Detection of 500-nucleotide DNA by Laser Desorption Mass Spectrometry

ARTICLE *in* RAPID COMMUNICATIONS IN MASS SPECTROMETRY · SEPTEMBER 1994

Impact Factor: 2.25 · DOI: 10.1002/rcm.1290080913 · Source: PubMed

RAPID COMMUNICATIONS IN MASS SPECTROMETRY, VOL. 8, 727–730 (1994)

Detection of 500-Nucleotide DNA by Laser Desorption Mass Spectrometry

K. Tang,[†] N. I. Taranenko, S. L. Allman, L. Y. Chang,[‡] and C. H. Chen[*]
Oak Ridge National Laboratory, Photophysics Group, Chemical Physics Section, Health Sciences Research Division, Oak Ridge, Tennessee 37831-6378, USA

SPONSOR REFEREE: Professor D. M. Lubman, University of Michigan, Ann Arbor, MI 48109-1055, USA

We report the first detection of DNA segments as large as 500 nucleotides by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry, using picolinic acid and 3-hydroxypicolinic acid mixtures as desorption matrices. The successful detection of 500-nucleotide DNA indicates that laser mass spectrometry is now emerging as a new biotechnology tool for DNA-related research. It should be possible to utilize fast detection of large DNA segments by laser mass spectrometry for rapid human genome sequencing.

Since the inception of the Human Genome Project, it has been realized that the development of fast sequencing technologies is essential to the success of this ambitious health research program.[1] The conventional DNA-sequencing approach usually employs the method developed by Sanger and his co-workers.[2] Four sets of DNA fragments are produced by four different enzymatic termination reactions to obtain ladder DNA products, which are subsequently separated by size with polyacrylamide gel electrophoresis. The sequence can be determined either by radioactive material tagging and autoradiography or by dye tagging and laser-induced fluorescence detection.[3–5] The length of DNA to be sequenced is usually limited to a few hundred base pairs. To expedite DNA separation in a semi-solid gel matrix, both capillary gel electrophoresis and ultra-thin gel electrophoresis have been pursued to speed up the separation process. Nevertheless, the preparation of the gel and the time-consuming process for separation by gel makes development of alternative methods critically important for fast human-genome sequencing. Mass spectrometry has been considered as a potential ultra-fast sequencing method.[6] Since the discovery of matrix-assisted laser desorption/ionization (MALDI) for obtaining parent-ion mass spectra of proteins by Hillenkamp and his co-workers,[7] MALDI-MS has been extensively pursued for detection of nucleic acids. However, the detection of singly-stranded DNA, containing all four of the different bases, by laser mass spectrometry is still limited to small DNA segments.[8–10] The detection sensitivity is usually in the region of a few picomoles for each size of DNA.[9] Recently, restriction enzyme digested DNA fragments, which were amplified by the polymerase chain reaction (PCR) from the human genome were also detected by MALDI time-of-flight (TOF) mass spectrometry, although the ion spectrum of DNA segments shows molecular weights equivalent to the masses of singly-stranded DNA. In order to achieve fast DNA sequencing by laser mass spectrometry, the following three conditions are required: (i) the size of the DNA fragment to be detected should be larger than 300 nucleotides; (ii) the sensitivity needs to be high enough to detect quantities typical of PCR production, usually in the femtomole region; and (iii) the resolution of the mass spectra needs to be sufficiently high for sequence determination. In this work, we report the successful detection of DNA segments as large as 500 nucleotides in length. The sample was produced by PCR processing, which is typically used to amplify genomic DNA for analysis. Thus, we have met two key conditions for using mass spectrometry for fast DNA sequencing. However, the resolution in our experiment is still not sufficient for separating oligonucleotides with single nucleotide differences. If the resolution can be significantly improved, fast DNA sequencing by laser mass spectrometry could become a practical reality in the near future.

EXPERIMENTAL

A pulsed Nd:YAG laser (Spectra Physics, Model DCR2, Mountain View, CA, USA) producing a 7 ns pulse with a wavelength of 266 nm by fourth harmonic generation was used for MALDI. The laser fluences used were between 20 and 200 mJ/cm^2. A linear time-of-flight mass spectrometer capable of 45 keV acceleration was found to be important in obtaining mass spectra of these long DNA strands. Desorbed ions with high ion energy was found to significantly enhance ion detection efficiency.[11] Other details are the same as reported previously.[12] Synthetic oligonucleotides used in this work were purchased from Oligo Etc. (Wilsonville, OR, USA) and used without further purification.

Double-stranded DNA segments of 246 base pairs were amplified by polymerase chain reaction (PCR) from pLB129 (a portion of a human immunoglobin V_λ germline gene) using a common primer VL3-6 of 29-mer size (sequence: 5′-GAG GGT GGC TGA GTC AGC TCA TAG GAG GC). The sequence of the other primers was 5′-AGG AAT TCC ATG GCC TGG ACC CCT CTC). Amplification was done routinely using 100 pmol of each primer in 100 μL solution that contained dNTP, buffer, etc., from a commercial kit, Gene Amp (Perkin Elmer, Norwalk, CT, USA) for 30 cycles in the Perkin Elmer DNA Thermal Cycler. The conditions for each cycle were: denaturation at 94 °C for 1 min, annealing at 52 °C for 1 min, and extension at 72 °C for 1.5 min. The PCR products were

[†] Graduate student, Vanderbilt University, Nashville, Tennessee, USA.
[‡] Present address: Cooperative Ward-Laboratory, Chinese Academia Sinica, Taipei, Taiwan.
[*] Author for correspondence.

Received 5 July 1994
Accepted 9 July 1994

purified by phenol/chloroform extraction and alcohol precipitation and redissolved in 10 μL of pure water. 1–2 μL of this DNA solution was mixed with 1 μL aqueous solution matrix directly on the stainless-steel probe tip and dried under forced air at ambient temperature for MALDI-TOF mass spectrometric analysis.

RESULTS AND DISCUSSION

The discovery of 3-hydroxypicolinic (3-HPA) as a good matrix for MALDI of oligonucleotides by Wu *et al.*[8] led to the detection of singly-stranded DNA of 67 bases in length. Tang *et al.*[9] used 3-HPA and detected a singly-stranded DNA of 150 bases. Recently, Tang *et al.*[13] discovered that picolinic acid (PA) as a matrix was better than 3-HPA for MALDI of oligonucleotides, and detected DNA segments up to 190 nucleotides in size. Lubman and his co-workers[14] demonstrated the detection of enzyme-digested double-stranded DNA larger than 200 base pairs in size by using Nafion as the substrate and 3-HPA as the matrix. Although mixed-chemical matrices for MALDI have been tried by Wu *et al.*[8] and Tang *et al.*,[15] significant improvements over a single matrix have not been reported. In this work, we discovered that the right mixture of PA and 3-HPA can extend the capability of MALDI-MS for detecting much larger DNA segments. The following three factors may be the key reasons for achieving MALDI-MS with the mixed PA and 3-HPA matrix superior to that with a single component matrix.

(i) Adding a small amount of 3-HPA to PA can result in structural changes (defects) of PA crystals. This could lead to a decrease in the amount of heat required to vaporize the composite matrix.
(ii) Adding 3-HPA creates a small shift of the absorption band of PA, which increases its absorption coefficient. Therefore, more laser energy can be absorbed.
(iii) 3-HPA can change DNA surroundings slightly, such that the bonding between DNA and matrix molecules will be different than that for pure PA matrix material.

A mass spectrum of mixtures of singly-stranded (ss) DNA of 10 bases and 150 bases using PA and 3-HPA mixtures as the matrix is shown in Fig. 1. The signal levels are much larger than when 3-HPA was used as the matrix.[9] Both singly- and doubly-charged ions were observed. The resolutions (defined as mass peak divided by mass width, i.e., $M/\Delta M$) for the 100-mer and the 150-mer were estimated as 27 and 23 respectively. Large poly-C and poly-G ss-DNA segments have been found to be difficult to detect by using 2,5-DHB.[16] With 3-HPA and PA mixtures as matrices, we found that long poly-C and poly-G ss-DNA can be detected. Experimental results are shown in Fig. 2, Fig. 3, and Fig. 4. A MALDI spectrum of double-stranded (ds) DNA with 246 base pairs with a resolution of 22 using the same matrix mixture is shown in Fig. 5. It was observed that the DNA ions detected in the mass spectrometer always correspond to the molecular weight of ss-DNA.[17] Although the molecular weights of each ss-DNA are not the same, the limited resolution prevents the observation of two separated peaks. The DNA segment used represents a partial sequence of a human immunoglobulin V_λ gene.[15] The quantity of

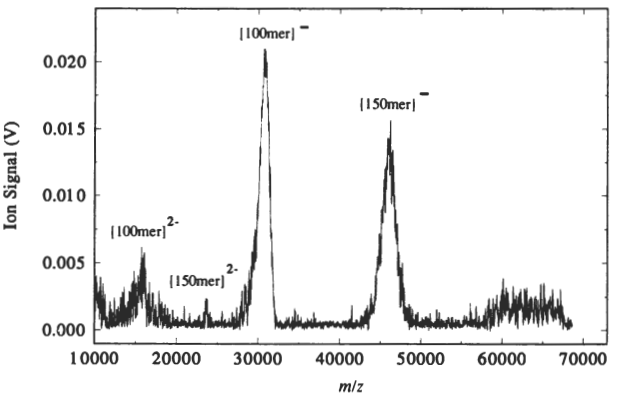

Figure 1. Negative-ion spectrum of synthetic deoxynucleotide 100-mer and 150-mer. The sample contained 15 pmol of each nucleotide, 1.2 μmol of picolinic acid, 0.15 μmol of 3-HPA. Laser wavelength was 266 nm and laser fluence was 100 mJ/cm². The resolution for the 100-mer and 150-mer are estimated as 27 and 23, respectively. The sequence of 100-mer oligonucleotide and 150-mer oligonucleotide are: 5′-TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG C-3′ and 5′-TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CCG ATT AGC CGA TTA GCC GAT TAG CTT-3′ respectively.

DNA in the sample was estimated as 200 fmol by measuring fluorescence intensity using a fluorometer (Hoefer Scientific Instruments, San Francisco, USA, Model TK0100). Figure 6 shows the mass spectrum of ds-DNA of 500 base pairs in size. The quantity of DNA in the sample was estimated as 100 fmol. Both singly and doubly charged negative ions were detected. However, the doubly charged ion signals are always much smaller than those for the singly charged ions. For DNA sequencing work in the future, the weak signals from the doubly charged ions in the mass spectra shown in Fig. 1 and Fig. 6 are likely to be observed by the prominent peak corresponding to the singly charged parent ions of DNA in appropriate sizes. It is conceivable that the first crucial criterion for the DNA sequencing work using laser mass spectrometry is met with the ability to detect DNA up to 500 base pairs in the current study.

The second condition concerning sensitivity appears also to be at least partially fulfilled, since the DNA

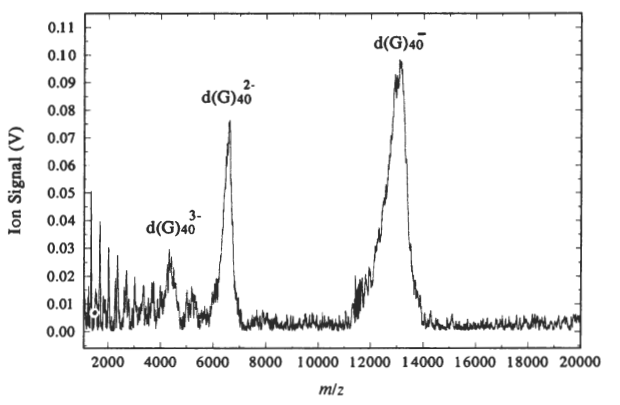

Figure 2. Negative-ion mass spectrum of $d(G)_{40}$ with picolinic acid + 3-HPA as matrix. The sample contained 30 pmol DNA, 2.4 μmol of picolinic acid and 0.05 μmol of 3-HPA. Laser wavelength was 266 nm and laser fluence was 50 mJ/cm².

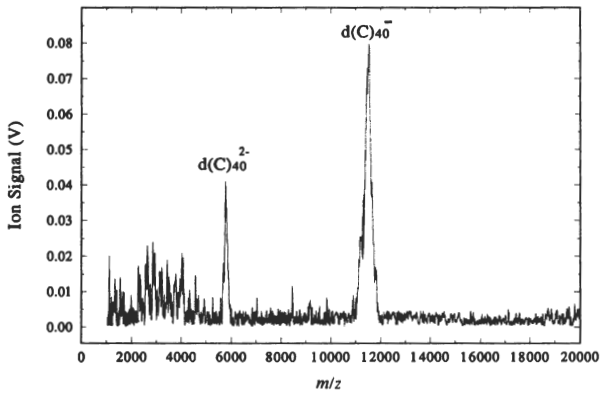

Figure 3. Negative-ion mass spectrum of $d(C)_{40}$ with picolinic acid + 3-HPA as matrix. The sample contained 30 pmol DNA, 2.4 μmol of picolinic acid and 0.05 μmol of 3-HPA. Laser wavelength was 266 nm and laser fluence was 50 mJ/cm^2.

fragments of 246 and 500 base pairs were generated by PCR for mass spectrometric analyses. The quantity of PCR products is normally produced at a picomole range and only a fraction of the amplified DNA (1/10–1/20) was routinely used for mass spectrometry. Thus, the detection sensitivity approaches the subpicomole region. Coupled with the PCR technique, laser mass spectrometry may be able to detect the sequencing products. However, the major obstacle is still the resolution of the mass spectra of the charged DNA ions. Assuming that better resolution can be obtained in the near future, then an example of mass spectrometry for DNA sequencing might be as follows. A small single-stranded oligonucleotide with the structure of GAA TTC CGC CGA TAC TGA CGG GCT CCA GGA GTC GTC GCC ACC is to be sequenced: Sanger's standard method will produce four different sets of DNA ladder products with selected terminal nucleotides such as A (adenine), T (thymine), C (cytosine), and G (guanine). These four sets of DNA ladder products are to be mixed with a mixture such as the mixture of PA and 3-HPA, and subsequently placed into the mass spectrometer for detection. Figure 7 is the illustration of the concept for using MALDI-MS for DNA sequencing. If the mass resolution can reach 500, then fast mass spectrometer sequencing of the 500-mer can be achieved. The time required for running gel electrophoresis is usually in the order of hours.

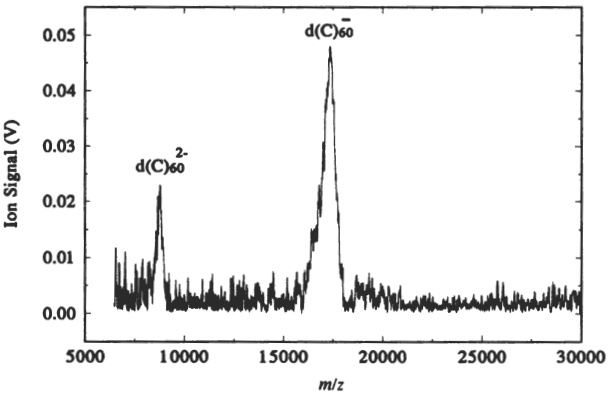

Figure 4. Negative-ion mass spectrum of $d(C)_{60}$ with picolinic acid + 3-HPA as matrix. The sample contained 30 pmol DNA and 2.4 μmol of picolinic acid and 0.05 μmol of 3-HPA. Laser wavelength was 266 nm and laser fluence was 50 mJ/cm^2.

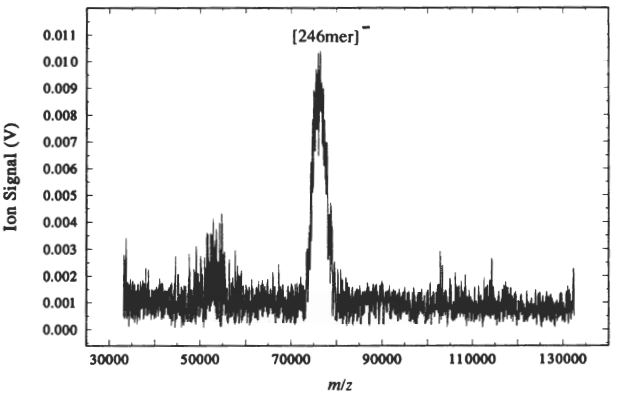

Figure 5. Negative-ion mass spectrum of a 246-base pair DNA amplified from pLB129. Total amount of DNA loaded was about 2 pmol. 1.6 μmol picolinic acid +0.2 μmol 3-HPA was used as a matrix. Laser wavelength was 266 nm and laser fluence was 100 mJ/cm^2. The resolution is 22. Although the molecular weights of each ss-DNA are not the same, the limited resolution prevents the observation of two separated peaks. The sequence of one strand of the DNA is 5′-AGG AAT TCC ATG GCC TGG ACC CCT CTC TGC CTC TCC TTC CTC ACC CTC TAC AAA GGT GAT GCC CCC AAC CCT GCC TTA GGC TCA GCC CTT ACA GGA TCC TGA GCT GGT CCT GCC CTG AAC CCT GAG CTC AGC CTA GGC ATA GCC TCA GGG TGA CAC CAC TGG AAA TGT GTT TGT TAT CTT CAA GCC CCC TCT CCT TTC CTC TCC TGC AGG CTC TGT GGC CTC CTA TGA GCT GAC TCA GCC ACC CTC-3′. The DNA segment used represents a partial sequence of a human immunoglobulin V_λ gene. The quantity of DNA in the sample was estimated to be 200 fmol by measuring fluorescence intensity using a fluorometer.

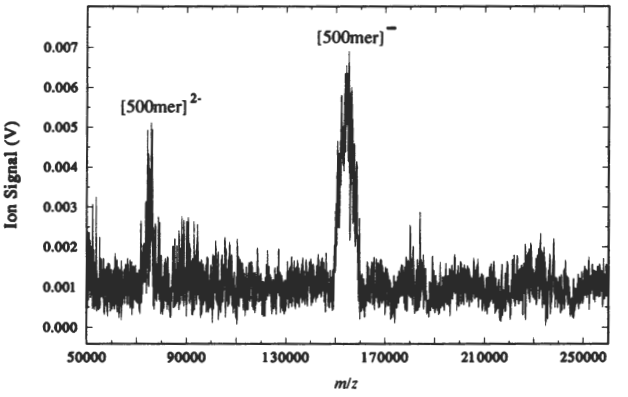

Figure 6. Negative-ion spectrum of a 500 bp DNA amplified from bacteriophage lambda genome. Total amount of DNA loaded was about 2.5 pmol. 2.4 μmol picolinic acid +0.3 μmol 3-HPA was used as a matrix. Laser wavelength was 266 nm and laser fluence was 200 mJ/cm^3. Both singly-charged and doubly-charged ions are observed. The sequence of one strand of this DNA is 5′-GAT GAG TTC GTG TCC GTA CAA CTG GCG TAA TCA TGG CCC TTC GGG GCC ATT GTT TCT CTG TGG AGG AGT CCA TGA CGA AAG ATG AAC TGA TTG CCC GTC TCC GCT CGC TGG GTG AAC AAC TGA ACC GTG ATG TCA GCC TGA CGG GGA CGA AAG AAG AAC TGG CGC TCC GTG TGG CAG AGC TGA AAG AGG AGC TTG ATG ACA CGG ATG AAA CTG CCG GTC AGG ACA CCC CTC TCA GCC GGG AAA ATG TGC TGA CCG GAC ATG AAA ATG AGG TGG GAT CAG CGC AGC CGG ATA CCG TGA TTC TGG ATA CGT CTG AAC TGG TCA CGG TCG TGG CAC TGG TGA AGC TGC ATA CTG ATG CAC TTC ACG CCA CGC GGG ATG AAC CTG TGG CAT TTG TGC TGC CGG GAA CGG CGT TTC GTG TCT CTG CCG GTG TGG CAG CCG AAA TGA CAG AGC GCG GCC TGG CCA GAA TGC AAT AAC GGG AGG CGC TGT GGC TGA TTT CGA TAA CC-3′. Of course, the other strand is the complement of this strand.

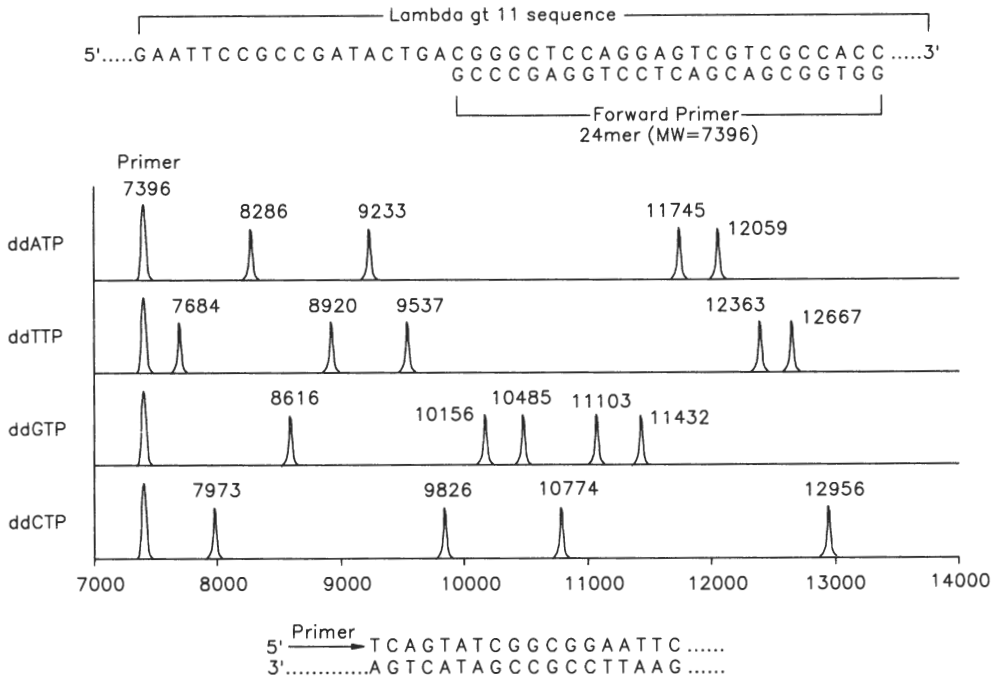

Figure 7. Conceptual mass spectra for DNA sequencing. If we need to sequence the part of the DNA with the sequencing structure of GAA TTC CGC CGA TAC TGA CGG GCT CCA GGA GTC GTC GCC ACC, a primer is needed for producing Sanger's ladder products. With ddNTP, different sizes of DNA with selected terminations are produced. There are five different sizes of DNA with the last base of T. Their sizes are 25-mer, 29-mer, 31-mer, 40-mer, and 41-mer. The number of different sizes of DNA with the last base as A, G, and C will be 4, 5, and 4 respectively. If no fragmented ions are produced, the mass spectra can easily be used to construct the sequencing information.

However, the time for analyzing a DNA sample by mass spectrometry is typically in the range of a few hundred microseconds.

We have clearly demonstrated that large sizes of DNA can be detected by laser desorption mass spectrometry. When the resolution of mass spectra of large DNA segments can reach a few hundred, fast DNA sequencing by laser desorption mass spectrometry may become a reality.

Acknowledgements

We would like to thank K. L. Lee, K. B. Jacobson, S. J. Kennel, L. M. Smith, C. H. Becker, and R. F. Haglund for very valuable discussions. We also thank N. V. Taranenko for measurements of the concentrations of DNA after the PCR process. Research was sponsored by the Office of Health and Environmental Research, U.S. Department of Energy under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc. Preparation of the manuscript by Darlene Holt is also acknowledged.

REFERENCES

1. F. Collins and D. Galas, *Science* **262**, 43 (1993).
2. F. Sanger, S. Nicklen and A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **74**, 5463 (1977).
3. U. Lardegren, R. Kaiser, C. T. Caskey and L. Hood, *Science* **242**, 229 (1988).
4. R. A. Mathies and X. C. Huang, *Nature* **359**, 167 (1992).
5. A. J. Kostichka *et al.*, *Bio/Technology* **10**, 78 (1992).
6. L. M. Smith, *Science* **262**, 530 (1993).
7. M. Karas, D. Backmann, U. Bahr and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes* **78**, 53 (1987).
8. K. J. Wu, A. Steding and C. H. Becker, *Rapid Comm. Mass Spectrom.* **7**, 142 (1993).
9. K. Tang, S. L. Allman and C. H. Chen, *Rapid Comm. Mass Spectrom.* **7**, 943 (1993).
10. K. Tang, S. L. Allman, R. B. Jones, C. H. Chen and S. Araghi, *Rapid Comm. Mass Spectrom.* **7**, 435 (1993).
11. P. W. Geno and R. D. MacFarlane, *Int. J. Mass Spectrom. Ion Processes* **92**, 195 (1989).
12. K. Tang, S. L. Allman, R. B. Jones, C. H. Chen and S. Araghi, *Rapid Comm. Mass Spectrom.* **7**, 63 (1993).
13. K. Tang, N. I. Taranenko, S. L. Allman, C. H. Chen, L. Y. Cháng and K. B. Jacobson *Rapid Comm. Mass Spectrom.* **8**, 673 (1994).
14. J. Bai, Y. H. Liu, D. M. Lubman and D. Siemieniak, *Rapid Comm. Mass Spectrom.* **8**, 687 (1994).
15. K. Tang, S. L. Allman and C. H. Chen, *Rapid Comm. Mass Spectrom.* **6**, 365 (1992).
16. G. R. Parr, M. C. Fitzgerald and L. M. Smith, *Rapid Comm. Mass Spectrom.* **6**, 369 (1992).
17. K. Tang, S. L. Allman, C. H. Chen, L. Y. Cháng and M. Schell, *Rapid Comm. Mass Spectrom.* **8**, 183 (1994).