Oligonucleotide analysis by MALDI-MS

L. Haff, P. Juhasz, S. Martin, M. Roskey, I. Smirnov, W. Stanick, M. Vestal and K. Waddell

Perseptive Biosystems, Inc., 500 Old Connecticut Path, Framingham, MA 01701, USA

Matrix-assisted laser desorption (MALDI) time-of-flight (TOF) mass spectrometry (MS) has been successfully used for oligonucleotide analysis. Delayed extraction is shown to dramatically improve the overall quality of mass spectra. Described applications include oligonucleotide sequence confirmation, determination of the extent of synthetic DNA failure sequences, characterization of the oxidation of phosphorothioate oligonucleotides, and use of MALDI-TOF for DNA sequencing by the Sanger dideoxy termination reaction.

Matrix-assisted laser desorption (MALDI) time-of-flight (TOF) mass spectrometry (MS) has been successfully used for oligonucleotide sequence confirmation by mass analysis of fragments generated by concentration-dependent exonuclease digestion. This approach is particularly well-suited to the analysis of synthetic DNA due to the relatively short length of the DNA strands. Additionally, an important application of MALDI-TOF is identification of the extent of failure sequences associated with automated oligonucleotide synthesis.

Oligonucleotides present a challenge for MALDI because their ionization is frequently accompanied by generation of a multiplicity of ions, including sodium and potassium adducts and fragments arising from cleavage of *N*-glycosidic bonds. The generation of these ions results in broadening of the peaks in the mass spectrum, compromising accurate mass determinations, sensitivity and consequently, sequence analysis. Ionization of oligonucleotides can lead to fragment ions and thus more complicated spectra. Suitable matrices for oligonucleotide analysis have been identified. These matrices promote the generation of mainly molecular ions. The matrix material 3-hydroxypicolinic acid, for example, imparts less intenal energy to the sample ions than other matrix materials, thereby reducing fragmentation but allowing sufficient energy for ionization of the oligonucleotide polymer.

In conventional MALDI-MS (continuous extraction), the ion-accelerating voltage is continuously applied. As a result, the ions formed exhibit a broad energy distribution. Additionally, when desorption occurs in a strong electrical field, the kinetic energy of the ions can be attenuated by collisions with other material in the neutral plume, resulting in further mass-dependent energy dispersion. Conventional MALDI analysis performs well for relatively short (20 – 30 mer) sequences, however this technique does not offer sufficient resolution for sequence determination of longer oligonucleotides. This has made necessary the development of higher resolution instruments for the analysis of longer oligonucleotides.

Delayed extraction (DE) technology improves instrument resolution considerably by separating the ion desorption process from the ion accelerating process (Fig. 1). In the delayed extraction mode of operation at the point of the laser impacting the sample, the sample plate and ion extraction grid are at similar potentials, so a potential gradient does not exist when the sample is ionized. After a user set time delay, a potential gradient is then applied between the sample plate and the grid. The time of flight for the ions is measured to yield the mass spectrum.

The DE-coupling of ion desorption from ion acceleration effectively minimizes the dependence of ion flight time on initial velocity. According to the theory of “time-lag energy focusing” originally proposed and developed by Wiley and McLaren [1], the dependence of ion flight time on initial velocity can be corrected to the first order by delaying the extraction of ions from the source. If higher order terms are insignificant, then the mass resolution should be determined by the ratio of the total flight time to the uncertainty in the time measurement. The observed mass resolution should then increase in proportion to the effective length of the ion flight path. Ion extraction can therefore be delayed until after the flight-density plume (formed by ablated matrix and non-ionized sample) has expanded, minimizing collisional energy loss and compensating for the range in ion energies during desorption/ionization. Thus, the ions produced by DE exhibit a narrow initial velocity distribution.

Article available at http://analusis.edpsciences.org or http://dx.doi.org/10.1051/analusis:1998260026

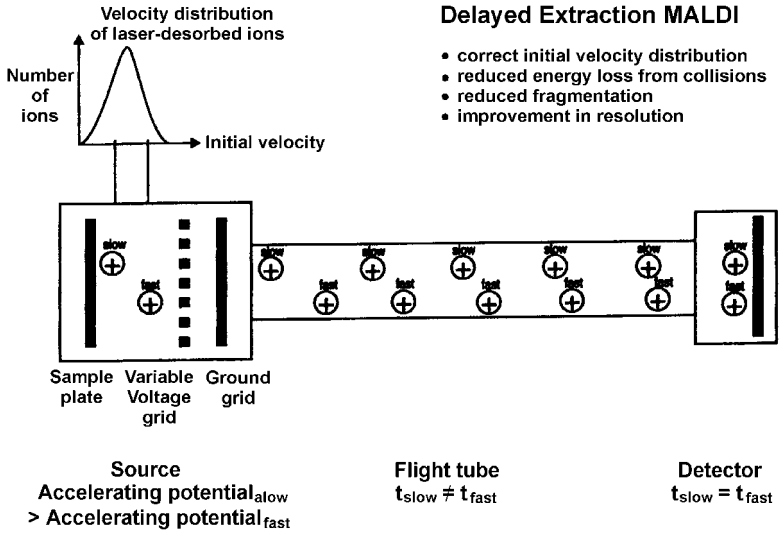

Figure 1. Delayed extraction (DE) MALDI.

The combined result of adjustments in extraction conditions and delay time improves resolution and mass accuracy. Suppressing matrix background improves the quality of the MALDI mass spectra by reducing chemical noise, and minimizing the effects of laser intensity.

The delayed extraction technique extends MALDI-TOF to the application of analysis of Sanger dideoxybase termination sequencing reactions. When coupled with a high-yield cycle sequencing protocol to enhance the amounts of termination products present in dideoxy-sequencing mixtures, DE MALDI-TOF MS provides sufficient sensitivity and resolution to enable the sequencing of dideoxy-terminated sequencing mixtures of model synthetic templates up to 40 – 50 bases in length. In a direct comparison of a mixed base 50-mer oligonucleotide analyzed in the continuous (or standard mode) *versus* the delayed extraction mode, we have found a significant improvement in sensitivity and a four-fold increase in resolution (Fig. 2). With larger mass fragments of DNA there is typically a concomitant decrease in resolution in standard MALDI, however with delayed extraction MALDI there is dramatic consistency in resolution upwards to 10 000 Da.

For DNA sequencing applications, the major advantages offered by the DE MALDI-TOF approach include very rapid data acquisition (MALDI-MS data acquisition occurs in less than a minute) and elimination of the need for sequencing gels and radioactive or fluorescent labeling reagents. The applications described in this text include examples of MALDI-TOF and DE MALDI-TOF MS – based analyses of relatively short, enzyme-digested oligonucleotides. The combination of an optimized high yield dideoxy-chain termination sequencing protocol and DE MALDI-TOF MS for DNA sequencing of DNA templates of up to 50 bases in length is also described.

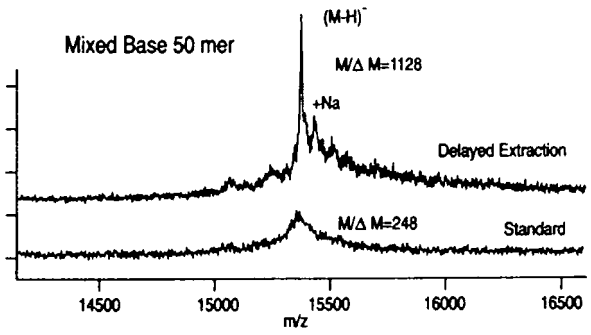

Figure 2. Comparison between standard and delayed extraction MALDI for oligonucleotide analysis.

Materials and methods

Oligonucleotide synthesis

The 31- and the 51-mer phosphodiester oligonucleotides, as well as the 16-mer fully-thioated oligonucleotide, were synthesized trityl-off using the standard 0.2 µmol protocol on a Perseptive Biosystems Expedite™ Nucleic Acid Synthesis System using base-protected ß-cyanoethyl phosphoramidite chemistry (Perseptive Biosystems, Inc.). Cleavage of the oligonucleotide from the CPG synthesis support was performed by the syringe method using 30% ammonium hydroxide at room temperature for 90 minutes. Following deprotection ovenight at 55 °C, the sample was dried under vacuum, then redissolved in water. The oligonucleotide was sequenced without further purification. Oligonucleotide concentrations were determined by measurement of optical density (OD) at 260 nm.

Other oligonucleotides were obtained in either HPLC purified or gel-purified form from Genosys (The Woodlands, TX) or Research Genetics (Huntsville, AL). The 13-base primer sequence was 5´-GCCAGGGTTTTCC-3´. Forty and 50-base templates were synthesized with the sequences 5´-biotin- ATA GCG CTG AAT TCT GCA TCG TGA CTG GGA AAA CCC C-3´´ and 5´-phosphate-TAC GTC GTA CAT AGC GCT GAA TTC TGC ATC GTG ACT GGG AAA ACC CTG GC-3´, respectively. (The 5´ modifications were made for reasons unrelated to this application.) The 50-mer was repurified by HPLC to reduce levels of contaminating failure sequences. Lyophilized oligonucleotides were resuspended in HPLC – grade water.

3´–5´ exonuclease oligonucleotide digestion

400 pmol (1 µL) of the crude trityl-off oligonucleotide was digested using the Sequazyme oligonucleotide sequencing kit (Perseptive Biosystems).

5´–3´ exonuclease oligonucleotide digestion

200 pmol (1 µL) of the oligonucleotide was digested using the Sequazyme oligonucleotide sequencing kit.

Cycle sequencing reactions

Sequencing ladders were generated by primer extension with Sequitherm™ DNA polymerase (Epicentre Technologies, Madison, WI) using the manufacturer's suggested protocol with the following modifications: 20 pmol of primer were added to 20 pmol of template in a buffer containing 50 mM Tris-HCl (pH 9.3), 2.5 mM $MgCl_2$, 0.7 mM $MnCl_2$ and 5 units of Sequitherm DNA polymerase in a total volume of 16 µL. Additionally, 0.7 mM $MnCl_2$ was included in the reaction mixture, which favored an even distribution of relatively small termination fragments between 14 and 40 bases in length. Four µL of this solution was mixed with 2 µL each of A, C, G, and T termination mixture. Reactions were thermally cycled on a Perkin-Elmer GeneAmp ™ PCR system 9600 thermal cycler.

Results and discussion

Oligonucleotide sequencing

This application describes exonuclease sequencing of a 51-mer phosphodiester oligonucleotide by MALDI-TOF on a Voyager Biospectrometry workstation. Prior to sequencing, the oligonucleotide was partially digested from the 3´-end with snake venom phosphodiesterase and in a separate reaction, from the 5´-end with bovine spleen phosphodiesterase. The resulting mixture of oligonucleotide fragments, each progressively one nucleotide shorter in length, was then analyzed by mass spectrometry. The base associated with each subsequent mononucleotide residue was then identified by comparing the masses of adjacent peaks, the difference in mass being equivalent to the mononucleotide removed by the enzymatic digestion (Fig. 3). This technique permits rapid, convenient sequence confirmation of single-stranded oligonucleotides in excess of 50 residues in length.

Sample preparation and MALDI-TOF analysis

2 µL aliquots of the above digestion mixtures were mixed with 4 µL of a 10:1 solution of 3-hydroxypicolinic acid. About 200 cation exchange beads were added to the sample (target) mixture and mixed for about 30 s. The beads were allowed to settle and the supernatant was collected for analysis. A total of 2 µL of the sample was spotted onto the Voyager Biospectrometry workstation (Perseptive Biosystems) sample plate and dried at ambient pressure and temperature. Mass spectra of the samples were produced in negative ion mode with an accelerating voltage of 25 kV with a nitrogen laser at 337 nm. The Voyager workstation was operating in linear mode.

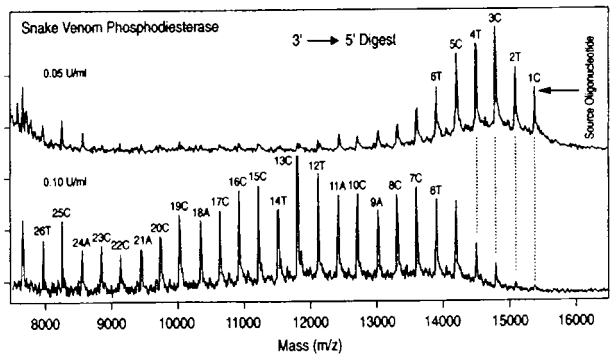

Figure 3. The mass spectrum of the 51-mer oligonucleotide following 20 min 3´–5´ digestion with snake venom phosphodiesterase. For the analysis described, enzyme concentration points were chosen that reconstructed the entire oligonucleotide sequence with a minimum number of profiles. The figures indicate the base assignments which correlate with the measured mass differences between adjacent peaks.

Analysis of oligonucleotide failure sequences

Due to incomplete nucleotide coupling during automated oligonucleotide synthesis, failure sequences are generally present as minor contaminants in the crude product. These failure sequences, generated at each step in the synthesis process, span the mass range between the 3´- mononucleotide residue and the full-length product.

Analysis of failure sequences in the crude oligonucleotide product mixture using DE MALDI-TOF MS provides a simplified method for sequence verification. In the experiment described, the sequence of a 31-mer synthetic oligonucleotide with a molecular mass of 9486.17 Da was determined by measurement of the mass differences between adjacent failure sequences. Using this method, the nucleotide product sequence can be determined down to the 3´- trinucleotide segment.

Sample preparation and DE MALDI-TOF MS analysis

The matrix 3-hydroxypicolinic (3-HPA) acid was dissolved in a 1:1 acetonitrile-water mixture at a concentration of 25 g/L; diammonium tartrate was added to a final concentration of 2.5 g/L to suppress sodium and potassium adduct formation in the mass spectrometer. The matrix solution was used to dilute the oligonucleotide to a final concentration of between 1 – 5 pmol/µL. A 2.0 µL aliquot of this solution was placed on the sample holder and dried. Mass spectra of the sample were produced in negative ion mode with an accelerating voltage of 25 kV and a nitrogen laser at 337 nm. Mass spectrometry was performed on a Voyager RP-mass spectrometer equipped with a DE ion source.

Oligonucleotide sequence analysis

Based on mass differences of adjacent failure sequences the sequence of the 31-mer oligonucleotide was determined, in agreement with calculated values, to be 3' TAG-TGG-TCA-ACG -ATT- GCC-TAC-CATCCG-TTA-G 5′ (Fig. 4). Table I shows the analysis by DE MALDI-TOF MS of failure sequences in the crude synthetic 31-mer oligonucleotide mixture depicted in figure 4. Accuracy of mass differences is adequate for sequence identification.

Table I. Analysis by DE MALDI-TOF MS of failure sequences. Listed values are for $(M–H)^-$ ions (reprinted with permission from [2], copyright 1996 American Chemical Society).

Calculated	*Observed*	*Mass difference*	*Base*
883.6	878.7		
1187.8	1183.2	304.5	T
1517.0	1513.0	329.8	G
1846.3	1842.5	329.5	G
2150.4	2147.3	304.8	T
2439.6	2436.8	289.5	C
2752.8	2750.4	313.6	A
3066.0	3063.0	312.6	A
3355.2	3352.9	289.9	C
3684.4	3682.6	329.7	G
3997.6	3996.5	313.9	A
4301.8	4300.3	303.8	T
4606.0	4605.0	304.7	T
4935.2	4934.7	329.7	G
5224.4	5223.3	288.6	C
5513.6	5512.3	289.0	C
5817.8	5817.6	305.3	T
6131.0	6129.6	312.0	A
6420.2	6419.7	290.1	C
6709.4	6709.4	289.7	C
7022.6	7021.8	312.4	A
7326.8	7327.9	306.1	T
7616.0	7616.5	288.6	C
7905.2	7905.4	288.9	C
8234.4	8237.3	331.9	G
8538.6	8542.0	304.7	T
8842.8	8844.9	302.9	T
9156.0	9156.6	311.7	A
9485.2	9485.2	328.6	G

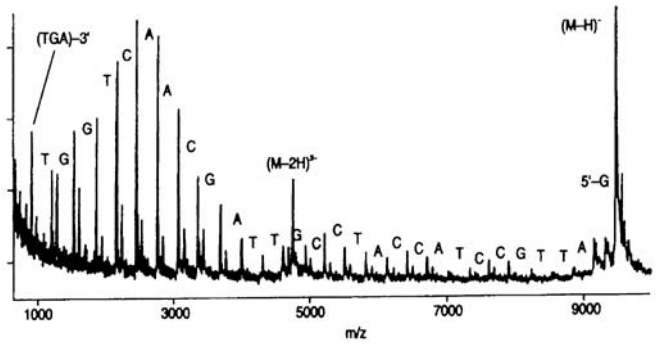

Figure 4. The MALDI mass spectrum of the 31-mer phosphodiester oligonucleotide (reprinted with permission from [2], copyright 1996 American Chemical Society).

Detection of oxidised phosphorothioate oligonucleotides

The improved mass resolution afforded by DE MALDI MS has proven useful for detection of oxidation of phosphorothioate oligonucleotides. Phosphorothioate DNA and RNA analogs have sulphur in place of oxygen as one of the non-bridging ligands bonded to phosphorus. Such modified oligos display increased resistance of the internucleotide phosphorothioate linkage to enzymatic cleavage, a property making these compounds attractive as potential antisense DNA therapeutics. Detection of contaminating oxidized by-products of these molecules by mass spectrometry offers a rapid and convenient alternative to other analytical methods such as ^{31}P NMR. The molecular mass will decrease by 16 daltons upon each oxidation event. The following experiment shows DE MALDI-TOF MS analysis of a fully-thioated 16-mer antisense phosphorothioate oligonucleotide. With DE, sufficient mass resolution was obtained to identify peaks differing by only 16 Da. DE MALDI-TOF MS analysis, using $(M–H)^-$ and $(M–2H)^{2-}$ ions of the 16-mer for internal calibration, gave an *m/z* consistent with oxidation of a single C = S bond (Fig. 5).

Sample preparation and MALDI-TOF MS analysis

The oligonucleotide was mixed with 3-hydroxypicolinic acid matrix (25 g/L solution in 1:1 acetonitrile/water) to a final oligonucleotide concentration of 5 pmol/µL. Approximately

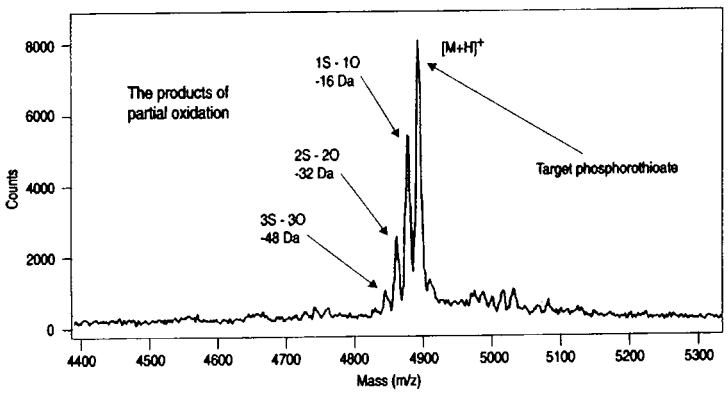

Figure 5. The DE MALDI mass spectrum of the 16-mer phosphorothioate. This resolution was sufficient to detect peaks consistent with sample components (note the –16 Da and the –32 Da peak) differing by only 16 Da in molecular weight. These components are consistent with substitution on the S in the phosphorothioate bond by an O.

0.5 µL of this mixture was deposited on the plate surface. Mass spectra of this sample were produced in the negative mode using –20 kV total accelerating voltages with the nitrogen laser at 337 nm. Linear MALDI-TOF mass spectra were recorded on an instrument with a linear drift length of 1.3 m, modified for DE. Reflector mass spectra were collected on a Perseptive Biosystems Voyager Elite instrument equipped with a single-stage reflector and DE. The ion path in this instrument in the reflector mode is 3 meters.

Sequencing DNA by the Sanger dideoxy termination reaction and DE MALDI-TOF MS

Mass spectrometry offers a number of potential advantages for sequencing DNA over the conventional electrophoretic technique, including far greater speed and the ability to sequence without radioactive or fluorescent labels. However, MALDI-TOF analysis of Sanger dideoxy termination reactions has in the past proven difficult, largely because of the relatively small concentrations of termination products produced and by the presence of buffer salts and other components which interfered with MALDI. A combination of instrumental improvements and modifications to the Sanger sequencing protocol however, now allow for sequencing of DNA. First, the advent of delayed extraction MALDI-TOF greatly increased instrumental sensitivity and resolution for DNA. Secondly by employing a cycle sequencing protocol with a thermostable DNA polymerase, it has proven possible to greatly increase the amount of termination products in a Sanger sequencing reaction. In the present technique, as much as 100 femtomole of termination fragments are produced in each 6 µL reaction mixture, and sequencing mixtures of model synthetic templates up to 50 bases in length can be achieved. Initial studies using serial dilutions of oligonucleotides indicated that the detection limit for DE MALDI was between 2 and 5 fmol, and that about 25 fmol per termination product was the minimum desirable level for sequencing with DE MALDI MS due to the increased complexity of the Sanger reaction products. Using the optimized protocol as described above, sufficient amounts of dideoxy-terminated products were generated to permit analysis. Figures 6A and 6B illustrate negative ion mass spectra of dideoxy-T terminated Sanger reactions comparing continuous *versus* delayed extraction. The amount of product generated in this reaction varied from a low of 20 fmol for the ddA terminated 29-mer to 100 fmol for the ddT terminated 38-mer.

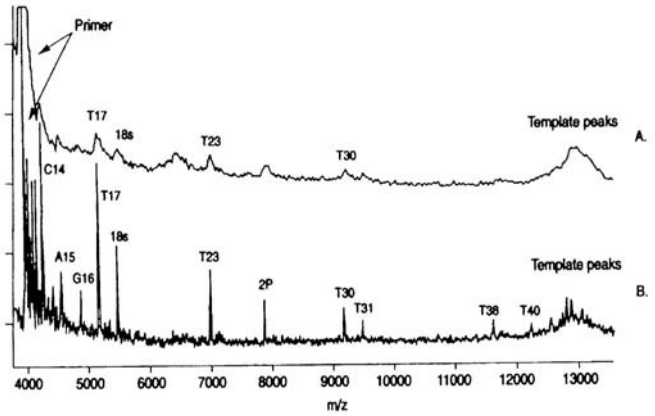

Figure 6. Negative ion mass spectra of dideoxy-T terminated Sanger reactions containing an average of 1 fmol of each terminated fragment generated by primer extension on a 40-base template. Trace A shows continuous extraction MALDI, and trace B, DE MALDI with a 400 ns delay time. The peak marked 18 s indicates nonspecific termination fragments at base position 18. The peak labeled 2P indicates a dimer of the primer. The peaks labeled as $(T - 2H)^{2-}$ represent doubly-charged template peaks. Peaks are numbered according to the total DNA fragment nucleotide length including primer from the 5´ end.

Sample preparation and MALDI-TOF analysis

Oligonucleotides: Matrix stock solutions consisted of (A) 50 g/L 3-hydroxypicolinic acid; (B) 50 g/L picolinic acid, in 50% acetonitrile/50% water (v/v) and (C) 50 g/L diammonium citrate in water. The final matrix solution was prepared immediately prior to use by mixing the three stock solutions 3-HPA, PA and diammonium citrate, in a ratio of 10:1:1. Oligonucleotides diluted to appropriate concentrations for analysis were then mixed with an equal volume of a slurry of 50W – X8 cation exchange polymer beads in the ammonium ion form. Following incubation for 5 min, the supernatant fraction was removed and adjusted to 1 µL in volume under vacuum. Oligonucleotides were mixed with 1 µL of matrix and applied to a MALDI plate previously coated with Nafion (Aldrich Chemical Co.). The plate was allowed to dry under ambient conditions.

Sequencing reaction products: Sequencing reaction products were adjusted to 50 µL with water and then desalted. The samples were then treated with ammonium-loaded cation exchange beads, as previously described, mixed with matrix and applied to the MALDI plate. Mass spectrometry was performed on a Voyager mass spectrometer equipped with delayed extraction and operating in the linear mode. The instrument has a drift length of 1.3 m and uses a nitrogen laser (Laser Science, Newton, MA) for sample desorption/ionization. Mass spectra were produced in negative ion mode with an accelerating voltage of –22 to –25 kV.

References

1. Wiley, W. C.; McLaren, I. H. *Rev. Sci. Instr.* **1955**, *26*, 1150-1157.
2. Juhasz, P.; Roskey, M. T.; Smirnov, I. P.; Haff, L. A.; Vestal, M.; Martin, S. A. *Anal. Chem.* **1996**, *68*, 941-946.