2488–2490 *Nucleic Acids Research, 1998, Vol. 26, No. 10*

Sequencing DNA using mass spectrometry for ladder detection

N. I. Taranenko, S. L. Allman, V. V. Golovlev, N. V. Taranenko[1], N. R. Isola and C. H. Chen*

Life Science Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6378, USA and
[1]Yale University Medical Center, New Haven, CT, USA

Received December 3, 1997; Revised and Accepted March 23, 1998

ABSTRACT

Sequencing of DNA fragments of 130 and 200 bp using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for DNA ladder detection was demonstrated. With further improvement in mass resolution and detection sensitivity, mass spectrometry shows great promise for routine DNA sequencing in the future.

INTRODUCTION

DNA sequencing has become an important technique for biomedical research and clinical applications during the past two decades. With the conventional sequencing approach different sizes of DNA ladders, produced by either the Sanger enzymatic method (1) or the Maxam–Gilbert chemical cleavage method (2), are separated by gel electrophoresis to read the sequences. Gel electrophoresis involves the use of either radioisotopes or fluorescent dyes for DNA band identification and is a time-consuming process that normally requires at least a few hours. Recently capillary gel electrophoresis (3) and ultra-thin gel electrophoresis (4) have made significant progress in increasing sequencing speed. Nevertheless, speed continues to be limited by the time required for labeling and the gel running time. Several innovative sequencing technologies have been proposed which do not require gels. These technologies include sequencing by hybridization (5), scanning tunneling microscopy (6), single molecule detection using flow cytometry (7) and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (8). Among these new approaches, MALDI is emerging as a promising technology with the potential for very fast sequencing (9). In this article we present results of sequence determination of both strands of a PCR-amplified 200 bp fragment. The results presented below indicate that the mass spectrometric method of DNA sequencing has advanced to the point of being useful in biomedical research. This method is still below the capability of gel or capillary electrophoretic methods, but the use of unlabeled primers and fast sequence determination make the MALDI approach an attractive alternative.

Since the discovery of MALDI in 1987 (10) sequencing of DNA by mass spectrometry has been proposed. However, in early studies the detection of DNA by MALDI was limited to very short synthetic oligonucleotides (11) due to fragmentation of longer species during the laser desorption step. New matrices better suited to DNA have been discovered (12–14) and single-stranded (ss)DNA of 500 nt and double-stranded (ds)DNA longer than 500 bp have been detected (15,16). DNA fragments generated from the polymerase chain reaction (PCR) have also been successfully detected by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) (17). The recent development of delayed pulsed ion extraction (18–20) significantly improves the resolution of proteins and small DNAs. However, the mass resolution of long DNA fragments (DNAs longer than 300 nt) by MALDI is still not high enough to obtain sequencing information.

During the past few years mass spectrometry for sequencing short DNA fragments has been pursued. In 1993 Smith and his colleagues (21) used MALDI-TOF-MS to analyze mock Sanger sequencing reactions containing mixtures of synthetic oligonucleotides. In 1995 Shaler *et al*. (22) used MALDI to sequence ladders of a 45mer ssDNA that had been prepared enzymatically. Recently Roskey *et al*. (23) reported the results of sequencing a synthetic ssDNA template of 50 nt. However, several unidentified peaks were observed and several peaks corresponding to positions on the DNA ladders could not be clearly identified. Koster *et al*. (24) also succeeded in sequencing synthetic ssDNA of 39 nt using streptavidin-coated magnetic beads for purification of DNA ladders. A synthetic 78mer ssDNA was also used as a template for sequencing, but only the G termination reaction was reported; serious fragmentation due to depurination was observed. Smith and co-workers (25) reported MALDI analysis of sequencing reactions from bacteriophage M13 and determined the sequence for samples up to 35 bases in length, however, a few false stops were observed. A complete sequencing effort for a ssDNA with 50 bases was demonstrated with no false stops, no serious fragmentation and no unidentified peaks by Taranenko *et al*. (26). Sequencing of DNA fragments of 100 nt was presented by Monforte and Becker (27), but with no experimental details.

The size of DNAs that can be sequenced using mass spectrometry is shorter than using the gel electrophoresis method because mass resolution and detection sensitivity decrease rapidly for longer DNAs. In order to compensate for poor detection sensitivity we have modified the protocol for DNA ladder preparation to produce larger quantities of long DNA ladders. Furthermore, the ratio of ddNTP to dNTP was optimized for production of these sizes of DNA ladders. In this work we present results using MALDI for sequencing of DNA templates of 130 and 200 bp.

MATERIALS AND METHODS

DNA segments to be sequenced were amplified by PCR. Amplification of selected DNA targets was performed using a Perkin Elmer DNA Thermal Cycler (Norwalk, CT). The entire

*To whom correspondence should be addressed. Tel: +1 423 574 5895; Fax: +1 423 576 2115; Email: chenc@ornl.gov

bacteriophage λ DNA served as the template for PCR using a GeneAmp PCR reagent kit (Perkin Elmer). Two primer pairs which define a 130 base (7131–7260) and 200 base (7131–7330) region of the phage genome were used for PCR amplification.

From the PCR kit a ssDNA 25mer, complimentary to λ nt 7131–7155, was used as primer 1 for preparing the 130mer and 200mer PCR fragments. The primer 1 sequence is d(GATGAGTT-CGTGTCCGTACAACTGG). Primer 2, a 20mer complimentary to λ nt 7242–7260, was purchased from Oligos Etc. (Wilsonville, OR) and used without further purification and in combination with primer 1 to produce the 130 bp PCR product. The sequence of primer 2 is d(TCACGGTTCAGTTGTTCACC). To produce the 200mer PCR product primer 1 was used in combination with primer 3, a 20mer that was purchased from Oligos Etc. The sequence of primer 3 is d(GTCATCAAGCTCCTCTTTCA). Amplification for 30 cycles was done routinely using generally applicable conditions and buffers. 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. To ensure the success of each PCR reaction the PCR products were checked by 2% agarose gel electrophoresis. Bands that corresponded to the expected PCR products were clearly observed. The amplified product was purified using QIAquick DNA purification (Qiagen Inc., Chatsworth, GA), ethanol precipitated and concentrated. The purified template DNA was dissolved in 50 µl sterile deionized water and was used as template for sequencing.

The sequencing reactions on dsDNAs of 130 and 200 bp were performed using a Thermo Sequenase DNA polymerase cycle sequencing formulation (Thermo Sequenase cycle sequencing kit; Amersham Life Science). Since the sequencing kit was not designed to obtain sequences close to the primer, a substantial modification of the sequencing reaction conditions was needed. The balance between chain extension and chain termination was optimized to obtain reliable results. For each set of these four sequencing mixes we scaled up the reaction with four times the vendor's recommended concentration in a single tube. The synthesis of DNA ladders was carried out in one step. Each reaction was performed using 100 pmol 20mer primer (1, 2 or 3 as appropriate) and 4 pmol template (ratio 25:1) with 1 mM each dNTP (7-deazaguanosine triphosphate was used to eliminate most secondary structures) and 2 mM each ddNTP. All reagents were purchased from Amersham Life Sciences (Arlington Heights, IL). The concentration of $MgCl_2$ in the reaction buffer was adjusted to balance the dNTPs and ddNTPs and the pH was 9.5. DNA polymerase (8 U) was added in accordance with the manufacturer's cycle sequencing protocol. Fifty cycles of reaction were performed in a Perkin Elmer thermal cycler to produce DNA ladders. The conditions for each cycle were: denaturation at 95 °C for 30 s, annealing and polymerization at 72 °C for 120 s. The reaction was stopped by cooling on ice.

Molecular weights and number of bases of the sequencing ladders generated from a 130 bp template with forward and reverse primers for A, C, T and G reactions were calculated and compared with the peaks obtained from experimental results. They were in good agreement, with variances <100 Da.

All sequencing reaction products were purified on a G-25 Sephadex spin column (QuickSpin; Boehringer Mannheim) following the manufacturer's protocol. After ethanol precipitation the DNA was dried under vacuum and then dissolved in 1 µl sterile deionized water. Sample preparation for MALDI analysis was done by mixing 1 µl analyte DNA fragments with 1 µl matrix solution directly on stainless steel plates and air drying at ambient temperature. The matrix solution was 3-hydroxypicolinic acid:picolinic acid:ammonium fluoride (molar ratio 9:1:1) in 50% aqueous acetonitrile.

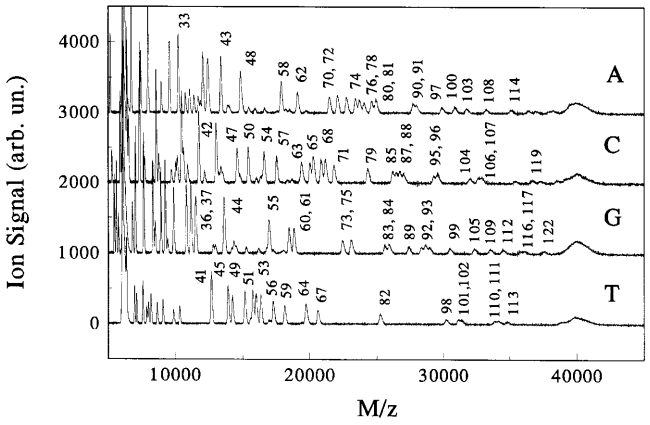

Figure 1. Negative ion mass spectra of Sanger sequencing ladders produced by dideoxynucleotide chain termination of the PCR product of a double-stranded 130 bp template (+ strand) with a 20 nt reverse primer. The sequence of the (–) strand 130mer is d(GATGAGTTCG TGTCCGTACA ACTGGCGTAA TCATG-GCCCT TCGGGGCCAT TGTTTCTCTG TGGAGGAGTC CATGACGAAA GATGAACTGA TTGCCCGTCT CCGCTCGCTG GGTGAACAAC TGAA-CCGTGA).

A linear TOF mass spectrometer (Voyager; PerSeptive Biosystems, Framingham, MA) equipped with a pulsed nitrogen laser with a wavelength of 337 nm for desorption and ionization was used in obtaining mass spectra of DNA samples. The laser fluence was kept just above the threshold for observing the DNA peaks. The typical laser fluence was 45–65 mJ/cm^2. The acceleration voltage was 28 125 V. A delayed pulsed ion extraction device was installed to increase mass resolution. The vacuum of the TOF chamber was typically 1.5×10^{-7} Torr. The signals of negative DNA ions were collected (typically for >100 laser pulses), digitized and averaged by a digital oscilloscope (Tektronix 520A) controlled by a laboratory computer.

RESULTS AND DISCUSSION

We used PCR to produce a template of 130 bp that was used as template with primer 1 or 2 for Sanger sequencing. The PCR product was checked by MALDI to ensure the success of PCR reactions (28). Figure 1 shows the negative ion mass spectra of DNA ladders from dideoxy A, C, G and T reactions using the reverse primer. Although the resolution is poor for some adjacent peaks (see peaks for 90mer and 91mer in Fig. 1), the sequence can still be determined from the lack of peaks in the spectra for the other reactions. Sequence identification was obtained up to 120 bases, which represents the sequence of 100 nt. The quantities of primers and templates used in producing DNA ladders are rather high, at a few picomoles, however, the detection sensitivity can be improved by making smaller sample spots. Since our laser spot illuminates <10% of the sample, improvement in detection sensitivity by a factor of 10 is certainly feasible (29).

Forward primers were also used for sequencing, with results similar to the data shown in Figure 1. Since sequencing using both forward and reverse primers was demonstrated to cover 100 nt, complete sequencing of 200 bp DNA can be achieved by operating from both ends. Figure 2 illustrates the experimental results of combining the mass spectra from forward and reverse primers to sequence 160 bp dsDNA. The DNA peaks from the

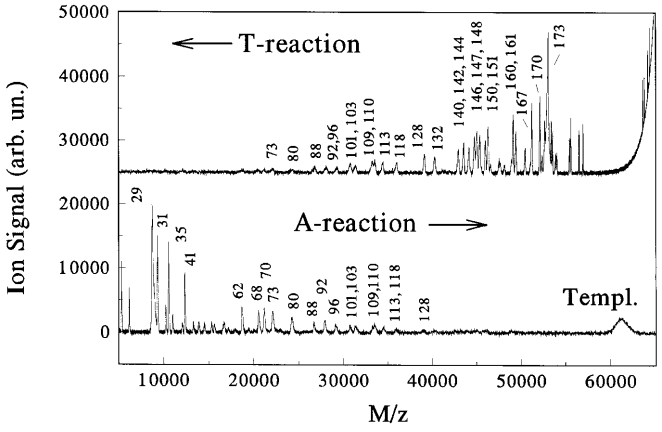

Figure 2. Negative ion mass spectra of Sanger sequencing ladders produced by dideoxynucleotide chain termination of the PCR product of a double-stranded 200 bp template. The results for the A reaction with reverse primer and T reaction with forward primer are presented. The sequence of the (–) strand 200 bp DNA is d(GATGAGTTCG TGTCCGTACA ACTGGCGTAA TCATGGCCCT TCGGGGCCAT TGTTTCTCTG TGGAGGAGTC CATGACGAAA GATG-AACTGA TTGCCCGTCT CCGCTCGCTG GGTGAACAAC TGAACCGT-GA TGTCAGCCTG ACGGGGACGA AAGAAGAACT GGCGCTCCGT GTGGCAGAGC TGAAAGAGGA GCTTGATGAC).

shorter DNA fragments, obtained from each end, are prominent and the signal is still adequate for the longer DNAs obtained in both directions.

At present automatic gel sequencers can read up to 1000 bp, but most routine sequencing ranges from 300 to 500 bases. We present results utilizing MALDI-TOF-MS to determine sequences up to 160 bp, indicating that mass spectrometry is emerging as a useful tool for DNA sequencing. Mass spectrometric detection utilizes unlabeled primers. It has the additional advantages of higher sequencing speed and the ability to obtain sequence information close to the primer, which is often difficult with gel techniques. We have also demonstrated the ability to sequence double-stranded templates generated by PCR amplification of specific target sequences. This capability offers a powerful tool for re-sequencing short regions of DNAs for diagnostic purposes, such as sequencing a *p53* exon to identify mutations. Mass spectrometric detection of DNA ladders can be used to obtain the sequence and identify point mutations. For short DNA fragments point mutations can be identified by mass spectrometric measurements of the sizes of DNA alone. However, the mass resolution is currently not adequate to detect the 9 Da mass difference between adenine (A) and thymine (T) in a 100 nt segment of ~30 kDa. Mass spectrometry can help resolve sequence ambiguities in DNAs with a high percentage of GC-rich regions and other secondary structures not resolvable by gel or capillary electrophoretic methods.

CONCLUSION

MALDI-TOF-MS has been successfully used for detection of the components of DNA ladders to achieve sequencing of dsDNA templates of 130 and 200 bp using 20mer primers. With improved resolution and sensitivity in the future, mass spectrometric DNA sequencing should become very valuable for *de novo* sequencing. With the rapid discovery of new genes, mass spectrometry should be even more valuable for re-sequencing for mutation detection, since the size of DNAs to be re-sequenced usually does not need to be long. We expect that mass spectrometry will complement the existing sequencing technologies by being a fast label-free method and will be useful in resolving sequence ambiguities.

ACKNOWLEDGEMENTS

We would like to thank K.Bruce Jacobson, Marvin Stodolsky, Steve Martin, Lan-Yang Chang and Larry Haff for valuable discussions. We would also like to acknowledge the loan of the Voyager mass spectrometer to ORNL by PerSeptive Biosystems Inc. This work was sponsored by the Office of Biological and Environmental Research, US Department of Energy under contract DE-AC05-96OR22464 with Lockheed Martin Energy Research Corporation. Preparation of the manuscript by Darlene Holt is also acknowledged.

REFERENCES

1 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
2 Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
3 Cohen,A.S., Najarian,D.R., Paulus,A., Guttman,A., Smith,J.A. and Karger,B.L. (1988) *Proc. Natl. Acad. Sci. USA*, **85,** 9660–9663.
4 Kostichka,A.J. *et al.* (1992) *BioTechnology*, **10,** 78–81.
5 Dramanac,R., Dramanac,S., Strezoska,Z., Paunesku,T., Labat,I. and Zeremski,M. (1993) *Science*, **260**, 1649–1652.
6 Hansma,P.K., Eling,V.B., Marti,O. and Bracker,C.E. (1988) *Science*, **242**, 209–216.
7 Keller,R.A., Ambrose,W.P., Goodwin,P.M., Jett,J.C., Martin,J.C. and Wu,M. (1996) *Appl. Spectrosc*., **50**, 12A–32A.
8 Karas,M. and Hillenkamp,F. (1988) *Anal. Chem*., **60**, 2299–2301.
9 Smith,L.M. (1996) *Nature Biotechnol*., **14**, 1084–1087.
10 Karas,M., Beckmann,D., Bahr,U. and Hillenkamp,F. (1987) *Int. J. Mass Spectrom. Ion Process*., **78**, 53–68.
11 Parr,G.R., Fitzgerald,M.C. and Smith,L.M. (1992) *Rapid Commun. Mass Spectrom*., **4**, 369–372.
12 Wu,K.J., Steding,A. and Becker,C.H. (1993) *Rapid Commun. Mass Spectrom*., **7**, 142–146.
13 Tang,K., Taranenko,N.I., Allman,S.L., Chen,C.H., Chang,L.Y. and Jacobson,K.B. (1994) *Rapid Commun. Mass Spectrom*., **8**, 673–677.
14 Taranenko,N.I., Tang,K., Allman,S.L., Chang,L.Y. and Chen,C.H. (1994) *Rapid Commun. Mass Spectrom*., **8**, 1001–1006.
15 Tang,K., Taranenko,N.I., Allman,S.L., Chang,L.Y. and Chen,C.H. (1994) *Rapid Commun. Mass Spectrom*., **8**, 727–730.
16 Bai,J., Liu,Y., Liang,X., Zhu,Y. and Lubman,D.M. (1995) *Rapid Commun. Mass Spectrom*., **9**, 1172–1176.
17 Tang,K., Allman,S.L., Chen,C.H., Chang,L.Y. and Schell,M. (1996) *Rapid Commun. Mass Spectrom.*, **8**, 183–183.
18 Vestal,M.L., Juhasz,P. and Martin,S.A. (1995) *Rapid Commun. Mass Spectrom*., **9**, 1044–1050.
19 Brown,J.S. and Lennon,J.J. (1995) *Anal. Chem*., **67**, 1998–2003.
20 Whittal,R.M. and Li,L. (1995) *Anal. Chem*., **67**, 1950–1954.
21 Fitzgerald,M.C., Zhu,L. and Smith,L.M. (1993) *Rapid Commun. Mass Spectrom*., **7**, 895–897.
22 Shaler,T.A., Tan,Y., Wickham,J.N., Wu,K.J. and Becker,C.H. (1995) *Rapid Commun. Mass Spectrom*., **8,** 537–540.
23 Roskey,M.Y., Juhasz,P., Smirnov,I.P., Takach,E.J., Martin,S.A. and Haff,L.A. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 4724–4729.
24 Koster,H., Tang,K., Fu,D., Braun,A., van der Boom,D., Smith,C.L., Cotter,R.J. and Cantor,C.R. (1996) *Nature Biotechnol*., **14**, 1123–1128.
25 Mouradian,S., Rank,D.R. and Smith,L.M. (1996) *Rapid Commun. Mass Spectrom*., **10**, 1475–1478.
26 Taranenko,N.I., Chung,C.N., Zhu,Y.F., Allman,S.L., Golovlev,V.V., Isola,N.R., Martin,S.A., Haff,L.A. and Chen,C.H. (1997) *Rapid Commun. Mass Spectrom*., **11**, 386–392.
27 Monforte,J.A. and Becker,C.H. (1997) *Nature Med*., **3**, 360–362.
28 Tang,K., Allman,S.L., Chen,C.H., Chang,L.Y. and Schell,M. (1994) *Rapid Commun. Mass Spectrom*., **8**, 183–186.
29 Little,D.P., Cornish,T.J., O'Donnel,M.J., Braun,A., Cotter,R.J. and Koster,H. (1997) *Proceedings of the 45th ASMS Conference on Mass Spectrometry and Allied Topics*, June 1–5, 1997, Palm Springs, CA, p. 11.