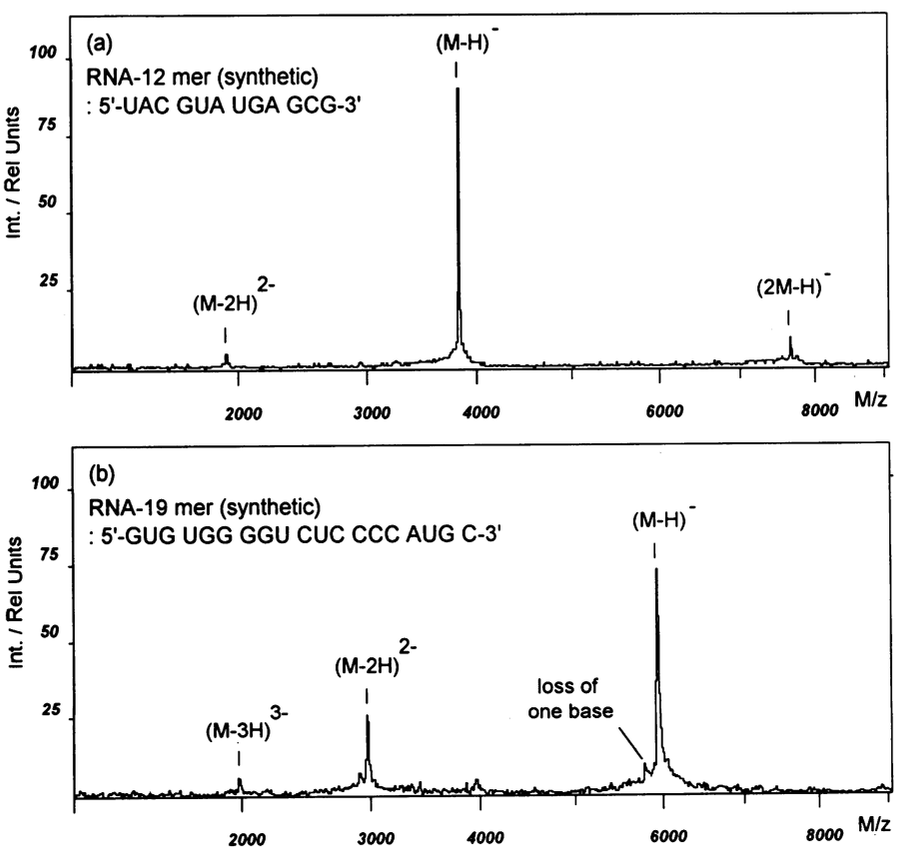

Figure 8. IR-MALDI mass spectra of synthetic oligoribonucleotides: **(a)** a 12-mer, **(b)** a 19 mer; total oligoribonucleotide load: 5 pmoles for each measurement. Both spectra are the sums of 10 single-shot spectra.

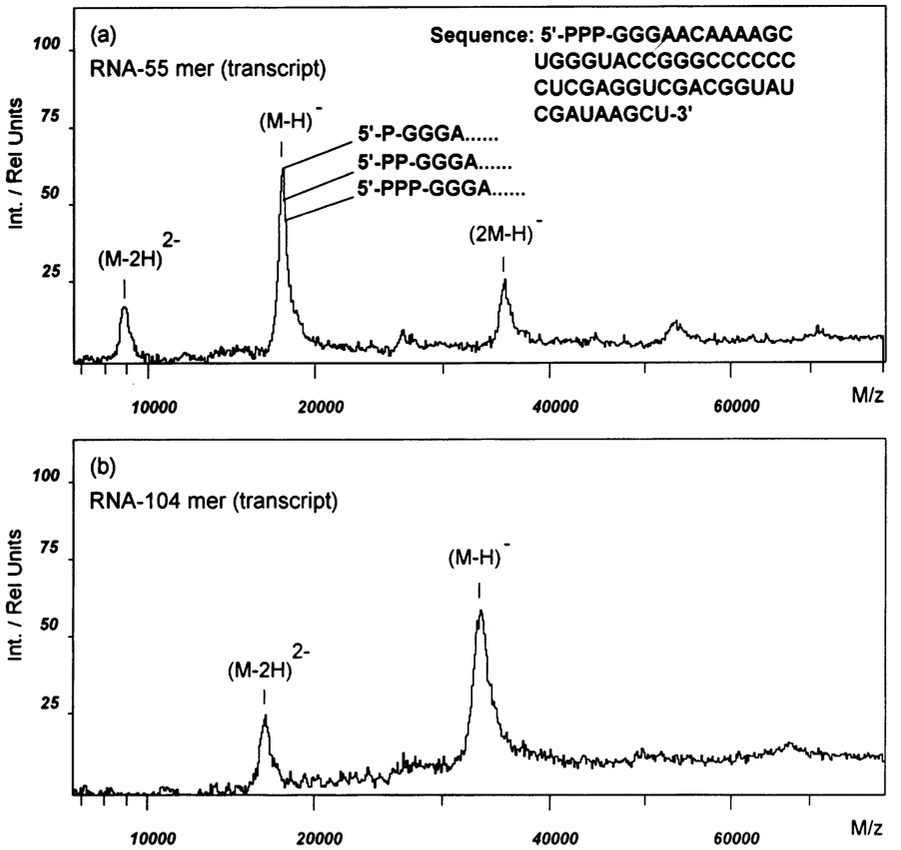

Figure 9. IR-MALDI mass spectra of RNA transcripts of **(a)** 55 nucleotides, and **(b)** 104 nucleotides. Total RNA load: 5 pmoles for each measurement. Both spectra are the sums of 10 single-shot spectra.

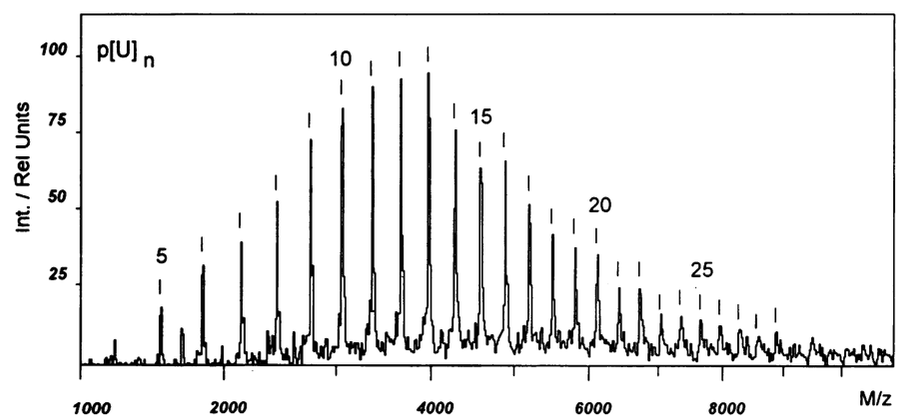

Figure 10. IR-MALDI mass spectrum of a mixture of oligouridylic acids generated by limited hydrolysis of Poly-U. Ca. 7 pmoles oligoribonucleotide were loaded. The spectrum is the sum of ten single-shot spectra.

from Fig. 10, was expected because uracil like thymine can not be protonated. However, for complex mixtures of DNA or RNA molecules, as encountered e. g. in sequencing applications, the fragmentation must be taken into account. Two different strategies could be pursued, depending on the size of the molecules in the mixture. Mixtures of small oligonucleotides (<25 nucleotides) could be analysed directly. Such mixtures, containing the complete sequence pattern, could be generated by limited degradation with a 3'- or 5'-end specific exonuclease. In this procedure, the mass differences between adjacent peaks in the mass spectrum is needed to identify the specific nucleotide unit. The mass resolution (+/− 1 Da) is sufficiently good to distinguish between any two nucleotides. The main advantage of this strategy is that rare or derivatized nucleotides can be identified by their exact mass.

In the case of larger DNA or RNA sequences, the dideoxy-nucleotide termination strategy is the method of choice and could be combined with MALDI-MS, whereby the MS-analysis would

replace the gel electrophoretic separation. Instead of four lanes on the gel, four mass spectra would be compared and the sequence could simply be read by going along the mass scale and see which mass spectrum presents the next signal. The requirement for the mass resolution in this scheme is greatly reduced, because the sequence is not derived from the exact mass differences between adjacent peaks. The main advantage of a mass spectrometric separation over the electrophoretic one would be the gain in analysis time. The time needed for sample preparation and analysis is typically a few minutes for a manual procedure and could easily be decreased to at most a few seconds in automated procedures.

The results presented show that sequencing of DNA by the use of MALDI-MS should be possible if an approach, based on 3′-deoxynucleotide termination of RNA transcripts of the target DNA, is used.

ACKNOWLEDGEMENTS

Part of this work was done in partial fulfilment of the requirement for the Ph. D. degree (of E. N.) at the University of Münster. This author would like to express his gratitude to Prof. Dr. A. Lezius for his advice and support. Financial support of this work by the German Fed. Dept. of Research and Technology (BMFT, grant No.: 13N5640) and the Commission of the European Communities (grant No.: GENO-CT91-0022 (SSMA)) is greatly appreciated.

REFERENCES

1. Karas,M. and Hillenkamp,F. (1988) *Anal. Chem.*, **60**, 2299–2301.
2. Karas,M., Bahr,U. and Hillenkamp,F. (1989) *Int. J. Mass Spectrom. Ion Proc.*, **92**, 231.
3. Hillenkamp,F. and Karas,M. (1990) *Methods Enzymol.*, **193**, 280.
4. Beavis,R.C. and Chait,B.T. (1989) *Rapid Commun. Mass Spectrom.*, **3**, 233.
5. Beavis,R.C. and Chait,B.T. (1989) *Rapid Commun. Mass Spectrom.*, **3**, 432.
6. Beavis,R.C. and Chait,B.T. (1989) *Rapid Commun. Mass Spectrom.*, **3**, 436.
7. Strupat,K., Karas,M. and Hillenkamp,F. (1991) *Int. J. Mass Spectrom. Ion Proc.*, **111**, 89–102.
8. Karas,M., Bahr,U., Ingendoh,A., Nordhoff,E., Stahl,B., Strupat,K. and Hillenkamp,F. (1990) *Anal. Chim. Acta*, **241**, 175.
9. Ehring,H., Karas,M. and Hillenkamp,F. (1992) *Organic Mass Spectrom.*, **27**, 472–480.
10. Stahl,B., Steup,M., Karas,M. and Hillenkamp,F. (1991) *Anal. Chem.*, **63**, 1463–1466.
11. Deppe,A., Bahr,U., Karas,M. and Hillenkamp,F. (1992) *Anal. Chem.*, **64**, 2866–2869.
12. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463.
13. Karas,M. and Bahr,T. (1990) *Trends Anal. Chem.*, **9**, 321.
14. Börnsen,K.O., Schär,M. and Widmer,H.M. (1990) *Chimica*, **44**, 412.
15. Huth-Fehre,T., Gosine,J.N., Wu,K.J. and Becker,C.H. (1991) *Rapid Commun. Mass Spectrom.*, **5**, 378.
16. Parr,G.R., Fitzgerald,M.C. and Smith,L.M. (1992) *Rapid Commun. Mass Spectrom.*, **6**, 369–372.
17. Hillenkamp,F., Karas,M., Ingendoh,A. and Stahl,B. (1990) In Burlingame,A.L. and McCloskey,J.A. (eds.), Biological Mass Spectrometry. Elsevier, Amsterdam, pp. 49–60.
18. Spengler,T., Pan,Y., Cotter,R.J. and Kan,L. (1990) *Rapid Commun. Mass Spectrom.*, **4**, 99.
19. Tang,K., Allman,S.L. and Chen,C.H. (1992) *Rapid Commun. Mass Spectrom.*, **6**, 365–368.
20. Nordhoff,E., Ingendoh,A., Cramer,R., Overberg,A., Stahl,B., Karas,M., Hillenkamp,F., and Crain,P.F. (1992) *Rapid Commun. Mass Spectrom.*, **6**, 771.
21. Tang,K., Allman,S.L., Jones,R.B., Chen,C.H. and Araghi,S. (1993) *Rapid Commun. Mass Spectrom.*, **7**, 63.
22. Nelson,R.W., Thomas,R.M. and Williams,P. (1990) *Rapid Commun. Mass Spectrom.*, **4**, 348.
23. Keough,T., Baker,T.R., Dobson,R.L.M., Lacey,M.P., Riley,T.A., Hassfield,J.A. and Hesselberth,P.E. (1993) *Rapid Commun. Mass Spectrom.*, **7**, 195.
24. Williams,P., Schieltz,D.M., Chou,C.-H., Luo,C.-W. and Thomas,R.M. (1992) *Rapid Commun. Mass Spectrom.*, **6**, 631.
25. Wu,K.J., Stedding,A. and Becker,C.H. (1993) *Rapid Commun. Mass Spectrom.*, **7**, 142.
26. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.
27. Ambrose,B.J.B. and Pless,R.C. (1987) *Methods Enzymol.*, **152**, 522–538.
28. Shapiro,R. and Danzig,M. (1972) *Biochemistry*, **11**, 23–29.
29. Zoltewicz,J.A., Clark,D.F., Sharpless,T.W. and Grahe,G. (1970) *J. Am. Chem. Soc.*, **92**, 1741–1750.
30. York,J.L. (1981) *J. Org. Chem.*, **46**, 2171–2173.
31. Cavalieri,L.F. and Rosenberg,B.H. (1957) *J. Am. Chem. Soc.*, **79**, 5352–5357.
32. Cox,R.A. and Peacocke,A.R. (1956) *J. Chem. Soc.*, 2499–2512.
33. Record,M.T. (1967) *Biopolymers*, **5**, 993–1008.
34. Cox,R.A. (1966) In Berg,H. (ed.), Elektrochemishe Methoden und Prinzipien der Molekular-Biologie. III. Jenaer Symposium. Akademie Verlag, Berlin, pp. 119–138.
35. Hald,M., Stein,P.C., Kirpekar,F. and Jacobsen,J.P. Manuscript in preparation.