JOURNAL OF MASS SPECTROMETRY, VOL. 33, 187–192 (1998)

Separate Analysis of Complementary Strands of Restriction Enzyme-digested DNA. An Application of Restriction Fragment Mass Mapping by Matrix-assisted Laser Desorption/Ionization Mass Spectrometry

Yoshinao Wada*
Department of Molecular Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, 840 Murodo-cho, Izumi, Osaka 590–02, Japan

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI/TOF-MS) of a restriction endonuclease digest determines the molecular mass of PCR-amplified DNA more easily than measurement of undigested DNA. With this method, a 664 bp region from the FAS gene could be analyzed and a two-nucleotide deletion in the L1CAM gene was detected in a restriction fragment of 105 nucleotides. Furthermore, the analysis of smaller fragments allowed separate detection of single-stranded oligonucleotides comprising individual digested fragments. This mixture analysis of restriction enzyme digests improves the resolution, sensitivity and accuracy of MALDI/TOF-MS of DNA and is thus expected to facilitate its application to genetic diagnosis.

J. Mass Spectrom. **33**, 187–192 (1998)

KEYWORDS: matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; DNA; mapping; polymerase chain reaction; restriction endonuclease

INTRODUCTION

Electrospray or matrix-assisted laser desorption is now capable of ionizing large intact molecules of DNA.[1–19] This advance has opened the way for DNA sequencing and mutation detection using mass spectrometry.[20–31] However, the large molecular size of DNA, which is composed of an ~300 Da nucleotide unit, presents a major obstacle. In contrast, an amino acid residue is much smaller, between 57 and 186 Da for glycine and tryptophan, respectively, and the detection of a single amino acid replacement in, for example, a 16 kDa protein such as globin is feasible.[32] A total of 146 amino acids constituting human β-globin are encoded by 438 nucleotides, the molecular mass of which exceeds 130 kDa and is thus too large for detection of even the maximum mass difference between nucleotides, i.e. the 40 units between guanosine and cytosine, in the intact DNA molecule.

Peptide mapping is a basic method designed to elucidate primary protein structures. It is also applied to localizing mutations and other structural changes to a discrete segment of the amino acid sequence. Taking advantage of its feasibility for mixture analysis, mass spectrometry has been applied to peptide mapping and has become a powerful tool for characterizing structural changes and modifications of proteins[33]

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI/TOF-MS) was introduced into the analysis of restriction enzyme-digested DNA shortly after the use of this ionization method became widespread[11,13,20–23,30] The present study demonstrates that this combination is an effective means for improving the sensitivity and accuracy of MALDI/TOF-MS of DNA and thus of facilitating its application to genetic diagnosis. An example of the resolved detection of individual complementary strands composing a fragment of DNA by MALDI/TOF is also presented.

EXPERIMENTAL

PCR and sample preparation

The amplification target was a 664 bp (base pair) segment in the coding region of the FAS gene[34], a 278bp segment involving exon 18 of the L1CAM gene[35–37] and a 50 bp segment in exon 2 of the human β-globin gene. PCR was carried out in 50 µl of solution containing 10 mM Tris–HCl at pH 9.0, 50 mM KCl, 0.1% Triton X-100, 1.6 mM $MgCl_2$, 0.2 mM each of dATP, dGTP, dCTP and dTTP, 0.4 mM of each primer and 2 U of r*Taq* DNA polymerase. The amplification template for the FAS gene was a cDNA constructed by

* Correspondence to: Y. Wada, Department of Molecular Medicine, Osaka Medical Center, 840 Murodo-cho, Izumi, Osaka 590-02, Japan.

CCC 1076–5174/98/020187–06 $17.50

Received 13 August 1997
Accepted 1 October 1997

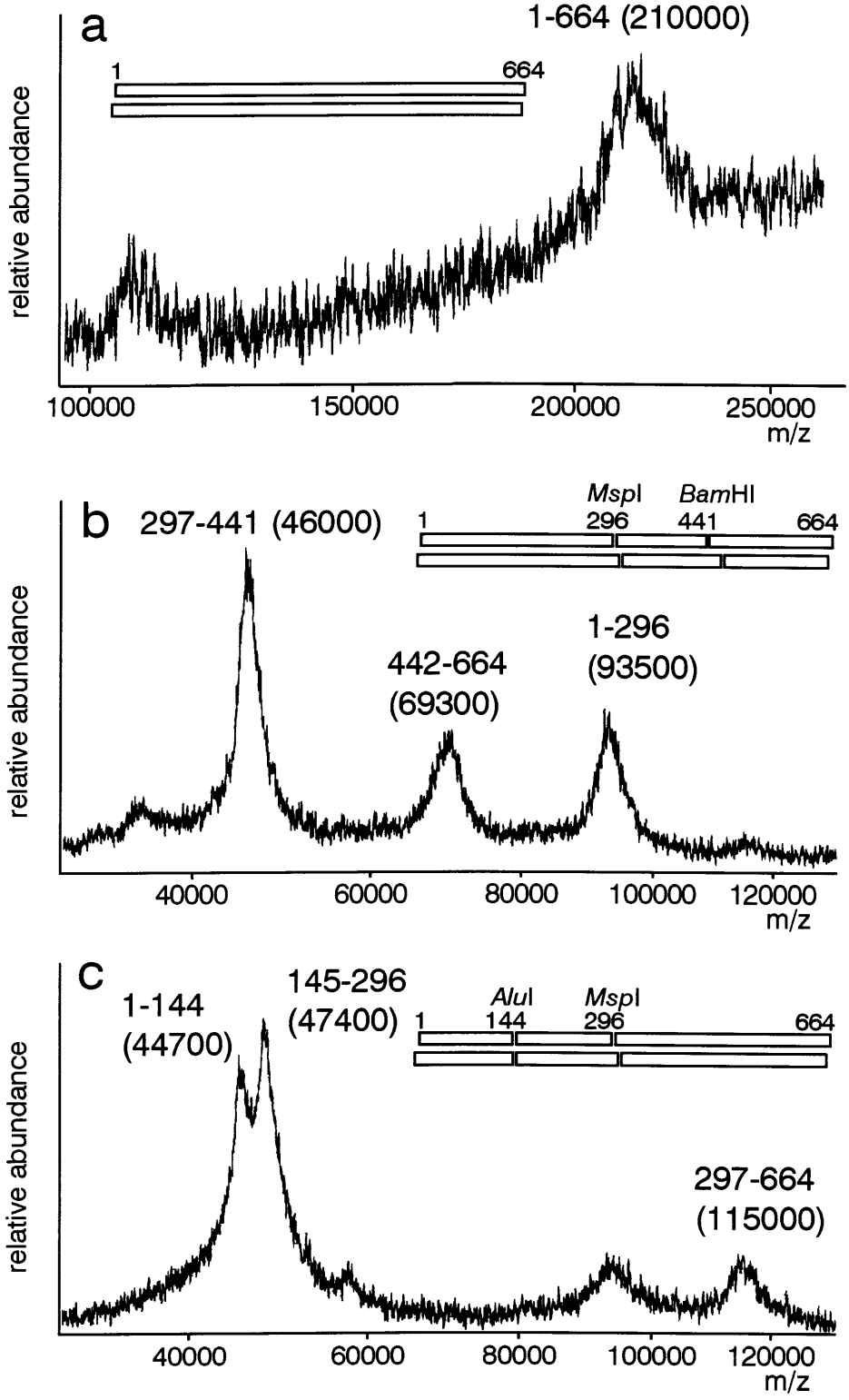

Figure 1. Positive ion MALDI/TOF mass spectra of a 664 bp DNA amplified from the FAS gene. Sequence numbers are given above the peaks with the measured values in parentheses. Cleavage sites are illustrated. (a) Undigested PCR product. The doubly charged ions are observed at *m/z* 110 000. (b) Digest produced by restriction endonucleases *Msp*I and *Bam*HI. (c) Digest produced by *Alu*I and *Msp*I. The peak at *m/z* 93 000 is probably an uncleaved DNA corresponding to sequence 1–296.

reverse transcription of the mRNA from normal human fibroblasts and those for the L1CAM and globin genes were genomic DNA extracted from peripheral leukocytes.

After 30 amplification cycles, the product DNA was extracted with phenol and precipitated with ethanol in the presence of 2 M ammonium acetate. The pellet was dried after washing with 70% ethanol. The purified DNA was dissolved in 5 μl of distilled water for mass spectrometry or subjected to digestion.

Enzymatic hydrolysis of DNA was carried out with a combination of restriction endonucleases; *Bam*HI/*Msp*I

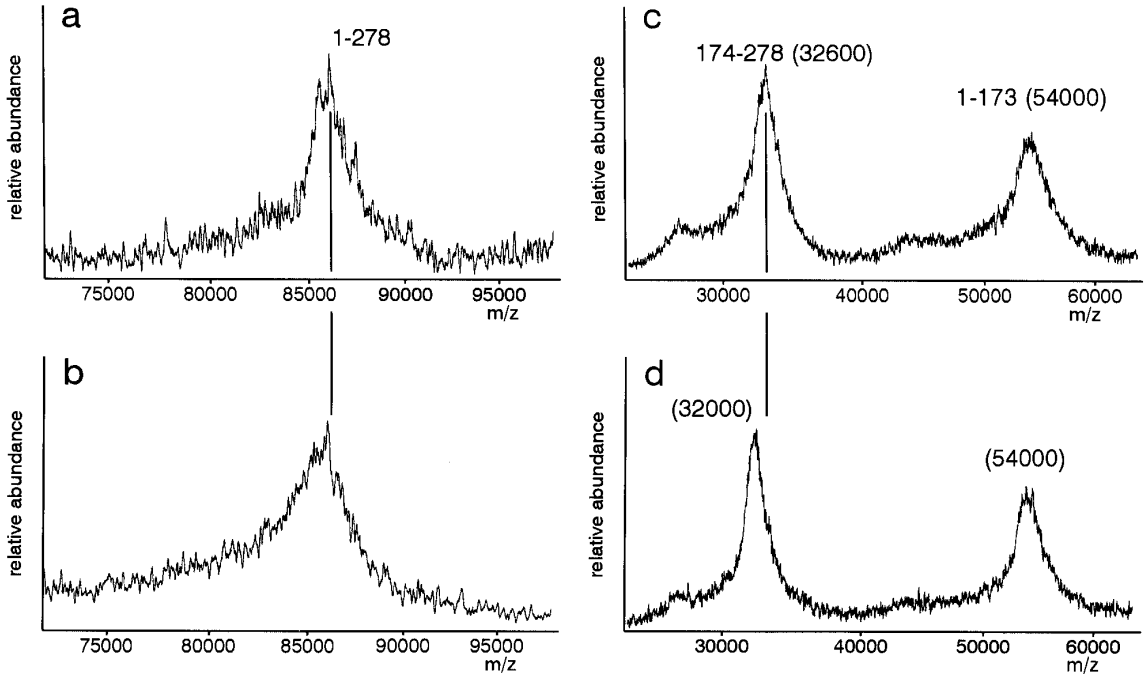

Figure 2. Positive ion MALDI/TOF mass spectra of a 278 bp DNA amplified from the L1CAM gene. Sequence numbers are given above the peaks with the measured values in parentheses. (a) Undigested PCR product from a normal control. (b) Undigested product from a patient with congenital hydrocephalus who has a two nucleotide deletion. (c) *Hin*cII digest of the 278 bp product from a normal control. (d) *Hin*cII digest of the amplified product from the patient. Vertical lines indicate the positions of normal peaks.

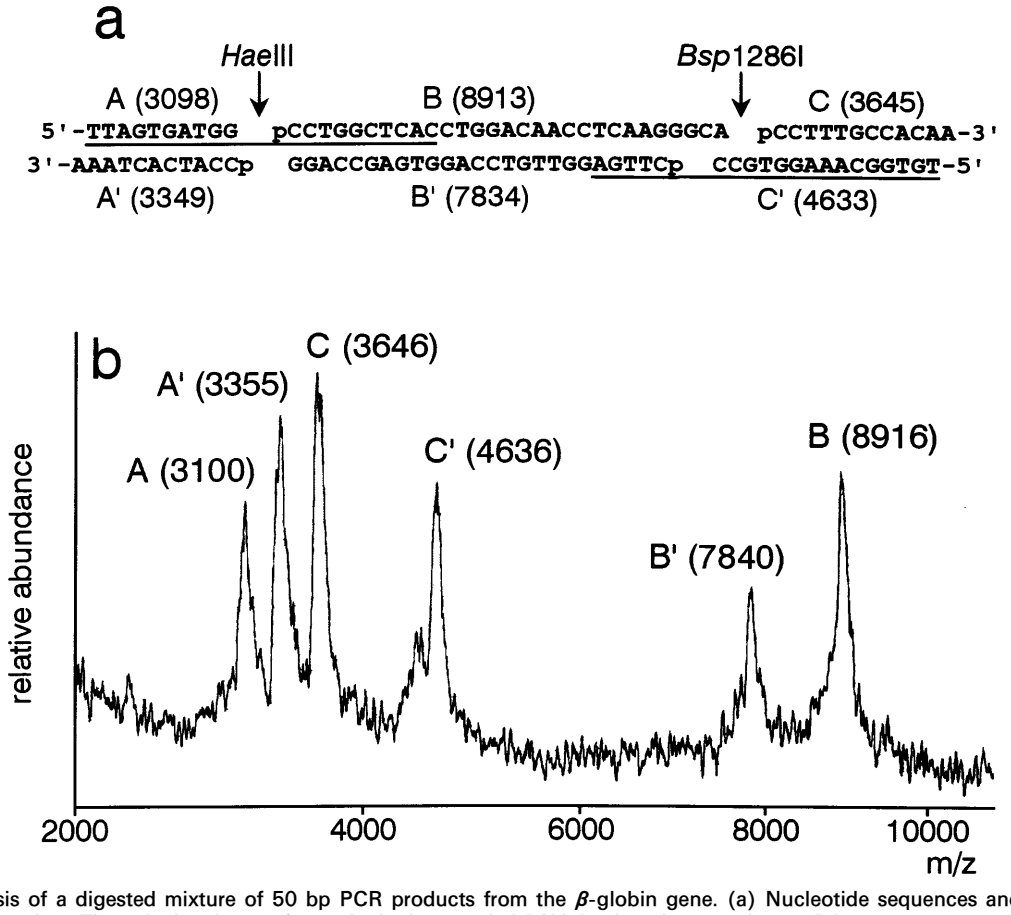

Figure 3. Analysis of a digested mixture of 50 bp PCR products from the β-globin gene. (a) Nucleotide sequences and the *Hae*III and *Bsp*1286I cleavage sites. The calculated mass for each single-stranded DNA is given in parentheses. Primer sequences are underlined. (b) Positive ion MALDI/TOF mass spectrum of the digest. The numbers in parentheses are the measured values for the peaks.

or *Alu*I/*Msp*I for FAS, with *Hin*cII for L1CAM and *Hae*III/*Bsp*1286I for globin. Typically, 1 pmol of DNA was incubated with 1 U of enzyme at 37 °C for 3 h. The digested DNA was extracted with phenol and precipitated with ethanol as described above.

Mass spectrometry

For MALDI/TOF-MS, 0.3 μl of the sample solution was mixed with the same volume of a 30 mg ml^{-1} solution of 3-hydroxypicolinic acid[9] on the probe tip of a Vision 2000 MALDI/TOF mass spectrometer (Thermo BioAnalysis, Hemel Hempstead, UK) equipped with a reflectron and a nitrogen laser (337 nm). Delayed extraction was not used. The sample amount loaded was estimated to be 0.1 pmol. The molecular mass was expressed as the chemical mass in the present study.

RESULTS

Analyses of a 664 bp product of the FAS gene

The calculated masses for a 664 bp segment of the FAS gene were 205 966/204 780 for sense/antisense strands with dA overhand at the 3′-ends[38]. In the MALDI/TOF mass spectrum, a broad signal corresponding to these DNA molecules was detected at m/z 210 000 with low intensity [Fig. 1(a)]. The PCR product was then digested with restriction endonucleases *Msp*I and *Bam*HI which cleave at positions 296 and 441, respectively. The MALDI/TOF mass spectrum of the digest revealed discrete molecular ion signals for three fragments, although the complementary strands were not resolved in each peak [Fig. 1(b)]. The signal for the fragment observed at m/z 46 000 was sharp compared with the larger one at m/z 93 500 and their peak widths at half-height were 2500 and 4000 units, respectively. Such peak broadening for the latter fragment was not ascribed to the mass difference between the two complementary strands, because the calculated difference was somewhat larger in the former, 699 of 44 770/45 469 vs 21 of 91 879/91 900 for the latter.

Subsequently, in order to cleave the largest fragment corresponding to sequence 1–296, the original PCR product was digested with *Alu*I and *Msp*I, which cleave at positions 144 and 296, respectively. The derived fragments were very clear at m/z 44 700 and 47 400 [Fig. 1(c)]. The calculated molecular masses were 44 590/44 653 and 47 307/47 265, respectively.

Detection of an L1CAM gene mutation

L1CAM is the gene responsible for X-linked congenital hydrocephalus.[36] A normally 278 bp region involving exon 18 of the L1CAM gene was amplified and analyzed by MALDI/TOF-MS. In the boy patient analyzed, two nucleotides at positions 221 and 222 were previously shown to be missing.[37] The molecular ion signal for the single-stranded DNA of the amplified product was observed, but the 0.7% molecular mass decrease corresponding to the deletion was barely detectable [Fig. 2(a) and (b)]. A mixture from a restriction endonuclease *Hin*cII digestion, i.e. cleaved at position 173, was analyzed. The deletion producing a 1.8%, or 600 units, decrease was clearly identified in a fragment corresponding to the region of sequence 174–278 [Fig. 2(c) and (d)].

Separation of complementary DNA strands

The two complementary oligonucleotides of DNA are ionized separately under the present conditions,[11] but their molecular ion signals usually overlap and are thus indistinguishable from each other.

When a 50 bp product of the β-globin gene was digested with *Hae*III and *Bsp*1286I, three restriction fragments were generated [Fig. 3(a)]. In the MALDI/TOF mass spectrum of the digest, there were six independent molecular ion signals corresponding to the single-stranded DNA component of these fragments [Fig. 3(b)]. Since *Bsp*1286I produced protruding cohesive termini, the complementary single-stranded DNAs were ~1000 U apart in both fragments B/B′ and fragment C/C′.

DISCUSSION

MALDI/TOF-MS of a mixture of restriction enzyme-digested DNA was first reported by Tang *et al.*,[11] who analyzed a mixture of fragments, all of which were less than 100 bp. The strategy was then extended to more complex mixtures[13,20] or was applied to the detection of mutations of three or four nucleotides and deletion/insertion of a <100 bp fragment.[22,30] Here, this strategy has been extended, allowing a clear separation of individual single-stranded DNAs comprising the fragments and detection of a two-base deletion in a 105 bp fragment. As is clearly demonstrated, mixture analysis of restriction enzyme-digested DNA or restriction fragment mass mapping is expected to overcome some of the problems encountered in the MALDI/TOF-MS of PCR-amplified DNA. Dealing with a number of digested DNA fragments rather than a large intact molecule confers, in principle, advantages similar to those of peptide mass mapping, although the two methods differ in some respects. These similarities and differences can be summarized as follows.

Advantages common to peptide mass mapping and restriction fragment mass mapping

First, good intensity and a high signal-to-noise ratio are more easily achieved than with the measurement of large undigested DNA molecules. Indeed, the signals of small fragments were intense and sharp compared with those of intact DNA or larger fragments, on analysis of the FAS gene (Fig. 1). Second, accurate and precise molecular mass determinations are relatively easy for

smaller molecules. Precise analysis is crucial for identifying and characterizing minute structural changes such as mutation and modification and is dependent upon both the peak width of the molecular ion signal and the resolution of the measurement. The peak width is primarily determined by the isotopic distribution and becomes larger with increasing molecular size. In addition, the molecular ion signals for large molecules tend to become broader than their theoretical distribution, probably owing to adduction of salts or to metastable decay induced by the MALDI process. This is remarkable in the MALDI/TOF-MS of DNA, as reported by Zhu *et al.*,[39] who described the resolution for DNA over 100-mer as being $\sim$100 even with delayed extraction unless advanced purification procedures are applied. In the present experiment, the resolution was 200 for the fragments of $<$30-mer [Fig. 3(b)], but 20 for the 150-mer DNA of an *Alu*I/*Msp*I digest of the FAS gene [Fig. 1(c)]. It is noteworthy, however, that even with such low resolution, a deletion-type mutation was detectable in a fragment exceeding 100 bp [Fig. 2(d)].

Third, digestion of PCR products improves the resolution in another way. *Taq* DNA polymerase has terminal deoxynucleotidyl transferase (TdT) activity which adds an additional single nucleotide which is usually, but not always, dA at the 3′-terminus. This activity leads to structural heterogeneity among the amplified molecules. Analysis of the restriction enzyme-digested fragments resolves this problem, because the cleaved 3′-terminal end is structurally discrete.

Differences between peptide mass mapping and restriction fragment mass mapping

Restriction fragment mass mapping differs from peptide mass mapping in a few important respects. As described above, the molecular sizes of nucleic acids are larger than those of amino acids. This presents a major obstacle in DNA analysis. However, a great variety of restriction enzymes, with different specificities, are available for DNA compared with only about 10 specific peptidases for protein. Furthermore, cleavage utilizing restriction enzymes is unaffected by the specific sequence adjacent to the cleavage site, unlike the case of proteins in which adjacent sequences may confer exceptional resistance to peptidase digestion.

Chemical properties, such as hydrophobicity, of peptides determine the intensity of molecular ion signals. In contrast, such properties are mostly homogeneous in DNA and the relative abundance of ions in the mass spectrum is primarily dependent on the size and amount of DNA molecule. This homogeneity facilitates the detection of component fragments covering the entire DNA sequence.

In the case of identifying a substitution mutation, the molecular mass shift of a mutated peptide, which can be detected and measured by peptide mass mapping, often specifies the site and type of mutation. Since there are only four nucleic acids in DNA, this strategy is rarely applicable to restriction fragment mass mapping, the exception being a mutation present at the cleavage site of a restriction enzyme.[20,21]

CONCLUSION

Restriction fragment mass mapping is expected to broaden the application of MALDI/TOF-MS to DNA. The separate detection of complementary strands presented in Fig. 3(b) illustrates the increased reliability of this strategy in genetic diagnosis.

REFERENCES

1. T. R. Covey, R. F. Bonner, B. I. Shushan and J. Henion, *Rapid Commun. Mass Spectrom.* **2**, 246 (1988).
2. J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong and C. M. Whitehouse, *Science* **246**, 64 (1989).
3. R. D. Smith, J. A. Loo, C. G. Edmonds, C. J. Barinaga and H. R. Udseth, *Anal. Chem.* **62**, 882 (1990).
4. B. Spengler, Y. Pan, R. J. Cotter and L. S. Kan, *Rapid Commun. Mass Spectrom.* **4**, 99 (1990).
5. F. Hillenkamp, M. Karas, R. C. Beavis and B. T. Chait, *Anal. Chem.* **63**, 1193A (1991).
6. K. Tang, S. L. Allman and C. H. Chen, *Rapid Commun. Mass Spectrom.* **6**, 365 (1992).
7. D. M. Schieltz, C.-W. Chou, C.-W. Luo, R. M. Thomas and P. Williams, *Rapid Commun. Mass Spectrom.* **6**, 631 (1992).
8. K. Tang, S. L. Allman, R. B. Jones, C. H. Chen and S. Araghi, *Rapid Commun. Mass Spectrom.* **7**, 63 (1993).
9. K. J. Wu, A. Steding and C. H. Becker, *Rapid Commun. Mass Spectrom.* **7**, 142 (1993).
10. K. Tang, S. L. Allman and C. H. Chen, *Rapid Commun. Mass Spectrom.* **7**, 943 (1993).
11. K. Tang, S. L. Allman, C. H. Chen, L. Y. Ch'ang and M. Schell, *Rapid Commun. Mass Spectrom.* **8**, 183 (1994).
12. K. Tang, N. I. Taranenko, S. L. Allman, C. H. Chen, L. Y. Ch'ang and K. B. Jacobson, *Rapid Commun. Mass Spectrom.* **8**, 673 (1994).
13. J. Bai, Y.-H. Liu, D. M. Lubman and D. Siemieniak, *Rapid Commun. Mass Spectrom.* **8**, 687 (1994).
14. K. Tang, N. I. Taranenko, S. L. Allman, L. Y. Ch'ang and C. H. Chen, *Rapid Commun. Mass Spectrom.* **8**, 727 (1994).
15. R. Chen, X. Cheng, D. W. Mitchell, S. A. Hofstadler, Q. Wu, A. L. Rockwood, M. G. Sherman and R. D. Smith, *Anal. Chem.* **67**, 1159 (1995).
16. Y. Naito, K. Ishikawa, Y. Koga, T. Tsuneyoshi, H. Terunuma and R. Arakawa, *Rapid Commun. Mass Spectrom.* **9**, 1484 (1995).
17. S. D. Fuerstenau and W. H. Benner, *Rapid Commun. Mass Spectrom.* **9**, 1528 (1995).
18. X. Cheng, D. G. Camp, II, Q. Wu, R. Bakhtiar, D. L. Springer, B. J. Morris, J. E. Bruce, G. A. Anderson, C. G. Edmonds and R. D. Smith, *Nucleic Acids Res.* **24**, 2183 (1996).
19. G. B. Hurst, M. J. Doktycz, A. A. Vass and M. V. Buchanan, *Rapid Commun. Mass Spectrom.* **10**, 377 (1996).
20. Y.-H. Liu, J. Bai, X. Liang, D. M. Lubman and P. J. Venta, *Anal. Chem.* **67**, 3482 (1995).
21. Y.-H. Liu, J. Bai, Y. Zhu, X. Liang, D. Siemieniak, P. J. Venta and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **9**, 735 (1995).
22. L.-Y. Ch'ang, K. Tang, M. Schell, C. Ringelberg, K. J. Matteson, S. L. Allman and C. H. Chen, *Rapid Commun. Mass Spectrom.* **9**, 772 (1995).
23. J. Bai, Y.-H. Liu, X. Liang, Y. Zhu and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **9**, 1172 (1995).
24. C. M. Bentzley, M. V. Johnston, B. S. Larsen and S. Gutteridge, *Anal. Chem.* **68**, 2141 (1996).

25. M. T. Roskey, P. Juhasz, I. P. Smirnov, E. J. Takach, S. A. Martin and L. A. Haff, *Proc. Natl. Acad. Sci. USA* **93**, 4724 (1996).
26. I. P. Smirnov, M. T. Roskey, P. Juhasz, E. J. Takach, S. A. Martin and L. A. Haff, *Anal. Biochem.* **238**, 19 (1996).
27. N. I. Taranenko, C. N. Chung, Y. F. Zhu, S. L. Allman, V. V. Golovlev, N. R. Isola, S. A. Martin, L. A. Haff and C. H. Chen, *Rapid Commun. Mass Spectrom.* **11**, 386 (1997).
28. T. Tsuneyoshi, K. Ishikawa, Y. Koga, Y. Naito, S. Baba, H. Terunuma, R. Arakawa and D. J. Prockop, *Rapid Commun. Mass Spectrom.* **11**, 719 (1997).
29. Y. F. Zhu, N. I. Taranenko, S. L. Allman, N. V. Taranenko, S. A. Martin, L. A. Haff and C. H. Chen, *Rapid Commun. Mass Spectrom.* **11**, 897 (1997).
30. J. R. Srinivasan, Y.-H. Liu, P. J. Venta, D. Siemieniak, A. A. Killeen, Y. Zhu and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **11**, 1144 (1997).
31. Y. Naito, K. Ishikawa, Y. Koga, T. Tsuneyoshi, H. Terunuma and R. Arakawa, *J. Am. Soc. Mass Spectrom.* **8**, 737 (1997).
32. Y. Wada, *Biol. Mass Spectrom.* **21**, 617 (1992).
33. Y. Wada, T. Matsuo and T. Sakurai, *Mass Spectrom. Rev.* **8**, 379 (1989).
34. N. Itoh, S. Yonehara, A. Ishii, M. Yonehara, S. Mizushima, M. Sameshima, A. Hase, Y. Seto and S. Nagata, *Cell* **66**, 233 (1991).
35. M. L. Hlavin and V. Lemmon, *Genomics* **11**, 416 (1991).
36. M. Jouet, A. Rosenthal, G. Armstrong, J. MacFarlane, R. Stevenson, J. Paterson, A. Metzenberg, V. Ionasescu, K. Temple and S. Kenwrick, *Nature Genet.* **7**, 402 (1994).
37. N. Okamaoto, Y. Wada and M. Goto, *J. Med. Genet.* in press.
38. M. Clark, *Nucleic Acids Res.* **16**, 9677 (1988).
39. Y. Zhu, L. He, J. R. Srinivasan and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **11**, 987 (1997).

