Genotyping of Apolipoprotein E by Matrix-assisted Laser Desorption/Ionization Time-of-flight Mass Spectrometry

Jannavi R. Srinivasan,[1] Maureen T. Kachman,[1] Anthony A. Killeen,[2] Nahida Akel,[2] David Siemieniak[3] and David M. Lubman[1]*

[1]Department of Chemistry, University of Michigan, Ann Arbor MI-48109, USA
[2]Department of Pathology, University of Michigan Medical Center, Ann Arbor MI-48109, USA
[3]Howard Hughes Medical Institute, University of Michigan, Ann Arbor MI-48109, USA

The genotyping of the various isoforms of Apolipoprotein E (apo E) has been performed using matrix-assisted laser desorption/ionization (MALDI-MS). The polymerase chain reaction was used to amplify the specific apo E gene sequence followed by digestion with Cfo I (*Clostridium formicoaceticum*), for generating restriction fragments for rapid and accurate mass analysis. An exonuclease I digestion step was introduced to remove the unused primers after PCR, which can otherwise interfere in the mass spectral analysis. By replacing the gel electrophoresis detection step with MALDI-MS, restriction isotyping of the apo E gene was achieved. Genotyping of an unknown sample obtained from an independent diagnostic laboratory demonstrated the validity of the MALDI-MS method for the routine clinical analysis of apo E. © 1998 John Wiley & Sons, Ltd.

Apolipoprotein E (apo E) plays an important role in lipid metabolism and also appears to be involved in cholesterol transport between different tissues. The apo E gene is polymorphic, with three common alleles ($\varepsilon2$, $\varepsilon3$, $\varepsilon4$) that control the expression of six phenotypes, namely, E2/2, E3/3, E4/4 homozygotes and E2/3, E2/4 and E3/4 heterozygotes. These isoforms differ from each other by a cysteine–arginine interchange that occurs at one or both of two positions (112 and 158) in the amino acid sequence. Apo E3 is the most common isoform (50–70%) in the general population whereas apo E2 and apo E4 isoforms typically contribute 8–10% and 10–15%, respectively to the gene pool.[1] Studies have shown that individuals with the apo E2 isoform have a lower than average cholesterol level, while those with the E4 isoform have a higher than average cholesterol level, and may be at a higher risk of developing atherosclerosis.[2]

The apo E polymorphism is an important risk factor in the development of Alzheimer's and cardiovascular diseases. Alzheimer's disease is a complex, heterogeneous, genetic disorder which involves the degeneration of neurons in the cerebral vessels causing dementia in the elderly. The $\varepsilon4$ allele has been shown to be associated with increased risk for the onset of Alzheimer's disease in many populations throughout the world.[3] Apo E polymorphism is also associated with Type III hyperlipoproteinemia, a disease causing the accumulation of triglycerides and cholesterol rich lipoproteins in the plasma which may contribute to the development of premature atherosclerosis.[4] Studies have shown a high percentage (90%) of the patients having the homozygote phenotype E2.

Apolipoprotein E genotyping is a very important diagnostic test. Apo E polymorphisms have been widely studied in numerous research laboratories and a number of methods to type the different isoforms have been developed. Apo E phenotyping by isoelectric focusing[5,6] and two dimensional gel electrophoresis[7,8] are two common methods of studying the protein. Some of the methods of genotyping that are being pursued include single strand conformation polymorphism (SSCP),[9] hybridization with allele specific oligonucleotides (ASO),[10,11] restriction isotyping with specific endonucleases[12,13] and sequencing.[14] All of these methods involve the amplification of the apo E genomic sequence containing the common polymorphic sites. Restriction isotyping is a commonly accepted method for genotyping of apolipoprotein E. This method involves the determination of apo E genotype by exploiting the specificity of a restriction endonuclease for certain regions of the amplified sequence. A major drawback of each of these techniques is the time required for the detection of the DNA by gel electrophoresis. Other disadvantages include the inability to detect small DNA fragments (≤ 40 bp) with ethidium bromide staining. Acrylamide gels, which are commonly used to detect small DNA fragments, contain a potential neurotoxin. In addition, because gel electrophoresis is not easily automated, only a relatively small number of samples can be analyzed in a day.[13] Hence a rapid, efficient and yet hazard free detection technique is needed.

In this work, MALDI-MS is used for genotyping of the apo E gene. MALDI-MS has been shown as a possible replacement for gel electrophoresis detection in DNA genetic screening studies. For example, MALDI-MS has been highly successful as a detection tool in DNA screening studies of the Tay–Sachs gene,[15] CFTR gene[16–18] and in the detection of single and two nucleotide mutations in the L1CAM gene.[19,20] MALDI-MS has also been used for the detection of specific regions of PCR amplified bacterial

*Correspondence to: D. M. Lubman, Department of Chemistry, University of Michigan, Ann Arbor MI-48109, USA.

```
apo E sequence

allele
ε2
ε3
ε4    5'-taagcttG GCA CGG CTG TCC AAG GAG CTG CAG GCG GCG CAG GCC CGG CTG GGC GCG GAC
             (primer 1 sequence)                          ↑                        ↑

ε2                      T
ε3                      T
ε4    ATG GAG GAC GTG CGC GGC CGC CTG GTG CAG TAC CGC GGC GAG GTG CAG GCC ATG CTC GGC
                        ↑

ε2
ε3
ε4    CAG AGC ACC GAG GAG CTG CGG GTG CGC CTC GCC TCC CAC CTG CGC AAG CTG CGT AAG CGG
                                    ↑                        ↑

ε2                                              T
ε3                                              C
ε4    CTC CTC CGC GAT GCC GAT GAC CTG CAG AAG CGC CTG GCA GTG TAC CAG GCC GGG GCgaat
                                            ↑         3'- CAC ATG GTC CGG CCC CGctta
                                                              (primer 2 sequence)
ε2
ε3
ε4    tctgttatttatattttatatttat
      agacaataaatataaaatataaata
```
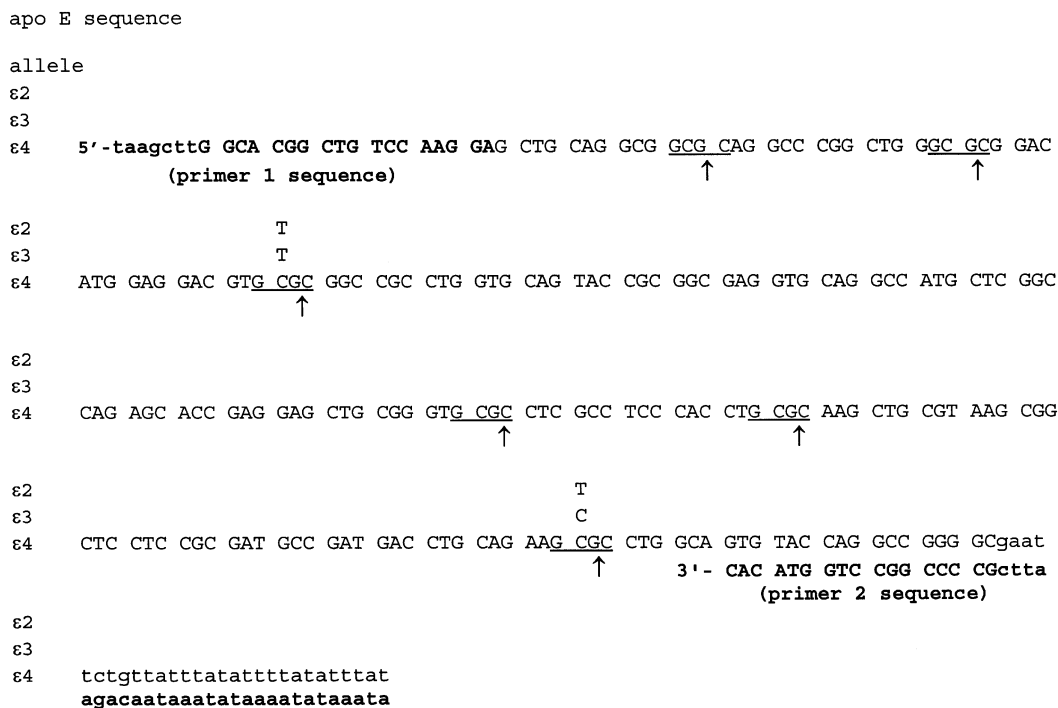
Figure 1. DNA sequence of PCR amplified apolipoprotein E and the locations of Cfo I cleavage sites. The sequence of one strand of the 264 bp amplified product of the ε4 allele is shown, and the variations due to ε2 and ε3 isoforms are shown above the ε4 sequence. ↑ represents the cut sites for the Cfo I restriction enzyme and the recognized sequence, GCGC, is underlined throughout. The primer sequences are shown in bold. Lower case letters represent bases that do not appear in the apo E genomic sequence, but were added to the primers as primer 'tails'.[12,13]

genome in a study involving Legionella species.[21] The use of peptide nucleic acid probes as DNA mimics for the MALDI-MS typing of human leukocyte antigen (HLA) has recently been reported by Girard and co-workers as a new MALDI-MS application for genetic screening problems.[22] More recently, Little and co-workers have demonstrated MALDI-MS as a viable technique for detecting apo E polymorphisms using PROBE reactions.[23] In this procedure, sequence specific primers are extended by a series of bases unique to the identity of a particular allele site, producing low mass products for MALDI-MS detection.

In this present work we demonstrate that MALDI-MS can be used as a routine detection method in a typical diagnostic lab where restriction isotyping is performed on a regular basis and the products generated are detected by gel electrophoresis. Here, the polymerase chain reaction (PCR) is used to generate the initial sequence of apo E gene for our experiments and the molecular weight of the isolated and amplified DNA is studied by MALDI-MS. The amplified genomic DNA is subjected to Exonuclease I digestion to remove excess primers after the PCR step that would otherwise result in saturation effects in the mass spectrometer. Subsequently, restriction fragment length polymorphism (RFLP) by a specific endonuclease is used

to determine the apo E genotype. DNA fragments from the digestion were studied by both gel electrophoresis and MALDI-MS. An unknown sample from an independent diagnostic laboratory was also successfully amplified and digested by PCR-RFLP and the resulting fragments were detected by mass spectrometry. Mass spectral data were confirmed later by gel electrophoresis performed by the diagnostic lab.

EXPERIMENTAL

Amplification of apo E sequences from genomic DNA

Genomic DNA was extracted from 1 mL whole blood using a commercial 'salt-extraction' kit (Puregene™, Gentra Systems Inc. Minneapolis, MN, USA). The final DNA preparations were dissolved in 0.33 mL of 10 mM Tris and 1 mM EDTA buffer (pH 8.0). DNA was amplified by PCR in a DNA thermocycler (M. J. Research, Watertown, MA, USA) using oligonucleotide primers that were synthesized at the DNA synthesis core facility at the University of Michigan. The primer design was based on those previously reported for the PCR amplification of the apo E gene.[12,14] The sequences of the primers as well as the amplified apo E

Table 1. The constant and polymorphic restriction sites for the Cfo I enzyme and the expected fragment lengths after digestion

ApoE allele	Restriction sequence and nucleotide position for Cfo I cleavage	Fragments, bp
ε2	GCGC at positions 38, 54, 145, 163	18, 38, 91, 101
ε2/ε3	GCGC at positions 38, 54, 145, 163, 211	18, 38, 48, 53, 91, 101
ε2/ε4	GCGC at positions 38, 54, 73, 145, 163, 211	18, 19, 38, 48, 53, 72, 91, 101
ε3	GCGC at positions 38, 54, 145, 163, 211	18, 38, 48, 53, 91
ε3/ε4	GCGC at positions 38, 54, 73, 145, 163, 211	18, 19, 38, 48, 53, 72, 91
ε4	GCGC at positions 38, 54, 73, 145, 163, 211	18, 19, 38, 48, 53, 72

gene are shown in Fig. 1. Each amplification reaction included ~400 ng of genomic DNA, 1X PCR buffer (supplied with Taq Polymerase), 15 mM MgCl₂, 200 nM each of sense and antisense primers, 10% DMSO, 100 μM of each dNTP and 5 units (5 U/μL) of Taq polymerase (GIBCO-BRL). Each reaction mixture was subjected to an initial denaturation step at 94 °C for 1 min, followed by 39 cycles of denaturation (45 s at 94 °C), primer annealing (30 s at 60 °C) an elongation (1 min at 70 °C). The amplified DNA were then subjected to a final extension at 72 °C for 5 min. The amplified sequence was purified from the interfering enzymes and buffers using Qiagen PCRquick Kit (Qiagen Inc., Chatsworth, CA, USA), prior to mass spectral analysis.

Exonuclease I digestion of primers

The amplified genomic DNA from various patient samples were set up for exonuclease I digestion in a thermocycler (M. J. Research, Watertown, MA, USA). The protocol included an initial incubation at 70 °C for 10 minutes to minimize dimerization of primers. Exonuclease I (4 U) was added to each reaction followed by incubation at 37 °C for 15 minutes. The enzyme was heat inactivated by incubating the mixture for an additional 10 minutes at 70 °C.

Restriction enzymatic digestion of amplified apo E

The amplified apo E sequences were then subjected to restriction enzymatic digestion by the addition of Cfo I directly to the PCR mixture in the presence of 1X restriction enzyme buffer. Complete digestion was achieved by incubating the mixture for ~6 hours at 37 °C. The cut sites for the Cfo I enzyme and the expected fragments are shown in Table 1. DNA fragments less than 100 bp in length could not be subjected to Qiaquick purification procedures. Instead, the restriction enzyme digests were purified by ethanol precipitation of the DNA. The ethanol precipitation protocol involved mixing the fragments with 0.5 volume of 7.5 M ammonium acetate and 3 volumes of ice cold ethanol. This mixture was allowed to stand at −75 °C overnight to increase the recovery of low mass DNA fragments. The DNA fragments were pelleted by spinning the precipitates at
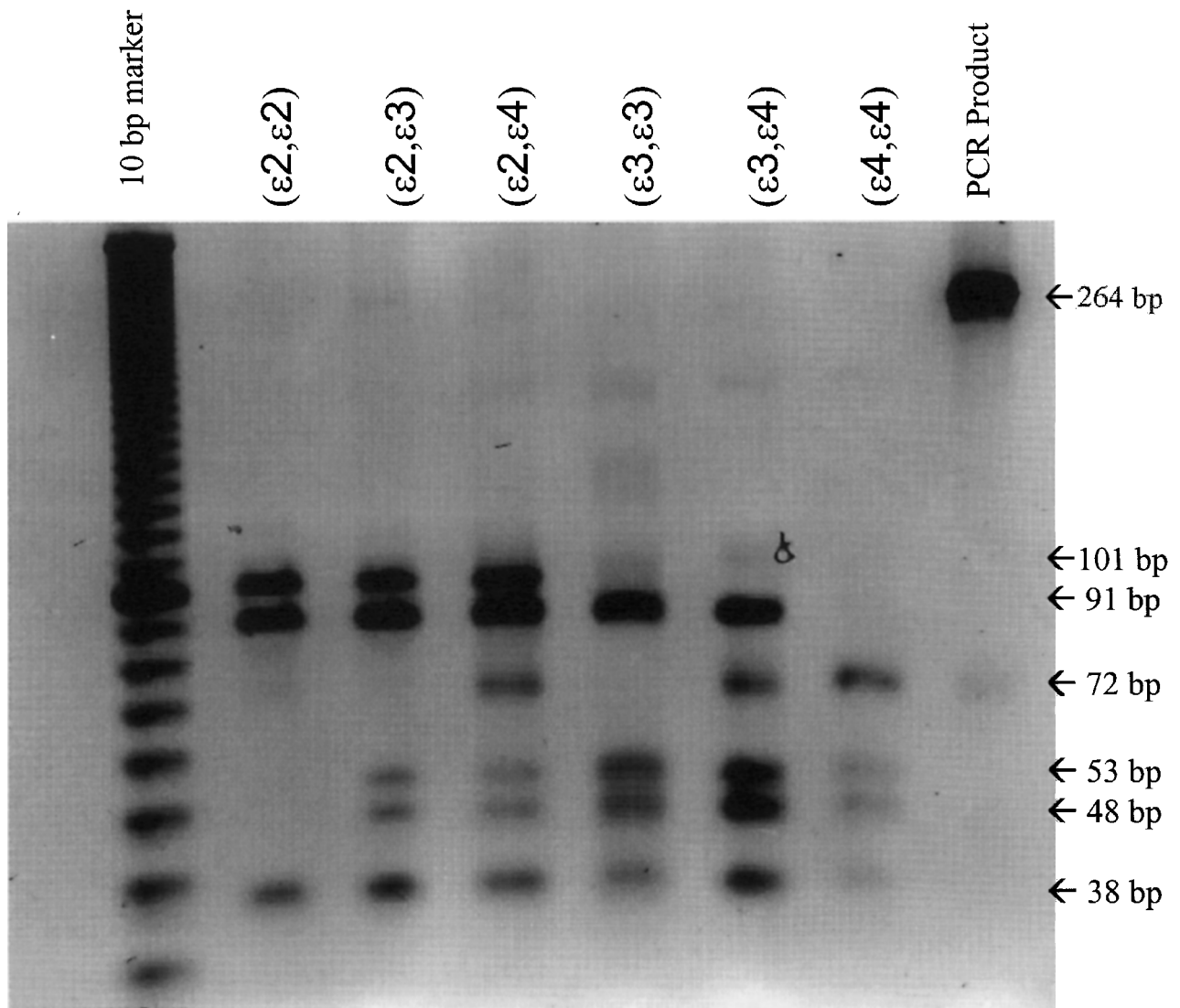

Figure 2. Gel electrophoresis result of the amplified apo E sequence and the Cfo I fragments. lane 1: molecular weight marker, 10 bp ladder, lane 2: ε2 isoform, lane 3: ε2/ε3 allele, lane 4: ε2/ε4 allele, lane 5: ε3 isoform, lane 6: ε3/ε4 allele, lane 7: ε4 isoform, and lane 8: 264 bp PCR product. The expected Cfo I restriction fragments are shown on the right.

~14 000 rpm for 30 min at 4 °C. The DNA pellets were then thoroughly washed with 0.5 volume of 70% ethanol and resuspended in DNAse free water (2–4 μL) for further MALDI-MS analysis.

Mass spectrometric analysis

All MALDI-MS spectra were collected on a linear time-of-flight mass spectrometer (R. M. Jordan Co., Grass Valley, CA, USA) as has been described previously.[24] Briefly, the MALDI ions were generated by a 355 nm focused output of a frequency tripled DCR-11 Nd:YAG laser (Spectra Physics, Mountainview, CA, USA). The ions formed were accelerated to 40 kV and current signals were detected by an 18 mm triple microchannel plate detector (Jordan) placed 1 m from the ion optics. Ion signals were amplified by the presence of a CuBe venetian blind conversion dynode (+15kV), but at the expense of resolution. The ion currents were measured by a LeCroy 9350M digital oscilloscope (LeCroy Corporation, Chestnut Ridge, NJ, USA) and processed on an IBM Pentium processor. An average of 30–50 mass spectra were recorded for each genotype studied. Pulsed delayed extraction (PDE) was not used for the spectra presented in this work. The resolution enhancement due to PDE was compromised by the loss of resolution due to the post acceleration stage, which was necessary for improved detection sensitivity. Since resolution is not an issue in this present methodology and higher sensitivity implied detection of the important restriction enzyme products, the post acceleration stage was used instead of the PDE.

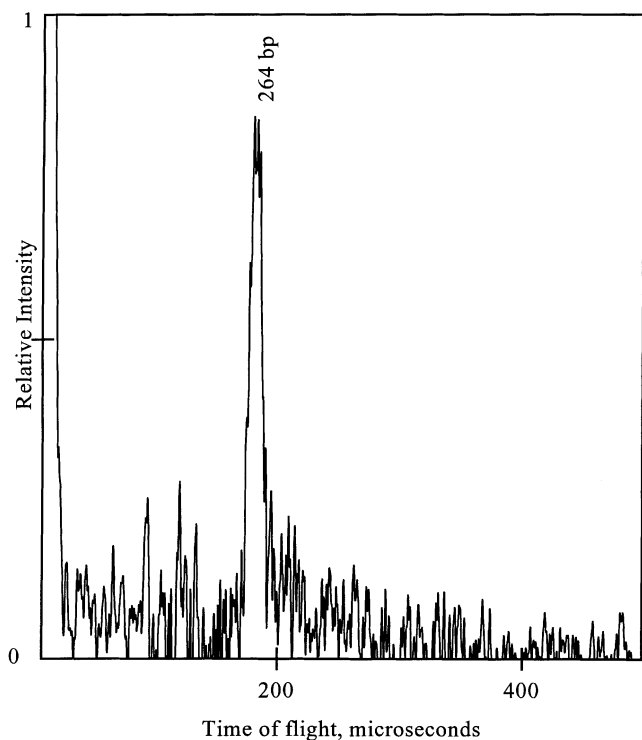The results presented herein used 2 μL of a 3-hydroxypicolinic acid and picolinic acid (3-HPA/PA) (4:1) matrix in 36% acetonitrile, as the matrix for desorption. The matrices were purchased from Sigma Chemical Co. (St. Louis, MO, USA) and used without further purification. The presence of nitrocellulose (Immobilon-LC Pure, Millipore, Bedford, MA, USA) for on-probe ion exchange purification was found to be essential in obtaining reproducible mass spectral data. Nitrocellulose (1–2 μL) was coated on to the stainless steel sample probe tip as a ~16 mg/mL solution in acetone. The precipitated DNA was redissolved in deionized water (2–4 μL), spotted onto the thin film and allowed to dry before the addition of the 3-HPA/PA matrix mixture. The entire mixture on the probe tip was allowed to dry at room temperature before being inserted into the mass spectrometer.
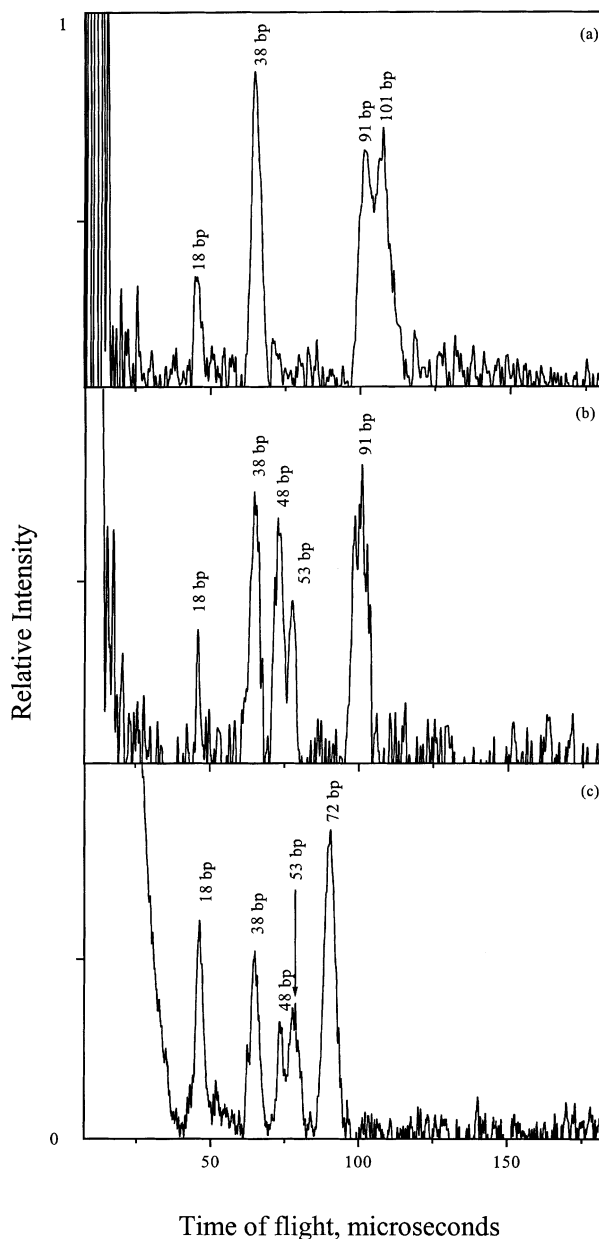

Figure 3. MALDI-MS of PCR amplified apo E from a 4:1 3-HPA/PA matrix. Nitrocellulose membrane was used for on-probe purification. The laser wavelength used was 355 nm.

Figure 4. MALDI-MS of Cfo I digested apo E (a) ε2, (b) ε3 and (c) ε4, isoforms using a 3-HPA/PA (4:1) matrix and 355 nm laser wavelength. The presence of 101 bp and 91 bp, the absence of 101 bp and 72 bp, and the presence of 72 bp are characteristic of the ε2, ε3 and ε4 alleles respectively.

RESULTS AND DISCUSSION

Analysis of amplified apo E gene

The genomic DNA samples from patients were amplified by polymerase chain reaction to a final length that was defined by the specific primers chosen. Figure 1 shows the PCR amplified sequence obtained for the different genotypes. A single prominent band at 264 bp was seen as the final product when detected by gel electrophoresis (Fig. 2). Figure 3 is the MALDI-MS observed for the PCR amplified DNA with the single stranded parent ion of the ds DNA arriving at 184 μs.

Restriction isotyping of amplified apo E

The restriction isotyping of apo E is accomplished by cleaving the PCR product with Cfo I in order to distinguish the six isoforms. Figure 1 shows the six cleavage sites in the amplified apo E sequence. The recognition sequence for the Cfo I enzyme is GCGC which is present at four constant and two polymorphic sites in the apo E sequence. The ε4 sequence is presented in the figure along with the nucleotide variations present in the ε2 and ε3 forms. The ε2 isoform differs from the ε4 due to a thymine substitution at nucleotide positions 71 and 211, thus eliminating the two enzyme cleavage sites (from GCGC to GTGC). The ε3 isoform produces five Cfo I cleavage sites by maintaining the GCGC sequence at position 211 but removing the polymorphic site at position 71. The resulting fragment patterns observed for the six isoforms of apo E are shown in Table 1. The gel electrophoresis results obtained for the Cfo I cleaved DNA samples representing the different homozygous and heterozygous combinations of apo E are shown in Fig. 2.

A primer digestion step was used before restriction enzyme digestion since the presence of one of the primers (46mer) caused occasional interference in detection of the higher mass fragments. In addition, the absence of the 46-mer primer simplifies the spectra in the region of the 48 bp diagnostic fragment. The primers were digested using Exonuclease I which is an exodeoxyribonuclease that digests single-stranded (ss) DNA in a 3′ to 5′ direction. This procedure was performed after the PCR step, but before the restriction enzyme was added.

Figure 4(a) is the MALDI mass spectrum obtained for an ε2 homozygote. The ε2 allele is characterized by the presence of the 91 bp and 101 bp fragments and the absence of the 48 bp, the 53 bp and the 72 bp fragments. The 18 bp and the 38 bp peaks are common to all the apo E allele combinations and are not generally detected by gel electrophoresis due to their small size.

The 91 bp and the 101 bp fragments result from cleavage at all four of the constant Cfo I sites but none of the polymorphic sites.

The ε3 allele is recognized by the presence of the 91 bp, 53 bp and 48 bp fragments (Fig. 4(b)). The 91 bp fragment reflects the absence of a cut site contributed by the mutation occurring at nucleotide position 71. The 48 bp and 53 bp fragments result from the polymorphic Cfo I site at position 211, thereby eliminating the 101 bp fragment observed in the ε2 allele.

In comparison, Fig. 4(c) represents the MALDI-MS of the PCR products from an ε4 homozygote. This pattern is characterized by the 72 bp fragment which results from the cleavage at nucleotide position 71. The other fragment generated due to the same mutation is 19 bp in length but is not typically used diagnostically due to its proximity to the 18 bp constant-site fragment. The ε4 allele also produces the 48 bp and the 53 bp fragments. The peak arriving at 53 μs belongs to one of the primers (25mer) and is most likely due to incomplete Exonuclease I digestion. The heterozygotic combinations that contained one of each allele resulted in three different combinations which were also studied using MALDI-MS. The results of the MALDI-MS typing of the ε2/ε3, ε2/ε4 and ε3/ε4 samples are shown in Fig. 5.

A test for the success of the MALDI-MS technique described above involved the determination of the apo E genotype of an unknown DNA sample. A sample of genomic DNA was obtained from the Molecular Diagnostics Laboratory, University of Michigan Medical Center, where apo E genotyping is performed on a routine basis by PCR-RFLP followed by gel electrophoresis detection.
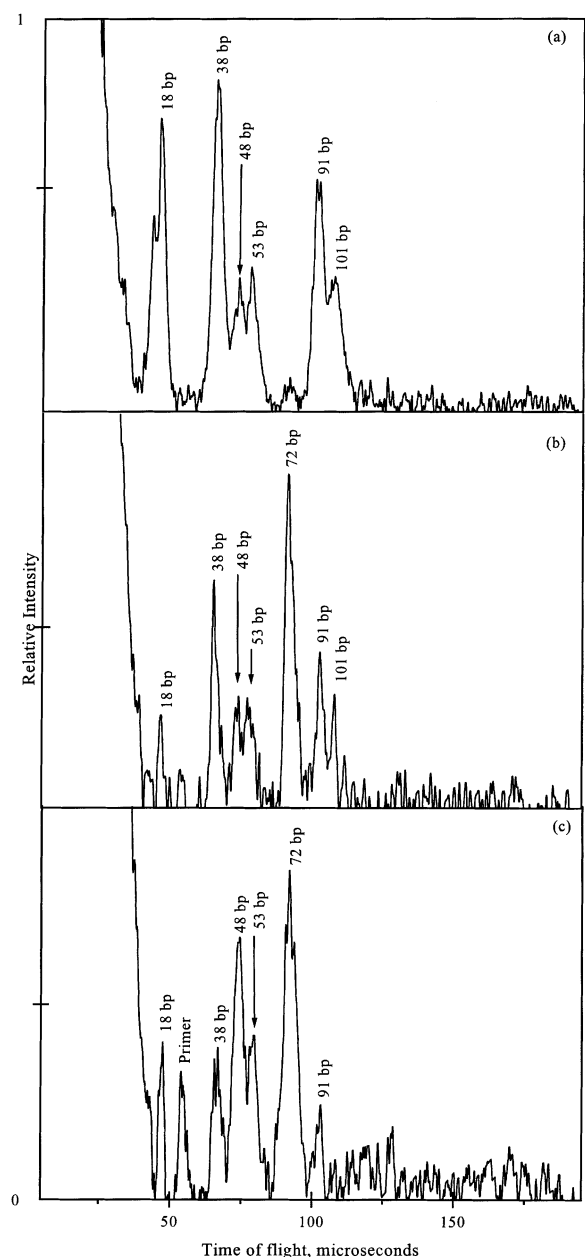

Figure 5. (a) MALDI-MS of Cfo I digested apo E (a)ε2/ε3, (b) ε2/ε4 and (c) ε3/ε4 alleles using a 3-HPA/PA (4:1) matrix and 355 nm laser wavelength.
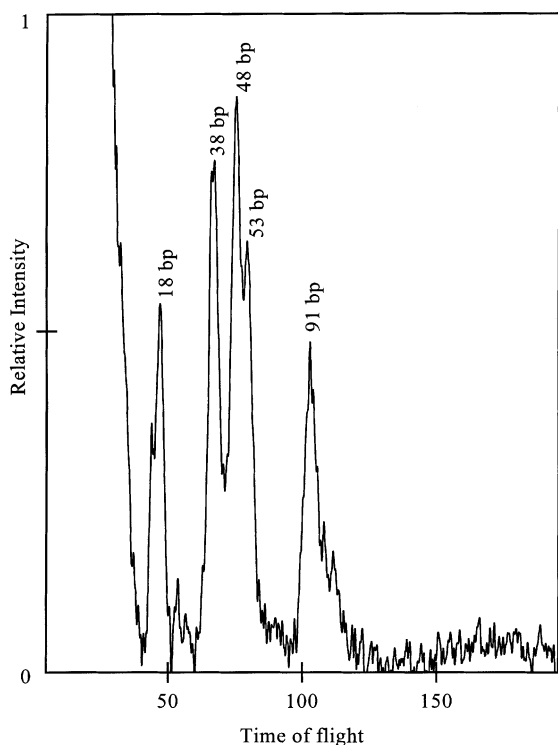
Figure 6. MALDI-MS of Cfo I digested unknown apo E sample using a 4:1 3-HPA/PA matrix and 355 nm laser wavelength. The presence of 91 bp, 53 bp and 48 bp are characteristic for the $\varepsilon 3$ allele.

Figure 6 is the MALDI-MS of the unknown DNA. The mass spectrum is characterized by the 91 bp, 53 bp and 48 bp fragments which clearly indicates that this patient is an $\varepsilon 3$ genotype. The mass spectral results were later confirmed with the gel electrophoresis data obtained by the Molecular Diagnostics Lab.

CONCLUSIONS

We have demonstrated the use of MALDI-MS as a clinical diagnostic tool for genotyping the apolipoprotein E gene. By replacing the standard gel electrophoresis technique for detection of the restriction isotyped fragments, we have shown the feasibility of MALDI-MS for studying double stranded DNA products generated for diagnostic purposes. The ability of MALDI-MS to genotype unknown DNA samples has also been demonstrated.

Acknowledgements

REFERENCES

1. R. W. Mahley and S. C. Rall Jr., *The Metabolic and Molecular Bases of Inherited Disease*, C. R. Scriver, A. L. Beaudet, W. S. Sly, *et al.* (Eds) McGraw-Hill, New York (1995), 1953.
2. J. Davignon, R. E. Gregg and C. F. Sing, *Arteriosclerosis*, **8,** 1–21 (1988).
3. A. D. Roses, W. J. Strittmatter, A. M. Saunders, D. E. Schmechel and M. A. Pericak-Vance, *Apolipoprotein E and Alzheimer's disease*, A. D. Roses, K. H. Weisgraber, and Y. Christen Springer-Verlag, pp. 1–10 (1996).
4. R. J. Havel, *Med. Clin. North Am.* **66,** 441–454 (1982).
5. G. Utermann, M. Hees and A. Steinmetz, *Nature* **269,** 604–607 (1977).
6. G. Utermann, A. Steinmetz and W. Weber, *Hum. Genet.* **60,** 344–351 (1982).
7. S. Visvikis, J. Steinmetz, I. Cuvelier, M. M. Galteau and G. Siest, *Progress recents en electrophorese bidimensionnelle*, M. M. Galteau and G. Siest (Eds), Presses Universitaires de Nancy, Nancy, France 159–166 (1986).
8. V. I. Zannis and J. L. Breslow, *Biochemistry* **20,** 1033–1041 (1981).
9. R. Aozaki, R. Kawaguchi, U. Ogasa, K. Hikiji, N. Kubo and I. Sokurabayashi, *Mol. Cell Probes* **8,** 51–54 (1994).
10. K. H. Weisgraber, Y. M. Newhouse and R. W. Mahley, *Biochem. Biophys. Res. Commun.* **157,** 1212–1217 (1988).
11. R. S. Holston, C. Snowden, F. Green, K. G. M. M. Alberti and S. E. Humphries, *Hum. Genet.* **83,** 364–368 (1989).
12. W. A. Reymer, B. E. Groenemeyer, R. van de Burg and J. J. P. Kastelein, *Clin. Chem.* **41,** 1046–1047 (1995).
13. J. E. Hixson and D. T. Vernier, *Arteriosclerosis* **31,** 545–548 (1990).
14. M. Emi, L. L. Wu, M. A. Robertson, R. L. Myers, R. A. Hegele, R. R. Williams, R. White and J.-M. Lalouel, *Genomics* **3,** 373–379 (1988).
15. 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,** 1144–1150 (1997).
16. Y.-H. Liu, J. Bai, Y. Zhu, X. Liang, D. Siemieniak, P. J. Venta and D. M. Lubman, *Rapid Commun. Mass Spectrom.* **9,** 735–743 (1995).
17. 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,** 772–774 (1995).
18. N. I. Taranenko, K. J. Matteson, C. N. Chung, Y. F. Zhu, L. Y. Chang, S. L. Allman, L. Haff, S. A. Martin and C. H. Chen, *Gen. Anal-Bio. Engg.* **13,** 87–94 (1996).
19. Y. Wada and M. Yamamato, *Rapid Commun. Mass Spectrom.* **11,** 1657–1660 (1997).
20. Y. Wada, *J. Mass Spectrom.* **33,** 187–192 (1998).
21. G. B. Hurst, M. J. Doktycz, A. A. Vass and M. Buchanan, *V. Rapid Commun. Mass Spectrom.* **10,** 377–382 (1996).
22. P. Jiang-Baucom, J. E. Girard, J. Butler and P. Belgrader, *Anal. Chem.* **69,** 4894–4898 (1997).
23. D. P. Little, A. Braun, B. Darnhofer-Demar and H. Koster, *Clin. Chem. Clin. Biochem.* **35,** 545–548 (1997).
24. J. Bai, Y. H. Liu, D. M. Lubman and D. Siemieniak, *Rapid Commun. Mass Spectrom.* **8,** 727 (1994).