DNA sequencing using biotinylated dideoxynucleotides and mass spectrometry

John R. Edwards[1,2], Yasuhiro Itagaki[3] and Jingyue Ju[1,2,*]

[1]Laboratory of DNA Sequencing and Chemical Biology, Columbia Genome Center, Columbia University College of Physicians and Surgeons, New York, NY 10032, USA, [2]Department of Chemical Engineering and [3]Department of Chemistry, Columbia University, New York, NY 10027, USA

ABSTRACT

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MS) has been explored widely for DNA sequencing. The major requirement for this method is that the DNA sequencing fragments must be free from alkaline and alkaline earth salts as well as other contaminants for accurately measuring the masses of the DNA fragments. We report here the development of a novel MS DNA sequencing method that generates Sanger-sequencing fragments in one tube using biotinylated dideoxynucleotides. The DNA sequencing fragments that carry a biotin at the 3′-end are made free from salts and other components in the sequencing reaction by capture with streptavidin-coated magnetic beads. Only correctly terminated biotinylated DNA fragments are subsequently released and loaded onto a mass spectrometer to obtain accurate DNA sequencing data. Compared with gel electrophoresis-based sequencing systems, MS produces a very high resolution of DNA-sequencing fragments, fast separation on microsecond time scales, and completely eliminates the compressions associated with gel electrophoresis. The high resolution of MS allows accurate mutation and heterozygote detection. This optimized solid-phase DNA-sequencing chemistry plus future improvements in detector sensitivity for large DNA fragments in MS instrumentation will further improve MS for DNA sequencing.

INTRODUCTION

With the completion of the first human genome sequence map, many areas that are highly polymorphic in both exons and introns of the genome will be known. Resequencing of the polymorphic areas that are linked to disease development will greatly contribute to the understanding of the disease and therapeutic development. Thus, high-throughput accurate methods for resequencing the highly variable intron/exon regions in the genome are needed in order to explore the potential of the complete human genome sequence map. The current state-of-the-art technology for high-throughput DNA sequencing, such as used in the Human Genome Project, is capillary array DNA sequencing using laser-induced fluorescence detection (1–4). Although this technology addresses the throughput and read length requirements of large-scale DNA-sequencing projects, the accuracy required for mutation detection needs to be improved for a wide variety of applications ranging from disease gene discovery to forensic identification. For example, electrophoresis-based sequencing methods have difficulty in detecting heterozygotes unambiguously and are not 100% accurate on a given base due to compressions in GC-rich regions (5,6). In addition, the first few bases after the priming site are often masked by the high fluorescence intensity from the excess dye-labeled primers or dye-labeled terminators, and are therefore difficult to identify.

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been explored widely for DNA sequencing (7–13). The Sanger dideoxy procedure is generally used to produce DNA-sequencing fragments. Compared with gel electrophoresis-based sequencing systems, MS produces a very high resolution of DNA-sequencing fragments, fast separation on microsecond time scales, and completely eliminates the compressions associated with gel electrophoresis. One challenge to sequencing DNA with MS is the stringent purity requirement for the DNA-sequencing fragments introduced to the mass detector, because the DNA sequences are determined by accurately measuring the mass of the DNA-sequencing fragments, which must be free from alkaline and alkaline earth salts as well as other contaminants.

Approaches for purifying DNA samples using the strong interaction of biotin and streptavidin (14) coupled with solid-phase methods are widely used (15–20). For DNA sequencing with MS, Monforte and Becker demonstrated read lengths up to 100 bp by purifying DNA-sequencing samples using a cleavable biotinylated primer (12). In this method, the extension fragments from the primer are captured by streptavidin-coated magnetic beads at the 5′-end of the extension fragments, while the other components in the sequencing reaction are washed away. Fu *et al.* reported the complete sequencing of exons 3 and 5 of the p53 gene using MS with an average read length of 35 bp (13). Here, the DNA-sequencing samples were processed through the use of immobilized DNA templates on a solid phase for one cycle extension. The extended DNA

*To whom correspondence should be addressed at: Room 405A, Russ Berrie Medical Science Pavilion, Columbia University College of Physicians and Surgeons, 1150 St Nicholas Avenue, New York, NY 10032, USA. Tel: +1 212 304 7369; Fax: +1 212 304 5515; Email: dj222@columbia.edu
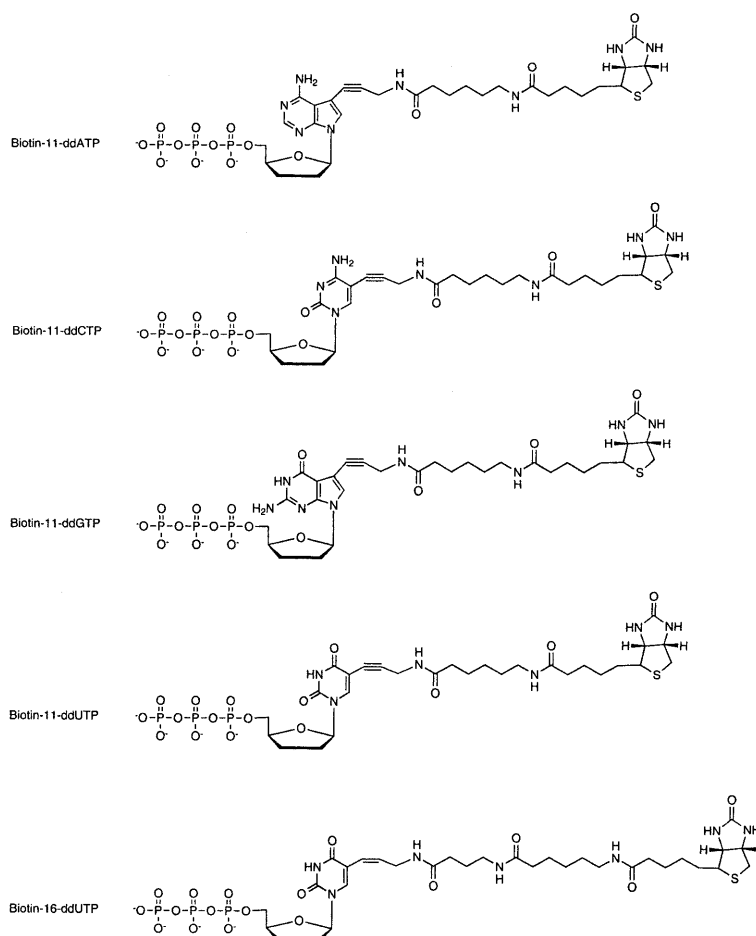
Figure 1. The structure of the biotinylated dideoxynucleotides used to generate Sanger sequencing products for MS analysis. The biotinylated ddUTPs function the same way as ddTTP in the Sanger DNA-sequencing reaction.

fragments were hybridized on the immobilized templates, while the other components in the sequencing reaction were eliminated. These efforts established the feasibility of using MALDI-TOF MS for high-throughput DNA sequencing up to 100 bp. However, in both methods, falsely stopped DNA-sequencing fragments are not eliminated and are introduced to the mass spectrometer. In addition, four separate reactions were performed, one for each dideoxynucleotide terminator analogous to the approach used in dye-labeled primer sequencing. Falsely terminated DNA fragments (false stops) are generated in Sanger-sequencing reactions when a DNA fragment terminates by incorporating a deoxynucleotide rather than a dideoxynucleotide. It has been shown that false stops and dimerized primers can produce extra peaks in the mass spectra preventing accurate base identification (8).

Ideally, for DNA sequencing with MALDI-TOF MS, one would like to establish a procedure that allows sequencing reactions to be performed in one tube to simplify sample preparation, uses cycle sequencing to increase the yield of the DNA-sequencing fragments, and isolates pure DNA-sequencing fragments free of false stops. We have developed a high-fidelity DNA-sequencing method using a dye-labeled primer and solid-phase capturable dideoxynucleotide terminators

(biotinylated ddNTPs) (21). After capture/release of the DNA-sequencing fragments at the 3′-end from the streptavidin-coated solid phase, only the pure DNA-sequencing fragments are loaded and detected on sequencing gels. This method is effective to remove falsely stopped DNA fragments for unambiguous genetic mutation detection. However, GC-rich compression issues still exist due to the use of gel electrophoresis. Here we explore the use of biotinylated dideoxynucleotides for the development of a high-fidelity DNA-sequencing method using MS. Through the use of biotinylated dideoxynucleotide terminators, our MS DNA-sequencing method utilizes the high affinity between the small biotin molecule and the protein streptavidin to isolate pure DNA-sequencing fragments. The approach is briefly described as follows. Once the DNA-sequencing fragments are terminated by biotinylated dideoxynucleotides at the 3′-end of the cycle-sequencing reaction, they are bound to a streptavidin-coated solid phase. All excess primers, salts and falsely terminated DNA fragments are washed away to provide pure DNA samples.

Previous reports using biotin and streptavidin for purifying DNA for MS detection involved attaching biotin to the 5′-end of the DNA (13), which cannot eliminate falsely stopped

DNA-sequencing fragments. The novelty of our method is to introduce a biotin moiety at the 3′-end of DNA-sequencing fragments through the use of biotinylated dideoxynucleotides. Thus, the DNA fragments correctly terminated by biotinylated ddNTPs will carry a biotin at the 3′-end, which will be captured by the streptavidin-coated solid phase. Falsely terminated DNA fragments will not carry a biotin and, therefore, will not be captured by the streptavidin-coated solid phase. This allows the isolation of only true DNA-sequencing fragments for MS detection. MALDI-TOF MS has been widely used for single nucleotide polymorphism (SNP) detection. The technique developed by Tang *et al.* involves immobilizing DNA templates on a chip and extends one to several bases for determining a particular SNP (22). Although this chip-based MS DNA analysis technique provides a reasonable throughput for SNP analysis, a cycle-sequencing method to amplify the extension fragments cannot be used in the chip format, because the extended DNA fragments must remain annealed to the DNA template that is immobilized on the chip. In summary, our method provides an efficient way to generate a high yield of DNA fragments using cycle sequencing and then to isolate only pure DNA-sequencing fragments for MS detection. This solid-phase DNA-sequencing chemistry will facilitate the development of a high-throughput and high-fidelity MS DNA-sequencing system to analyze genetic polymorphisms for genetic disease association studies (23).

MATERIALS AND METHODS

Biotinylated dideoxyadenosine-5′-triphosphate (Biotin-11-ddATP), biotinylated dideoxyguanosine-5′-triphosphate (Biotin-11-ddGTP), biotinylated dideoxyuridine-5′-triphosphate (Biotin-11-ddUTP) and biotinylated dideoxycytidine-5′-triphosphate (Biotin-11-ddCTP) were obtained from NEN Life Science (Boston, MA). A biotinylated dideoxyuridine-5′-triphosphate (Biotin-16-ddUTP) with a longer linker arm was purchased from Enzo Diagnostics, Inc. (Farmingdale, NY). Oligonucleotides were synthesized on a PerSeptive Biosystems Expedite nucleic acid synthesizer. Thermo Sequenase was from Amersham Pharmacia Biotech (Piscataway, NJ). Streptavidin-coated magnetic beads were obtained from Dynal Inc. (Oslo, Norway).

Sanger DNA-sequencing reaction

The structures of the biotinylated ddNTPs used for generating Sanger-sequencing fragments are shown in Figure 1. A synthetic template with the sequence 5′-ACTTTTTACTGTC-CGATCCCTGTATCTTAGAGCTCGCTATTCCGAGCTTA CACGT-3′, and the corresponding primer 5′-TAAGCTCG-GAAT-3′, were used to test the sequencing method using Biotin-11-ddNTPs. DNA-sequencing fragments for a synthetic template mimicking a portion of HIV protease gene, 5′-TAAAGCTATAGGTACAGTATTAGTAGGACCTACA-CCTGTCAA**G**ATAATGGTCCAGGTCGGT-3′, were generated using Biotin-16-ddUTP and Biotin-11-dd(A,C,G)TPs. The underlined G was also changed to a T and C to simulate a polymorphic site at this position. The corresponding primer sequence for the HIV synthetic template is 5′-ACCGACCT-GGACC-3′. Sanger-sequencing reactions consisted of 1.5 µl Thermo Sequenase reaction buffer (Amersham Pharmacia Biotech); 1125 pmol each of dATP, dTTP, dCTP, dGTP;
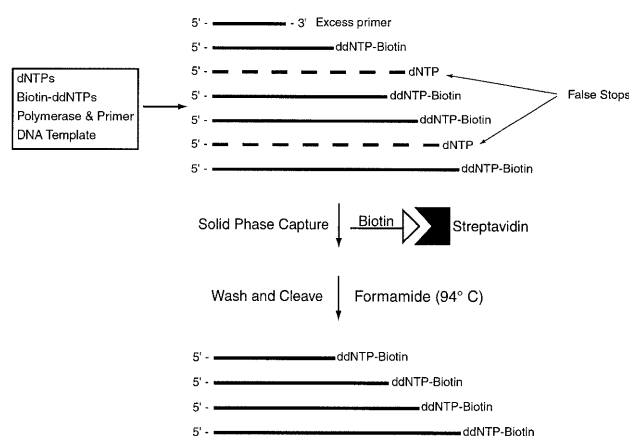

Figure 2. Scheme used to purify DNA-sequencing fragments for analysis by MS. False stops, excess primers and salts are eliminated from the reaction by capturing all correctly terminated DNA-sequencing fragments with streptavidin-coated magnetic beads. The purified fragments are then cleaved from the beads with 98% formamide for analysis by MALDI-TOF MS.

125 pmol each of Biotin-11-ddATP, Biotin-11-ddCTP, Biotin-11-ddGTP, and either Biotin-11-ddUTP or Biotin-16-ddUTP; 16 U of Thermo Sequenase; 30 pmol single-stranded DNA template; 200 pmol primer and 32 µl of water. The sequencing reactions were subjected to 25 cycles of 94°C for 20 s, 35°C for 40 s and 60°C for 1 min.

Solid-phase purification of DNA-sequencing products for mass spectrometry measurements

A schematic of the solid-phase purification of DNA-sequencing fragments is shown in Figure 2. Sanger sequencing reactions are used to produce the biotinylated DNA-sequencing fragments. To purify these fragments from false stops, excess primers and salts, the biotinylated DNA fragments are captured with streptavidin-coated magnetic beads. The beads are subsequently washed to remove all impurities. The purified DNA-sequencing fragments are then cleaved from the solid phase and analyzed by MS to produce accurate sequencing data. The detailed procedure is described as follows. DNA-sequencing products were combined with 70 µl streptavidin-coated magnetic beads that were resuspended in 34 µl of binding and washing (B/W) buffer (2 M NaCl, 0.5 mM Tris–HCl buffer, 1 mM EDTA, pH 7.0) and allowed to incubate for 20 min. The magnetic beads containing the biotinylated DNA fragments were subsequently washed twice with a modified B/W buffer (1 M NH_4Cl instead of 2 M NaCl), twice with 0.2 M triethyl ammonium acetate (TEAA) buffer and twice with deionized water. The magnetic beads were then resuspended in 8 µl of 98% formamide solution containing 2% 0.2 M TEAA buffer and were heated to 94°C for 5 min. This process denatures the binding between the biotin and streptavidin to release the biotinylated DNA-sequencing fragments from the magnetic beads. After removing the formamide by precipitation with EtOH, the pure biotinylated DNA-sequencing products were resuspended in 2 µl of water and 2 µl of matrix solution for MS analysis. The matrix solution was made by dissolving 35 mg 3-hydroxypicolinic acid (3-HPA) with 6 mg ammonium citrate in 350 µl 50% acetonitrile and then diluting
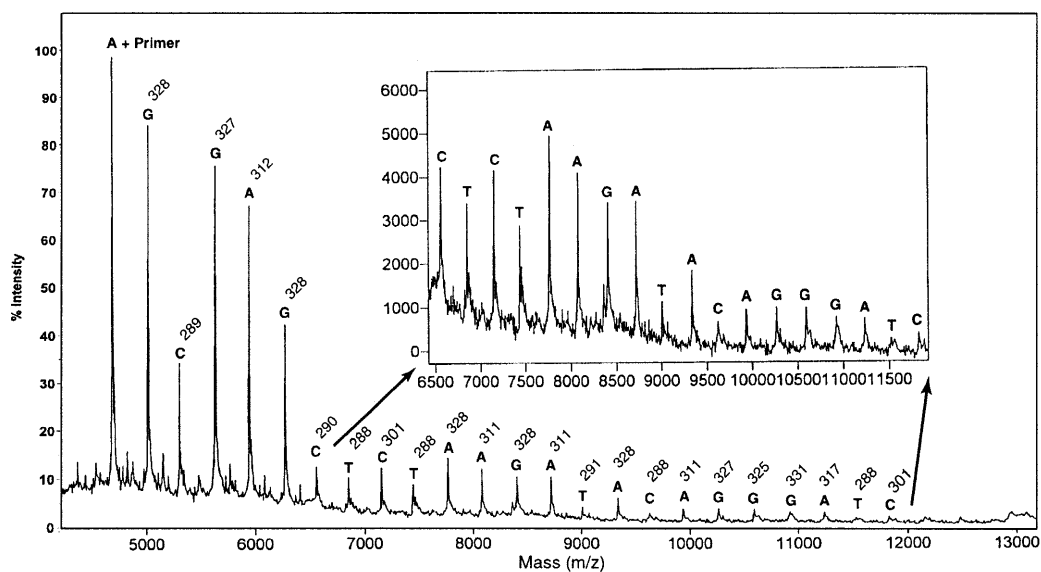
Figure 3. Mass-sequencing spectrum generated using Biotin-11-ddNTPs. The inset shows a magnification of the lower intensity region. The mass difference (indicated above each base) between a peak and the previous one is used to determine the identity of the nucleotide.

with 175 µl of water. All MS measurements were made on a PE Biosystems Voyager DE Pro MALDI-TOF mass spectrometer. One microliter of DNA matrix solution was spotted onto a stainless steel MALDI sample target plate. All measurements were taken in linear positive ion mode. A total accelerating voltage of 25 kV was used and typically between 100 and 200 shots were taken per spectra. The spectra were smoothed using the Voyager data analysis package.

RESULTS AND DISCUSSION

The structure of each of the five biotinylated dideoxynucleotides used in the DNA-sequencing reactions is shown in Figure 1. Biotin is attached to the five position of pyrimidines and the seven position of the purines through a linker arm. It has been shown previously that when these respective positions were modified with bulky organic dyes, DNA polymerase can still incorporate the modified nucleotides into the growing DNA strand (24). The linker arm consisting of 11–16 covalent bonds for attaching the biotin to the nucleotide was shown to be ideal for balancing capture efficiency by streptavidin and incorporation efficiency by polymerase (25,26).

The scheme for isolating pure biotinylated DNA-sequencing products is shown in Figure 2. This method was developed to efficiently eliminate salts and false stops from the DNA-sequencing products to produce clean MS spectra. Alkaline and alkaline earth salts can easily form adducts with DNA fragments that can interfere with accurate peak detection. It was shown previously that falsely terminated fragments created extra peaks in the mass-sequencing spectra interfering with accurate base identification (8). The use of biotinylated dideoxynucleotide terminators eliminates all falsely terminated fragments, salts and other components in the DNA-sequencing reaction.

As a first experiment, DNA-sequencing samples for MALDI-TOF MS analysis were prepared using four Biotin-11-ddNTPs. The resulting mass spectrum is shown in Figure 3. The first peak in the spectrum is the primer peak plus the first nucleotide that is complementary to the corresponding nucleotide in the DNA template. From here the difference in mass between each peak can be measured to determine the identity of the nucleotide corresponding to that peak, as each nucleotide (A, C, G, T) has a unique molecular weight. No primer peak is seen in the mass spectrum in Figure 3, since the primers are not biotinylated and are eliminated after the solid-phase capture. It was shown that excess primers could dimerize to form false peaks in the mass-sequencing spectra; thus, removing excess primers will produce cleaner sequencing data (8). There are no peaks due to false stops in the spectrum either. If standard terminators (ddNTPs) are used, peaks in the mass-sequencing spectrum due to some falsely stopped fragments can be indistinguishable from correctly terminated fragments. For instance, a fragment terminated falsely with a deoxy-adenine and a correctly terminated dideoxy-guanine residue would have the same mass. When solid-phase capturable terminators are used, these false stops are eliminated from the reaction mixture and cannot interfere with the obtained mass spectrum. The sequencing read length and the resolution in mass differences for this spectrum is similar to what has been reported previously (10).

As the mass difference between Biotin-11-ddUTP and Biotin-11-ddCTP is 1 Da, nucleotide T and C cannot be unambiguously differentiated in the mass spectrum in Figure 3. To resolve this issue, we examined the use of Biotin-16-ddUTP containing a longer linker arm to generate DNA-sequencing products for MS analysis. Use of Biotin-16-ddUTP and Biotin-11-dd(A,C,G)TPs shifts the mass difference between cytosine and thymidine to >80 Da, resulting in a better resolution in sequence determination. The mass-sequencing spectrum of
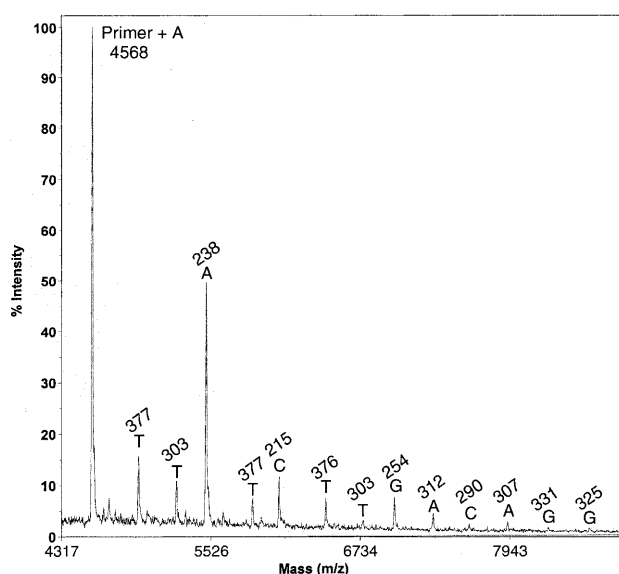
Figure 4. Mass-sequencing spectrum generated from Biotin-16-ddUTP and Biotin-11-dd(A,G,C)TPs. The spectrum is free of primers, primer adducts and false stops. Each base (A, C, G, T) is unambiguously determined due to the different molecular weights of the biotinylated dideoxynucleotide terminators.

Figure 5. Mass-sequencing spectrum generated from a DNA template containing a polymorphic site (T/C) at the sixth base after the primer to model a heterozygote. Inset is a magnified view of the heterozygote peak. The signal of each DNA fragment is very distinct and the heterozygote (A/G) can be easily recognized.

this experiment is shown in Figure 4. The result indicates that Biotin-16-ddUTP can be incorporated into the DNA fragments during the Sanger sequencing reactions to produce high-resolution mass-sequencing spectra. By using biotinylated terminators with different molecular weights, the smallest mass difference between any two sequencing fragments is 16 Da, whereas with standard terminators the smallest mass difference is only 9 Da. It has been shown that as the DNA fragment size increases, mass spectral peak width will increase, resulting in a diminished resolving capacity of the mass spectrometer for larger DNA fragments. Thus, using biotinylated terminators with increased mass differences for generating DNA-sequencing products will improve sequence identification by MS.

This increase in mass resolution also helps in heterozygote detection (27). DNA fragments with large mass typically have larger associated peak widths in their mass spectra. This makes heterozygote detection difficult if standard ddNTP terminators are used. We tested the ability to detect a heterozygote using biotinylated terminators [Biotin-11-dd(A,C,G)TPs and Biotin-16-ddUTP] and a synthetic template containing a polymorphic site (50% T and 50% C) at one specific position. As can be seen from the data in Figure 5, two distinct peaks with near equal intensity are generated in the mass-sequencing spectrum. This demonstrates that the use of biotinylated terminators with different molecular weights to generate Sanger sequencing fragments will allow distinct detection of heterozygotes. In a MS DNA-sequencing method reported previously, four reactions (one for each dideoxynucleotide) were separately performed. Using biotinylated terminators and capturing the DNA-sequencing fragments at the 3′-end of the fragments with a solid phase, we have developed a one-tube MS cycle-sequencing method. The use of solid-phase capturable
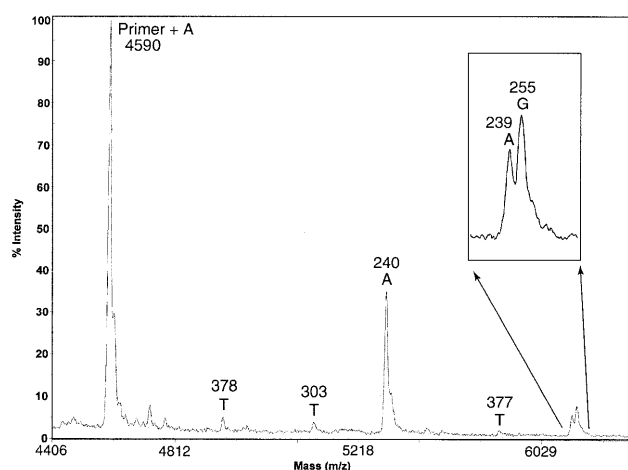
terminators, such as biotinylated dideoxynucleotides, provides a quick and easy method for generating clean Sanger sequencing products. This method can accurately detect SNPs and heterozygotes. With improvement in mass spectrometer detector sensitivity (28), this method will allow further improved detection of larger DNA fragments, and eventually will provide a platform for DNA sequencing unparalleled in speed and accuracy.

ACKNOWLEDGEMENTS

REFERENCES

1. Smith,L.M., Sanders,J.Z., Kaiser,R.J., Hughes,P., Dodd,C., Connell,C.R., Heiner,C., Kent,S.B.H. and Hood,L.E. (1986) Fluorescence detection in automated DNA sequencing analysis. *Nature*, **321**, 674–679.
2. Ju,J., Ruan,C., Fuller,C.W., Glazer,A.N. and Mathies,R.A. (1995) Energy transfer fluorescent dye-labeled primers for DNA sequencing and analysis. *Proc. Natl Acad. Sci. USA*, **92**, 4347–4351.
3. Ju,J., Glazer,A.N. and Mathies,R.A. (1996) Cassette labeling for facile construction of energy transfer fluorescent primers. *Nucleic Acids Res.*, **24**, 1144–1148.
4. Kheterpal,I., Scherer,J., Clark,S.M., Radhakrishnan,A., Ju,J., Ginther,C.L., Sensabaugh,G.F. and Mathies,R.A. (1996) DNA sequencing using a four-color confocal fluorescence capillary array scanner. *Electrophoresis*, **17**, 1852–1859.
5. Bowling,J.M., Bruner,K.L., Cmarik,J.L. and Tibbetts,C. (1991) Neighboring nucleotide interactions during DNA sequencing gel electrophoresis. *Nucleic Acids Res.*, **19**, 3089–3097.
6. Yamakawa,H. and Ohara,O. (1997) A DNA cycle sequencing reaction that minimizes compressions on automated fluorescent sequencers. *Nucleic Acids Res.*, **25**, 1311–1312.
7. Fitzgerald,M.C., Zhu,L. and Smith,L.M. (1993) The analysis of mock DNA sequencing reactions using matrix-assisted laser desorption/

ionization mass spectrometry. *Rapid Commun. Mass Spectrom.*, **7**, 895–897.

8. Roskey,M.T., Juhasz,P., Smirnov,I.P., Takach,E.J., Martin,S.A. and Haff,L.A. (1996) DNA sequencing by delayed extraction-matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Proc. Natl Acad. Sci. USA*, **93**, 4724–4729.

9. Taranenko,N.I., Allman,S.L., Golovlev,V.V., Taranenko,N.V., Isola,N.R. and Chen,C.H. (1998) Sequencing DNA using mass spectrometry for ladder detection. *Nucleic Acids Res.*, **26**, 2488–2490.

10. Kirpekar,F., Norhoff,E., Larsen,L., Kristiansen,K., Roepstorff,P. and Hillenkamp,F. (1998) DNA sequence analysis by MALDI mass spectrometry. *Nucleic Acids Res.*, **26**, 2554–2559.

11. Nordhoff,E., Luebert,C., Thiele,G., Heiser,V. and Lehrach,H. (2000) Rapid determination of short DNA sequences by the use of MALDI-MS. *Nucleic Acids Res.*, **28**, e86.

12. Monforte,J. and Becker,C. (1997) High-throughput DNA analysis by time-of-flight mass spectrometry. *Nature Med.*, **3**, 360–362.

13. Fu,D.J., Tang,K., Braun,A., Reuter,D., Darnhofer-Demar,B., Little,D.P., O'Donnell,M.J., Cantor,C.R. and Koster,H. (1998) Sequencing exons 5 to 8 of the p53 gene by MALDI-TOF mass spectrometry. *Nat. Biotechnol.*, **16**, 381–384.

14. Langer,P.R., Waldrop,A.A. and Ward,D.C. (1981) Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. *Proc. Natl Acad. Sci. USA*, **78**, 6633–6637.

15. Hawkins,T.L. (1992) M13 single-strand purification using a biotinylated probe and streptavidin coated magnetic beads. *DNA Seq.*, **3**, 65–69.

16. Hawkins,T.L., O'Connor-Morin,T., Roy,A. and Santillan,C. (1994) DNA purification and isolation using a solid-phase. *Nucleic Acids Res.*, **22**, 4543–4544.

17. Uhlen,M. (1989) Magnetic separation of DNA. *Nature*, **340**, 733–734.

18. Hultman,T., Bergh,S., Moks,T. and Uhlen,M. (1991) Bidirectional solid-phase sequencing of *in vitro*-amplified plasmid DNA. *Biotechniques*, **10**, 84–93.

19. Tong,X. and Smith,L.M. (1992) Solid-phase method for the purification of DNA sequencing reactions. *Anal. Chem.*, **64**, 2672–2677.

20. Tong,X. and Smith,L.M. (1993) Solid-phase purification in automated DNA sequencing, *DNA Seq.*, **4**, 151–162.

21. Ju,J. (1999) US Patent no. 5 876 936.

22. Tang,K., Fu,D., Julien,D., Braun,A., Cantor,C.R. and Koster,H. (1999) Chip-based genotyping by mass spectrometry. *Proc. Natl Acad. Sci. USA*, **96**, 10016–10020.

23. Roses,A. (2000) Pharmacogenetics and the practice of medicine. *Nature*, **405**, 857–865.

24. Prober,J., Trainor,G., Dam,R., Hobbs,F., Robertson,C., Zagursky.R., Cocuzza,A., Jensen,M. and Baumeister,K. (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, **238**, 336–341.

25. Gebeyehu,G., Rao,P., SooChan,P., Simms,D.A. and Klevan,L. (1987) Novel biotinylated nucleotide—analogs for labeling and calorimetric detection of DNA. *Nucleic Acids Res.*, **15**, 4513–4534.

26. Herman,T., Lefever,E. and Shimkus,M. (1986) Affinity chromatography of DNA labeled with chemically cleavable biotinylated nucleotide analogs. *Anal. Biochem.*, **156**, 48–55.

27. Fei,Z., Ono,T. and Smith,L.M. (1998) MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs. *Nucleic Acids Res.*, **26**, 2827–2828.

28. Hilton,G.C., Martinis,J.M., Wollman,D.A., Irwin,K.D., Dulcie,L.L., Gerber,D., Gillevet,P.M. and Twerenbold,D. (1998) Impact energy measurement in time-of-flight mass spectrometry with cryogenic microcalorimeters. *Nature*, **391**, 672–675.